Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 21 (2019), Issue 3 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoints

A Human(e) Factor in Clinical Decision Support Systems (e11732)
Tim Bezemer, Mark de Groot, Enja Blasse, Maarten ten Berg, Teus Kappen, Annelien Bredenoord, Wouter van Solinge, Imo Hoefer, Saskia Haitjema. . . . 6

Social Media Intervention Design: Applying an Affordances Framework (e11014)
Megan Moreno, Jonathan D'Angelo. . . . 15

The Rise of the Empowered Physician in the Digital Health Era: Viewpoint (e12490)
Bertalan Mesko, Zsuzsa Gy rffy. . . . 24

The Case for a Hippocratic Oath for Connected Medical Devices: Viewpoint (e12568)
Beau Woods, Andrea Coravos, Joshua Corman. . . . 723

Original Papers

Internet-Based Cognitive Behavioral Therapy in Stepped Care for Chronic Fatigue Syndrome: Randomized Noninferiority Trial (e11276)
Margreet Worm-Smeitink, Anthonie Janse, Arno van Dam, Andrea Evers, Rosalie van der Vaart, Michel Wensing, Hans Knoop. . . . 32

User-Centered Design of a Web-Based Tool to Support Management of Chemotherapy-Related Toxicities in Cancer Patients (e9958)
Rebecca Prince, Anthony Soung Yee, Laura Parente, Katherine Enright, Eva Grunfeld, Melanie Powis, Amna Husain, Sonal Gandhi, Monika Krzyzanowska. . . . 47

User Experience of an App-Based Treatment for Stress Urinary Incontinence: Qualitative Interview Study (e11296)
Ina Asklund, Eva Samuelsson, Katarina Hamberg, Göran Umefjord, Malin Sjöström. . . . 69

The Association Between Medication Adherence for Chronic Conditions and Digital Health Activity Tracking: Retrospective Analysis (e11486)
Tom Quisel, Luca Foschini, Susan Zbikowski, Jessie Juusola. . . . 80

Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention (e12441)
Kevin Wright, Carla Fisher, Camella Rising, Amelia Burke-Garcia, Dasha Afanaseva, Xiaomei Cai. . . . 90

XSL•FO
RenderX

Characterizing Media Content and Effects of Organ Donation on a Social Media Platform: Content Analysis (e13058)
Xiaoya Jiang, Wenshi Jiang, Jiawei Cai, Qingdong Su, Zhigang Zhou, Lingnan He, Kaisheng Lai. . . . 106

Health-Related Internet Usage and Design Feature Preference for E-Mental Health Programs Among Men and Women (e11224)
Rachel Smail-Crevier, Gabrielle Powers, Chelsea Noel, JianLi Wang. . . . 115

Do Search Engine Helpline Notices Aid in Preventing Suicide? Analysis of Archival Data (e12235)
Qijin Cheng, Elad Yom-Tov. . . . 127

Patient Rounds With Video-Consulted Relatives: Qualitative Study on Possibilities and Barriers From the Perspective of Healthcare Providers (e12584)
Christina Østervang, Lene Vestergaard, Karin Dieperink, Dorthe Danbjørg. . . . 136

Assessment of the Relationship Between Ambient Temperature and Home Blood Pressure in Patients From a Web-Based Synchronous Telehealth Care Program: Retrospective Study (e12369)
Ching-Chang Huang, Ying-Hsien Chen, Chi-Sheng Hung, Jen-Kuang Lee, Tse-Pin Hsu, Hui-Wen Wu, Pao-Yu Chuang, Ming-Fong Chen, Yi-Lwun Ho. . . . 147

Experiences With Video Consultations in Specialized Palliative Home-Care: Qualitative Study of Patient and Relative Perspectives (e10208)
Karen Funderskov, Mette Raunkiær, Dorthe Danbjørg, Ann-Dorthe Zwisler, Lene Munk, Mia Jess, Karin Dieperink. . . . 163

Use of Online or Paper Surveys by Australian Women: Longitudinal Study of Users, Devices, and Cohort Retention (e10672)
David Fitzgerald, Richard Hockey, Mark Jones, Gita Mishra, Michael Waller, Annette Dobson. . . . 173

A Web-Based Form With Interactive Charts Used to Collect and Analyze Data on Home Births in Italy (e10335)
Michele Zanetti, Rita Campi, Paola Olivieri, Marta Campiotti, Alice Faggianelli, Maurizio Bonati. . . . 184

Factors Influencing Web-Based Survey Response for a Longitudinal Cohort of Young Women Born Between 1989 and 1995 (e11286)
Deborah Loxton, Melissa Harris, Peta Forder, Jennifer Powers, Natalie Townsend, Julie Byles, Gita Mishra. . . . 195

Computer-Mediated Communication to Facilitate Synchronous Online Focus Group Discussions: Feasibility Study for Qualitative HIV Research Among Transgender Women Across the United States (e12569)
Andrea Wirtz, Erin Cooney, Aeysha Chaudhry, Sari Reisner, American Cohort To Study HIV Acquisition Among Transgender Women. . . . 204

Electronic Health Behaviors Among US Adults With Chronic Disease: Cross-Sectional Survey (e11240)
Lillian Madrigal, Cam Escoffery. . . . 214

Characterizing the Digital Health Citizen: Mixed-Methods Study Deriving a New Typology (e11279)
John Powell, Ulrike Deetjen. . . . 229

The Impact of Superfast Broadband, Tailored Booklets for Households, and Discussions With General Practitioners on Personal Electronic Health Readiness: Cluster Factorial Quasi-Randomized Control Trial (e11386)
Philip Abbott-Garner, Janet Richardson, Ray Jones. . . . 239

XSL•FO
RenderX

Cancer-Related Internet Use and Its Association With Patient Decision Making and Trust in Physicians Among Patients in an Early Drug Development Clinic: A Questionnaire-Based Cross-Sectional Observational Study (e10348)
Goldy George, Eucharia Iwuanyanwu, Adrianna Buford, Sarina Piha-Paul, Vivek Subbiah, Siqing Fu, Daniel Karp, Shubham Pant, Christina Hinojosa, Kenneth Hess, Charles Cleeland, Elmer Bernstam, Funda Meric-Bernstam, David Hong. 254

Peer Adoption and Development of Health Innovations by Patients: National Representative Study of 6204 Citizens (e11726)
Pedro Oliveira, Leid Zejnilovic, Salomé Azevedo, Ana Rodrigues, Helena Canhão. 264

Design and Preliminary Findings From a New Electronic Cohort Embedded in the Framingham Heart Study (e12143)
David McManus, Ludovic Trinquart, Emelia Benjamin, Emily Manders, Kelsey Fusco, Lindsey Jung, Nicole Spartano, Vik Kheterpal, Christopher Nowak, Mayank Sardana, Joanne Murabito. 275

Automated Analysis of Domestic Violence Police Reports to Explore Abuse Types and Victim Injuries: Text Mining Study (e13067)
George Karystianis, Armita Adily, Peter Schofield, David Greenberg, Louisa Jorm, Goran Nenadic, Tony Butler. 287

Patient Judgments About Hypertension Control: The Role of Variability, Trends, and Outliers in Visualized Blood Pressure Data (e11366)
Victoria Shaffer, Pete Wegier, KD Valentine, Jeffery Belden, Shannon Canfield, Sonal Patil, Mihail Popescu, Linsey Steege, Akshay Jain, Richelle Koopman. 299

The Role of Frontline Leaders in Building Health Professional Support for a New Patient Portal: Survey Study (e11413)
Sari Kujala, Iiris Hörhammer, Tarja Heponiemi, Kim Josefsson. 313

Patient Engagement as Measured by Inpatient Portal Use: Methodology for Log File Analysis (e10957)
Timothy Huerta, Naleef Fareed, Jennifer Hefner, Cynthia Sieck, Christine Swoboda, Robert Taylor, Ann McAlearney. 322

Disparities in Patient-Reported Interest in Web-Based Patient Portals: Survey at an Urban Academic Safety-Net Hospital (e11421)
Shobha Sadasivaiah, Courtney Lyles, Stephen Kiyoi, Piera Wong, Neda Ratanawongsa. 338

The Impact of a Web-Based Course Concerning Patient Education for Mental Health Care Professionals: Quasi-Experimental Study (e11198)
Anna Laine, Maritta Välimäki, Eliisa Löyttyniemi, Virve Pekurinen, Mauri Marttunen, Minna Anttila. 346

Evaluation of an E-Learning Training Program to Support Implementation of a Group-Based, Theory-Driven, Self-Management Intervention For Osteoarthritis and Low-Back Pain: Pre-Post Study (e11123)
Deirdre Hurley, Alison Keogh, Danielle Mc Ardle, Amanda Hall, Helen Richmond, Suzanne Guerin, Tara Magdalinski, James Matthews. 363

Clinical Virtual Simulation in Nursing Education: Randomized Controlled Trial (e11529)
José Padilha, Paulo Machado, Ana Ribeiro, José Ramos, Patrício Costa. 381

WeChat as a Platform for Problem-Based Learning in a Dental Practical Clerkship: Feasibility Study (e12127)
Wei Zhang, Zheng-Rong Li, Zhi Li. 390

Using Paid and Free Facebook Methods to Recruit Australian Parents to an Online Survey: An Evaluation (e11206)
Shannon Bennetts, Stacey Hokke, Sharinne Crawford, Naomi Hackworth, Liana Leach, Cattram Nguyen, Jan Nicholson, Amanda Cooklin. . .
4 0 2

XSL•FO
RenderX

Searching for Information on the Risks of Combined Hormonal Contraceptives on the Internet: A Qualitative Study Across Six European Countries (e10810)
Paula Gomes Alves, Irene Petersen, Fiona Stevenson. . . . 417

Exploring the Challenges of Implementing a Web-Based Telemonitoring Strategy for Teenagers With Inflammatory Bowel Disease: Empirical Case Study (e11761)
Alie Dijkstra, Anke Heida, Patrick van Rheenen. . . . 427

Exploring How Evidence is Used in Care Through an Organizational Ethnography of Two Teaching Hospitals (e10769)
Bryn Lander, Ellen Balka. . . . 440

Online Health Information Seeking and eHealth Literacy Among Patients Attending a Primary Care Clinic in Hong Kong: A Cross-Sectional Survey (e10831)
David Wong, Man-Kuen Cheung. . . . 449

Genomic Common Data Model for Seamless Interoperation of Biomedical Data in Clinical Practice: Retrospective Study (e13249)
Seo Shin, Seng You, Yu Park, Jin Roh, Jang-Hee Kim, Seokjin Haam, Christian Reich, Clair Blacketer, Dae-Soon Son, Seungbin Oh, Rae Park. . . . 462

The Effect of Online Effort and Reputation of Physicians on Patients' Choice: 3-Wave Data Analysis of China's Good Doctor Website (e10170)
Zhaohua Deng, Ziying Hong, Wei Zhang, Richard Evans, Yanyan Chen. . . . 475

Investigating the Effect of Paid and Free Feedback About Physicians' Telemedicine Services on Patients' and Physicians' Behaviors: Panel Data Analysis (e12156)
Hualong Yang, Xiaofei Zhang. . . . 489

Reporting of Patient Experience Data on Health Systems' Websites and Commercial Physician-Rating Websites: Mixed-Methods Analysis (e12007)
Tara Lagu, Caroline Norton, Lindsey Russo, Aruna Priya, Sarah Goff, Peter Lindenauer. . . . 501

Acceptability and Feasibility of Implementing Accelorometry-Based Activity Monitors and a Linked Web Portal in an Exercise Referral Scheme: Feasibility Randomized Controlled Trial (e12374)
Jemma Hawkins, Joanna Charles, Michelle Edwards, Britt Hallingberg, Linda McConnon, Rhiannon Edwards, Russell Jago, Mark Kelson, Kelly Morgan, Simon Murphy, Emily Oliver, Sharon Simpson, Graham Moore. . . . 524

Detecting Hypoglycemia Incidents Reported in Patients' Secure Messages: Using Cost-Sensitive Learning and Oversampling to Reduce Data Imbalance (e11990)
Jinying Chen, John Lalor, Weisong Liu, Emily Druhl, Edgard Granillo, Varsha Vimalananda, Hong Yu. . . . 542

Artificial Intelligence and the Future of Primary Care: Exploratory Qualitative Study of UK General Practitioners' Views (e12802)
Charlotte Blease, Ted Kaptchuk, Michael Bernstein, Kenneth Mandl, John Halamka, Catherine DesRoches. . . . 557

Physician Confidence in Artificial Intelligence: An Online Mobile Survey (e12422)
Songhee Oh, Jae Kim, Sung-Woo Choi, Hee Lee, Jungrak Hong, Soon Kwon. . . . 567

Design and Evaluation of Personalized Motivational Messages by a Virtual Agent that Assists in Post-Traumatic Stress Disorder Therapy (e9240)
Myrthe Tielman, Mark Neerincx, Willem-Paul Brinkman. . . . 580

XSL•FO
RenderX

Associations Between the Use of eHealth and Out-of-Hours Services in People With Type 1 Diabetes: Cross-Sectional Study (e13465)
Anne Hansen, Tor Claudi, Eirik Årsand. . . . 601

Sustainable Adoption of Digital Health Innovations: Perspectives From a Stakeholder Workshop (e11922)
Michelle Van Velthoven, Carlos Cordon. . . . 615

Improving the Quality of Consumer Health Information on Wikipedia: Case Series (e12450)
Shira Weiner, Jill Horbacewicz, Lane Rasberry, Yocheved Bensinger-Brody. . . . 623

Detecting the Interdisciplinary Nature and Topic Hotspots of Robotics in Surgery: Social Network Analysis and Bibliometric Study (e12625)
Lining Shen, Shimin Wang, Wei Dai, Zhiguo Zhang. . . . 631

Offline Digital Education for Medical Students: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration (e13165)
Bhone Kyaw, Pawel Posadzki, Gerard Dunleavy, Monika Semwal, Ushashree Divakar, Vasilis Hervatis, Lorainne Tudor Car. . . . 687

Reviews

eHealth Apps Replacing or Complementing Health Care Contacts: Scoping Review on Adverse Effects (e10736)
Wilhelmina Stevens, Rob van der Sande, Lilian Beijer, Maarten Gerritsen, Willem Assendelft. . . . 60

Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis (e12053)
Amelia Romeo, Sarah Edney, Ronald Plotnikoff, Rachel Curtis, Jillian Ryan, Ilea Sanders, Alyson Crozier, Carol Maher. . . . 510

Medical Doctors' Offline Computer-Assisted Digital Education: Systematic Review by the Digital Health Education Collaboration (e12998)
Hayfaa Wahabi, Samia Esmaeil, Khawater Bahkali, Maher Titi, Yasser Amer, Amel Fayed, Amr Jamal, Nasriah Zakaria, Amna Siddiqui, Monika Semwal, Lorainne Car, Paul Posadzki, Josip Car. . . . 648

Digital Education for Health Professions on Smoking Cessation Management: Systematic Review by the Digital Health Education Collaboration (e13000)
Monika Semwal, Penny Whiting, Ram Bajpai, Shweta Bajpai, Bhone Kyaw, Lorainne Tudor Car. . . . 662

Health Professions' Digital Education: Review of Learning Theories in Randomized Controlled Trials by the Digital Health Education Collaboration (e12912)
Shweta Bajpai, Monika Semwal, Ram Bajpai, Josip Car, Andy Ho. . . . 674

Serious Gaming and Gamification Education in Health Professions: Systematic Review (e12994)
Sarah Gentry, Andrea Gauthier, Beatrice L'Estrade Ehrstrom, David Wortley, Anneliese Lilienthal, Lorainne Tudor Car, Shoko Dauwels-Okutsu, Charoula Nikolaou, Nabil Zary, James Campbell, Josip Car. . . . 703

Corrigenda and Addenda

Author Contribution Correction: An Integrated Influenza Surveillance Framework Based on National Influenza-Like Illness Incidence and Multiple Hospital Electronic Medical Records for Early Prediction of Influenza Epidemics: Design and Evaluation (e13699)
Cheng-Yi Yang, Ray-Jade Chen, Wan-Lin Chou, Yuarn-Jang Lee, Yu-Sheng Lo. . . . 729

XSL•FO
RenderX

Viewpoint

A Human(e) Factor in Clinical Decision Support Systems

Tim Bezemer[1], MSc; Mark CH de Groot[1], PhD; Enja Blasse[1], BSc; Maarten J ten Berg[1], PhD; Teus H Kappen[2], MD, PhD; Annelien L Bredenoord[3], PhD; Wouter W van Solinge[1], PhD; Imo E Hoefer[1], MD, PhD; Saskia Haitjema[1], MD, PhD

[1]Laboratory of Clinical Chemistry and Haematology, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[2]Department of Anesthesiology, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[3]Department of Medical Humanities, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

Corresponding Author:
Saskia Haitjema, MD, PhD
Laboratory of Clinical Chemistry and Haematology
University Medical Center Utrecht
Room G03.550
Heidelberglaan 100
Utrecht, 3584 CX
Netherlands
Phone: +31 88 755 5555
Email: S.Haitjema@umcutrecht.nl

Abstract

The overwhelming amount, production speed, multidimensionality, and potential value of data currently available—often simplified and referred to as *big data* —exceed the limits of understanding of the human brain. At the same time, developments in data analytics and computational power provide the opportunity to obtain new insights and transfer data-provided added value to clinical practice in real time. What is the role of the health care professional in collaboration with the data scientist in the changing landscape of modern care? We discuss how health care professionals should provide expert knowledge in each of the stages of clinical decision support design: data level, algorithm level, and decision support level. Including various ethical considerations, we advocate for health care professionals to responsibly initiate and guide interprofessional teams, including patients, and embrace novel analytic technologies to translate *big data* into patient benefit driven by human(e) values.

(J Med Internet Res 2019;21(3):e11732) doi:10.2196/11732

KEYWORDS

clinical decision support; big data; artificial intelligence; machine learning; deep learning; precision medicine; expert systems; data science; health care providers

Introduction

Although medical data collection and interpretation used to be the domain of health care professionals, the broad availability of health data in unprecedented amounts has significantly and irrevocably changed the landscape of modern care. Even patients now start to collect their own health data using, for instance, smart watches or apps, which may become an important source of health data in the future.

The craft of translating information into the right diagnosis and corresponding treatment is daily routine for health care professionals. It entails collecting the relevant data for each individual patient, integrating this information with pre-existing knowledge, drawing a conclusion, and initiating appropriate treatment in dialogue with the patient. A significant portion of medical training is dedicated to learning how to distinguish relevant from irrelevant information to ultimately make the best decision possible. Yet, the overwhelming amount, production speed, multidimensionality, and potential value of data currently available (often simplified and referred to as *big data*) exceed the limits of understanding of the human brain.

Conversely, developments in data analytics and computational power provide the opportunity to obtain new insights and transfer data-provided added value to clinical practice in real time. Such systems are called clinical decision support (CDS) and can broadly be defined as “information systems designed to aid in the clinical decision-making process, by integrating different sources of health information such as Electronic Health Records, laboratory test results, etc” [1]. CDS systems come in many forms and functions, but all share the aim of generating clinically relevant outcomes based on input data. A decision

XSL•FO
RenderX

can be supported by a rule or a model as simple as an if-then rule (eg, built-in reference values for laboratory measurements) or a complex prediction model (eg, artificial intelligence [AI] pointing radiologists to possible incidental findings). The corresponding output of a CDS system varies from showing the generated prediction as input for a clinical decision (eg, automatically generated early warning scores) to acting upon the decision without human interference (eg, an implantable cardioverter defibrillator).

Recent reports on CDS systems in radiology and pathology are promising. Computers can, for example, support radiologists in interpreting mammograms or help pathologists in the classification of brain tumors [2,3]. Google recently also received the Food and Drug Administration's approval for the introduction of a diabetic retinopathy algorithm based on retinal imaging [4]. Moreover, development of complex algorithms now starts to transcend beyond imaging specialties [5].

If the computer seems to know better anyhow, should we fully abolish medical curricula and focus on data scientists who develop CDS systems, with lay people gathering the information required for them, entitling the computer to do the interpretation instead? Probably not. There are at least two human beings present when a medical decision is made: a patient and a health care professional. Recently, the role of the patient as a disease *experience* expert [6] and his or her role in shared decision making have come into awareness. Here, we focus on the role of health care professionals and their expert knowledge. Throughout this paper, we will briefly touch upon various ethical issues. However, we strongly feel ethical considerations pertaining to algorithmic decision making deserve a discussion of their own, and kindly refer the reader to a recent overview on this topic [7]. In this paper, we show that a well-designed CDS system needs expert knowledge of health care professionals in all 3 phases of development: data, algorithm, and decision support (Table 1). Moreover, in the era of CDS, we advocate for health care professionals to responsibly initiate and guide interprofessional teams, including patients, and embrace novel analytic technologies to translate *big data* into patient benefit driven by human(e) values.

Table 1. This table shows the 3 levels in the building process of a clinical decision support system and some examples of where clinical expert knowledge of health care professionals plays a role in each of these levels.

Level and example of issue	Example of expert knowledge
Data level	
Laboratory thresholds	Hemoglobin reference range to diagnose anemia
Derived measurements[a]	Body mass index
Diagnostic codes	Grouping of related diagnoses in a study population
Jargon	Same abbreviations having different meanings
Temporality	Glucose values are highly dependent on the time of day (eg, pre- or postprandial)
Algorithm level	
Methodological choices	How to handle missing data (eg, missing not at random)
Feature engineering[a]	Constructing relevant derived variables from raw data (eg, torsades de pointes, Wolff-Parkinson-White syndrome)
Artifacts	For example, oxygen saturation of zero caused by a slipping pulse oximeter, switched leads in an electrocardiogram
Decision support level	
Interpretation of model output	Risk probability of 0.75 requires a warning (*amber light*) in a CDS[b] system
Degree of autonomy	Tuning of implantable cardioverter defibrillator
Knowledge on *usefulness*	Weighing a CDS system's advice to treat while considering quality of life versus treatment burden in elderly cancer patients in a shared decision-making context

[a]Derived measurements may occur at the data level but also at the algorithm level; the former being undesirable because any manipulation at the data level may result in a loss of information.

[b]CDS: clinical decision support.

XSL•FO
RenderX

Table 2. Table comparing different types of clinical data on some points important to clinical decision support systems.

Clinical decision support issues	Electronic health record free-text/unstructured data (eg, clinical notes)	Registry/trial data (eg, case record forms case record forms and questionnaires)	Structured data/electronic health record (eg, lab values and smoking status)
Context completeness	Excellent: contextual information can be included.	Poor: context is essentially absent as a priori interpretation is an integral part of recording data in case record forms.	Depends on implementation. Context may be lost because of predetermined categorization.
Machine readability	Poor: information is mostly useful for case-specific usage by humans. May require text mining/text retrieval to convert to a machine-readable format.	Good: data are uniformly formatted and can be parsed by computers.	Excellent: data can be parsed or directly used by computers.
Translatability (between institutions)	Poor: free text contains jargon-specific, ambiguous abbreviations (eg, PCI: percutaneous coronary intervention/prophylactic cranial irradiation).	Excellent: trial data are usually collected using a standardized protocol, allowing for interoperability between institutions.	Good: lab values can be converted using reference values. Structured data, such as smoking and hypertensive status, can be reformatted for interoperability.
Noise resistance	Very poor: These type of data are very sensitive to *interobserver* noise (eg, personal abbreviations, spelling mistakes, and personal focus in recording certain types of information).	Excellent: data are recorded in a standardized way, designed to prevent noise.	Good: data are often machine-derived or recorded in a standardized way. However, bias because of differences in information-recording habits among health care professionals may arise.
Availability for reuse/general applicability	Excellent: these type of data are readily available, contain a lot of context (see Context completeness), and can thus be repurposed for a variety of applications.	Limited: trials are designed and conducted for one specific research question.	Excellent: these type of data are readily available and can thus be used for a plethora of purposes.
Design flexibility	Excellent: study design can be revisited if unanticipated bias effects arise. In this sense, bias could be *corrected* by altering the data selection.	Poor: study design is *hit-or-miss*. Bias cannot be *corrected* after the data recording process.	Excellent: study design can be revisited if unanticipated bias effects arise. In this sense, bias could be *corrected* by altering the data selection.

Data

Developing a CDS system starts with data. Data come in many forms and sets (Table 2). Structured data such as numeric data (eg, laboratory measurements and blood pressure) or categorical data (eg, hypertension yes/no or educational level) are easiest to work with in a model. This is the first point at which expert knowledge of health care professionals may enter CDS development process. However, a substantial part of day-to-day clinical decisions is based on unstructured free-text entries, encompassing, for example, patient history and physical examination observations by doctors or regular notes from nurses. Although discouraged in modern electronic health record (EHR) systems, unstructured free-text clinical notes still provide irreplaceable information and context to health care professionals. Using free text introduces a number of challenges. Aside from the obvious ones, such as writing style and typos, medical text is incredibly site specific and can be highly biased. This phenomenon ranges from language- and country-specific abbreviations to jargon differences between 2 wards within the same hospital (eg, AF for atrial flutter and amniotic fluid or MS for mitral stenosis and multiple sclerosis). This is an understandable effect of rapid communication between health care professionals or of health care professionals taking personal notes to capture their train of thought. However, this leads to a given phrase, term, acronym, or abbreviation being context specific and having different meanings in different situations. Free-text interpretation, therefore, heavily depends on contextual expert knowledge.

Data Sources

Widely used datasets for CDS systems include clinical trials and medical registries. Data collected within trials are of importance for a predefined research question. They are usually of high quality, may be stored in great detail, and are often richly annotated with expert knowledge (diagnostic codes and predetermined disease severity classifications). Medical registries are developed for quality control and research purposes. They are used to record a predefined limited number of variables for a specific group of patients, often focusing on particular conditions and diagnoses. Careful maintenance of research databases and registries allows for the collection of data from patients in a clean and systematic way according to protocol, preventing missingness and loss to follow-up as much as possible. However, because of their restrictive nature, research datasets and medical registries discard valuable contextual information, such as free-text notes, about included patients. Therefore, they show a limited, predefined scope of the patient's condition. Furthermore, women and minorities are underrepresented in research datasets, and patients who are included can suffer from the Hawthorne effect (ie, altered behavior because of the fact that one is a study subject) [8]. The concept of research datasets and medical registries does not allow for flexibility in study design; the decision on what

XSL•FO
RenderX

information to collect (and in what way) is single and final. Moreover, information beyond the scope cannot be added without considerable effort at a later moment (if anonymization or informed consent regulations do not prevent this at all).

Due to this rather *artificial* way of collecting data as compared with clinical care, research databases and registries are unsuited for the creation of broadly applicable CDS systems using increasingly complex models. Moreover, CDS systems preferably apply information that is already available to the care provider to aid in the clinical process without impeding it by requiring the collection of various additional data. Data from EHRs contain real-world data from clinical practice. EHR-based datasets are, therefore, more suitable for CDS system development. At the same time, EHR systems were designed as a virtual patient chart and not necessarily for reuse of the data they capture. As such, turning them into valuable EHR-based datasets takes careful and skilled data processing. For example, EHR data require more data cleaning (eg, how to handle *not at random* missing data—also a prime example of where clinical expert knowledge plays a vital role), careful assessment of informed presence bias (ie, acknowledgment of the bias introduced by the medical process), and decisions are to be made about how certain variables are derived from often unstructured data such as free text in EHR systems or clinical notes (eg, define diabetes mellitus and define hypertension) [9]. As only health care professionals themselves know about these inherent biases of working in an EHR, expert knowledge is indispensable. Table 2 compares different types of clinical data on a number of points important to CDS development. The Utrecht Patient Oriented Database in the University Medical Center Utrecht, the Netherlands, is an example of a routinely updated EHR-based database, containing data from multiple hospital sources of about 2.3 million patients (Multimedia Appendix 1, [10]). Utrecht Patient Oriented Database is curated by clinicians who use their expert knowledge in the design of the database to counter the known biases that are inherent to EHR data. Furthermore, they assist their clinical colleagues in transforming relevant data into meaningful variables to answer clinically relevant research questions and to develop CDS systems.

Data Preprocessing

Before the data can be used to build a model, they need to be preprocessed. Preprocessing steps define variables from raw data that a model can use. During preprocessing steps, the expert knowledge of health care professionals is important to derive meaningful variables and values from the data. For example, disease activity variables need to be constructed because research guidelines and accompanying questionnaires are not regularly applied in clinical care. Furthermore, health care professionals may direct data scientists away from composite endpoints (eg, a patient has a 50% increased risk of pneumonia, pulmonary embolism, or chronic obstructive pulmonary disease) as they are less useful for CDS than specific endpoints that require specific actions (eg, a 50% increased risk of pneumococcal pneumonia). Moreover, the extraction of features from the data, such as differences in laboratory values over time, requires expert knowledge to determine appropriate time windows. Although the accuracy of algorithms generally increases if missing values and outliers are removed, the absence of data can carry value that only a health care professional is able to acknowledge, and the same holds true for outliers.

Algorithm

After the selection of the right data to develop the CDS system, the next phase is to develop a model (ie, the *recipe* that describes the relationship between variables and outcome in the data) by using an algorithm (a predetermined computational method to derive such a *recipe* from the data). Depending on the complexity of the modeling task, model development usually contains a phase of model training and phase of model validation. In the training phase, a model that best fits the data (ie, makes the best predictions on the training data) is developed, and in the validation phase, tests are carried out to check whether the model is correct (ie, generalizes to the population). What constitutes a *good* prediction is dependent on the (clinical) research question (ie, identify all positive diagnoses at any cost or find a trade-off between cost and efficacy). It is common practice to test the model on a new dataset in the validation phase. This can be a previously unseen part of the total dataset or an entirely new dataset. Although modeling and algorithm development are not the natural habitat for most health care professionals, their knowledge and input are invaluable in this phase.

Simple Models

In simple models, the input of expert knowledge of health care professionals is well established. As mentioned before, the simplest form of decision models is *if-then* rules. Examples of such models include laboratory reference values based on statistical distributions of patient measurements (eg, if fasting glucose >11 mmol/L, then the patient probably has diabetes mellitus), medical risk scores (eg, if Glasgow Coma Scale is lower than 9, then consider intubating the patient), rule-based warnings for medication (eg, if the patient has impaired kidney function, then do not allow prescription for metformin), and alarms on the intensive care when vital sign thresholds are violated (eg, sound an alarm if saturation levels drop below 95%) [11]. When building these simple models into CDS systems, the thresholds and reference values need to be provided by health care professionals.

Complex Models

These traditional models and clinical scores are generally straightforward (Apgar score and Glasgow coma scale) to make them easily actionable, even in stressful situations. The beauty of their simplicity has ensured their broad application, but their sensitivity and specificity are unavoidably limited and usually include a substantial gray area. Moreover, most current models are based on regression or correlation measures that are less able to capture complex relationships in the data. The availability of machine learning offers novel approaches for developing medical models and risk scores. Machine learning refers to a group of statistical techniques that can be used to discern even complex patterns or regularities in data. They do so through an iterative process (in other words, the patterns are learned, hence machine *learning*) and produce a prediction model based on the learned patterns, which can then be

XSL•FO
RenderX

incorporated in clinical support tools. [12,13]. In this complex type of modeling, input of expert knowledge from health care professionals may seem less obvious. However, model development is not a neutral process and even the values of health care professionals may be of additional benefit [7].

Machine learning algorithms can be roughly divided into 2 classes: (1) supervised learning algorithms that make use of prior (expert) knowledge about outcomes to guide the process and (2) unsupervised learning algorithms that aim to discover data patterns irrespective of model outcome.

Input data for supervised learning algorithms need to be labeled and selected manually (eg, positive/negative diagnosis, benign/malign, and concentration of inflammatory marker X) before modeling, and these data then constitute the outcome variable to predict for new cases. In other words, supervised learning systems rely heavily on expert knowledge [12,14]. *Supervision* is not only needed for the prelabeling of cases and noncases but also for statistical and methodological choices. Such choices include, for example, the choice of which algorithm to use and whether to normalize/standardize the data, and more algorithm-specific choices, such as the number of layers and nodes in a neural network or number of splits in a decision tree. Making appropriate choices on these aspects requires input from data scientists and medical scientists alike and will significantly affect the validity of the model. When the input variables in supervised models are selected by health care professionals and are based on prior knowledge and scientific evidence, supervised machine learning models may provide a safe ground for decision tools.

Unsupervised learning algorithms aim to uncover regularities in data without being guided by a prelabeling of the data (ie, clustering algorithms). The scope of this technique is often to discover novel subgroups within data and populations [12,14]. This approach is useful when information on the characteristics needed to discriminate between patients and controls is not yet available, or when one aims to find starting points for more fundamental scientific research. Therefore, this approach is usually used to find novel patterns in the data instead of making predictions and is thus generally more exploratory in nature. An advantage is, thus, that it allows for hypothesis-free or agnostic detection of patterns even when expert knowledge on the difference between subgroups is missing. Nevertheless, unsupervised systems can still profit from expert knowledge in the modeling process, as for example, clustering algorithms often require the user to preset the desired number of clusters, a decision that may be based on evidence of a known pattern in the population of interest.

A group of supervised and unsupervised techniques that is currently the state-of-the-art in machine learning is called deep learning. These techniques often involve artificial neural networks and attempt to learn increasingly *deep* representations of associations in the data. Deep neural networks (*deep nets*) are capable of automatically determining how to represent the input data in the best way for the question at hand. Theoretically, deep nets do not even require manual data preparation. In essence, the only requirements are to standardize numeric data and to encode categorical data into a numerical format interpretable by the algorithm. Deep learning is often used to recognize patterns in complex datasets that can subsequently be used by (supervised) machine learning models, for example, using clusters of a disease as outcome variables. In some experiments, deep learning methods have been shown to have superior prediction accuracy compared with other methods [15]. With EHRs as input, deep learning can improve prediction performance in modeling tasks that can be used for CDS [16]. Deep nets have, for example, already been proven useful in the computer-aided assessment and interpretation of medical images [4]. However, the rule of thumb *garbage in is garbage out* also applies when it comes to deep learning. If death is included in a model to predict readmission, it will probably come up as the most predictive variable, yet this might not be the actionable insight a health care professional is looking for. Guidance of health care professionals is needed when applying machine learning in the medical field.

Modeling Characteristics

Although all models remain specific for a given question (eg, what decision to support), building a CDS system is not a static process. It often includes rounds of major and minor changes of variables included and algorithmic fine-tuning. Moreover, some algorithms are never really finished. These algorithms are called *self-learning* and are designed to incorporate newly acquired data over time into their modeling processes. One of the reasons one may want to adapt an algorithm is spurious associations. Agnostic analyses that do not incorporate current understanding are prone to implement clinically irrelevant or even false associations with potentially deleterious consequences [17]. Outcomes of CDS models, therefore, need critical appraisal from experts regularly [17,18].

As health care professionals are responsible for the decisions they make, they highly value transparency of a model's decision process and its development [19]. Whether or not the respective variables are shared with the health care professionals can be up for debate, as sharing of certain variables may lead to undesirable side effects. For example, an algorithm that states that a certain keyword in a patient history carries value as a warning, one may not want the health care professional to know this keyword to prevent it from being stated just to indicate a warning. A CDS system can be transparent to a greater or lesser extent. A CDS system that contains an algorithm that is too complex to comprehend can result in a so-called black box situation, where it is difficult or even impossible for a human brain to understand how the prediction model works. This renders validation of these black box algorithms extremely important. Unfortunately, there is a trade-off between attainable model complexity and model interpretability [20]. The opportunities that more complex models may provide should not be underestimated. To fully benefit from complex data and incorporate it into clinical practice, health care professionals may need to accept that the ultimate goal of thoroughly clinically validated predictive models in CDS systems may not be to be fully and completely interpretable or transparent but rather to be useful to a clinical purpose and influence patient outcomes. The process of model *development*, including choices that are made based on expert knowledge of both health care professional and data scientist, should always be as transparent

XSL•FO
RenderX

as possible for all parties involved. Transparency of the development process may ultimately be critical for acceptance of CDS systems in clinical practice.

Moreover, depending on the complexity of the algorithm, internal (ie, on the same data) and external (ie, on other data) validation steps are vital before patients are exposed to the output. To what extent CDS systems need to be assessed as medical devices, according to their intended use, is still up for debate [21,22]. If-then CDS systems and CDS systems written to combine data into a visually attractive interface may be excluded from extensive clinical research but still need quality checks, regular revisiting of the algorithm, and piloting in clinical practice to ensure the right information is displayed for the right patient (does the algorithm take the most recent value from the table and did column names change). Furthermore, scientific evidence on validity and added value of the CDS system likely increases its use by health care professionals. Fortunately, such research is currently gaining traction in the medical community [23].

Decision Support

Implementation and use of a CDS system consist of multiple steps, including presenting the algorithm output in a specific way, interpretation by the health care professional, and eventually, the medical decision that is made. A CDS system is not a bare model producing just an output (eg, *65%*). Almost always, it contains some level of interpretation. For example, laboratory measurements are often displayed in black, red, or blue to indicate whether they fall within or are higher/lower than a reference range. Risk percentages may be accompanied by a traffic light coloring scheme, indicating risk compared with a standard disease course. These manners of presentation (eg, how to report variables and what kind of user interface) are probably the most intuitive place to integrate the expert knowledge of a health care professional. Indeed, health care professionals and patients are often included in the user experience or user interface design phases to discuss implementation. However, this is frequently perceived too late in the CDS development process, and therefore, may yield an opposite effect.

Furthermore, model results have to be interpreted in a specific medical context before the CDS system can provide the actual tailored CDS and lead to action. This step is usually supervised by health care professionals. For instance, a cardiologist will double check the automated interpretation of the electrocardiogram (ECG) machine, and it is very likely that an eye specialist will supervise Google's new diabetic retinopathy algorithm before any treatment is started [4]. Supervision of interpretation does not necessarily need a thorough understanding of the algorithm itself. Rather, it is the human intervention of integrating the contextual knowledge of the health care professional and, even more importantly, the patient's wishes, before acting upon the algorithm's output.

Most CDS systems do not (yet) act autonomously, so they need the attention of a health care professional to be effective. The highly technologically supported intensive care units provide ample evidence that getting the attention of a health care professional can be a challenge, as too many alarms can lead to alarm fatigue. Tweaking alert settings on an individual patient basis to make them meaningful as opposed to being perceived as a nuisance has been shown to improve compliance in critical care [11]. Incorporation of the multitude of emerging CDS systems into clinical practice needs streamlining and thorough knowledge of the medical process. Health care professionals should take initiative to lead interprofessional teams, deciding how and when to report CDS.

Unfortunately, systematic scientific evidence outlining what requirements a useful CDS system should meet is missing [24] and has been replaced by more or less anecdotic or empirical recommendations for many years. The *Ten Commandments of Clinical Decision Support* [25] lists factors as speed, anticipation of information need, integration into the workflow, or general ease-of-use type of advice in alerts. Moreover, negative advice, for example, an advice not to perform or order a particular test, is rarely accepted when no alternative is suggested, and the method of alert presentation has been found to be crucial to alert compliance [26].

Although such experience-based recommendations remain useful, the most important evidence for the usefulness of a CDS system that will influence acceptation by health care professionals will be its ability to influence clinical outcomes. Unfortunately, the evidence for CDS systems' frequently purported advantages over *old-fashioned* clinical decision in improving clinical outcome, workload, and economic cost is scarce. However, CDS systems have been shown to improve health care processes and are the best way to decrease unnecessary clinical testing [27].

Acceptation of CDS by health care professionals depends on the degree to which they feel autonomous in their decision making. Rather than choosing colors for a user interface, being part of the development process, identifying the appropriate data, discussing model design, and validation may help health care professionals to feel in control in the midst of forces that are transforming daily clinical practice. A supportive organization with inspiring leadership encouraging involvement of health care professionals in the development of CDS systems stimulates this transformation.

Discussion

With the rise of machine learning, and especially deep learning in CDS systems, it is perhaps tempting to let IT and data experts build CDS systems, redirecting health care professionals to merely gathering data. However, we have shown that human(e) health care professionals are still of paramount importance, as all phases of development and use of a CDS system requires the extensive expert knowledge of health care professionals. Health care professionals should not just be involved in implementing the CDS system into clinical practice but should be part of an interprofessional CDS development team from the start, initiating and guiding development through clinical demand and expert knowledge. They bring in the clinical decision they want the CDS system to support and help to understand the context in which variables are collected during routine care. Their interpretation is vital in extracting relevant variables from raw data and in avoiding the implementation of

XSL•FO
RenderX

spurious associations in CDS systems. Moreover, as health care professionals want the best for their patients, they may even develop a sense of moral obligation to embrace strategies that unravel data complexities beyond their comprehension, as relying on methods that do not use the data to their full potential leads to potentially unused value for their patients. At the same time, both data scientists and health care professionals should be alert to cognitive biases provided by pre-existing expert knowledge. Indeed, interprofessional CDS development teams should be as inclusive as possible, as the values and preferences of the people involved influence the underlying model [28].

The hallmark of biology, variability, is complex to capture in a static algorithm, and a medical decision is not based on objective single data points but on subjective, context-sensitive longitudinal observations made by health care professionals during patient contact. This has consequences for the acceptance of autonomous CDS systems. Single measurements are not likely to lead to an acceptable *autonomous* action by a CDS system except for when doing nothing is likely to cause more harm (eg, an automated defibrillator that decides whether electric cardioversion is indicated based on ECG input and applies the appropriate therapy itself). Other accepted autonomous CDS systems, such as pacemakers or insulin pumps, gather continuous data, and thereby, learn and improve their efficacy for the individual patient. As longitudinal data collected by sensors inside and outside the hospital are becoming more ubiquitous, the value and applicability of accepted autonomous CDS systems enriched with these data are likely to improve. Still, the decision to implement a pacemaker and monitoring and tweaking its action and settings during follow-up are a doctor's job, including contextual expert knowledge and the patient's own preferences into the CDS system. Therefore, clinical reasoning is unlikely to ever be completely replaced by fully automated decision making through machine learning without human intervention. Even when expert knowledge is already embedded in medical data, and when the CDS system is clinically validated, the need for health care professionals to navigate the intricacies of incorporating expert knowledge in the ultimate clinical decision must not be underestimated. This includes implicit knowledge or *gut feeling* for which computability is limited [29]. Moreover, the final decision (how) to use the CDS is up to the health care professional and their patient.

Therefore, most CDS systems do not aim to replace health care professionals but are designed to support them. For example, the recent algorithms in the field of medical imaging preprocess data and take over tedious and simple tasks so that radiologists and pathologists can focus on more complex cases, acting more creatively. Given the recent developments in data protection regulations, health care professionals cannot easily be replaced by AI. The European General Data Protection Regulations state "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." This regulation appears tailored to the medical profession. Health care professionals (but basically everybody processing data) have a responsibility to guard their patients against irresponsible implementation of data-driven technologies. This especially holds true for self-learning algorithms that self-adapt to the patient population without human intervention, which may autonomously change considerably over time (so-called *algorithmic drift*). What is the meaning of human intervention in this sense, if the health care professional has no insight at all into the opaque model? In this light, how can health care professionals still justify their actions? How does this translate to accountability?

Importantly, the decision of how to respond to a CDS system is a moral one, and moral considerations regarding when to treat or not to treat are the expertise of human beings rather than that of AI systems. A CDS system cannot decide whether *primum non nocere* applies to a specific situation as harm and good and quality of life depend on personal judgement, context, and preferences of human beings. Some patients may be willing to take a risk that others would not, including application of a CDS system with a black box algorithm to their specific case. This way, cultural difference may indicate the need for locally tweaked systems. People, whether patients or their loved ones, should participate in shared decision making, tailoring the usage and outcomes of CDS systems to their wishes. What is *best* for the patient depends on more than just the output of a CDS system.

In conclusion, it is of paramount importance that health care professionals initiate and guide the development and implementation of CDS in clinical care, as opposed to waiting to be overwhelmed by current technological advancements. Most data scientists are not medical experts, and vice versa. Therefore, data scientists and health care professionals should team up in an interprofessional fashion, preferably also including patients. Data scientists who enthusiastically welcome recent innovations in AI pose a bold claim and carry the burden of proof to equip health care with suitable CDS tools. Once health care professionals can be convinced of the added benefit of CDS for their patients, they may acknowledge the necessity and value of data collection, interpretation, and curation, so they may embrace their expanding role and further evolve from doctor knows best to doctor does best.

Acknowledgments

TB is funded through the Netherlands CardioVascular Research Initiative CVON2017-22 ARGUS. This project is financed by the PPP Allowance made available by Top Sector Life Sciences & Health to Hartstichting to stimulate public-private partnerships.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

This table shows a number of indicative characteristics of the Utrecht Patient Oriented Database (UPOD) as of February 2018.

[PDF File (Adobe PDF File), 11KB - jmir_v21i3e11732_app1.pdf]

References

1. Moja L, Kwag KH, Lytras T, Bertizzolo L, Brandt L, Pecoraro V, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. Am J Public Health 2014 Dec;104(12):e12-e22 [FREE Full text] [doi: 10.2105/AJPH.2014.302164] [Medline: 25322302]
2. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, CADET II Group. Single reading with computer-aided detection for screening mammography. N Engl J Med 2008 Oct 16;359(16):1675-1684. [doi: 10.1056/NEJMoa0803545] [Medline: 18832239]
3. Cruz-Roa A, González F, Galaro J, Judkins AR, Ellison D, Baccon J, et al. A visual latent semantic approach for automatic analysis and interpretation of anaplastic medulloblastoma virtual slides. Med Image Comput Comput Assist Interv 2012;15(Pt 1):157-164. [Medline: 23285547]
4. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. J Am Med Assoc 2016 Dec 13;316(22):2402-2410. [doi: 10.1001/jama.2016.17216] [Medline: 27898976]
5. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. npj Digital Med 2018 May 8;1(1):1-1 [FREE Full text] [doi: 10.1038/s41746-018-0029-1]
6. Towle A, Brown H, Hofley C, Kerston RP, Lyons H, Walsh C. The expert patient as teacher: an interprofessional health mentors programme. Clin Teach 2014 Jul;11(4):301-306. [doi: 10.1111/tct.12222] [Medline: 24917101]
7. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: Mapping the debate. Big Data & Society 2016 Dec 19;3(2). [doi: 10.1177/2053951716679679]
8. Fernald DH, Coombs L, DeAlleaume L, West D, Parnes B. An assessment of the Hawthorne Effect in practice-based research. J Am Board Fam Med 2012;25(1):83-86 [FREE Full text] [doi: 10.3122/jabfm.2012.01.110019] [Medline: 22218628]
9. Agniel D, Kohane IS, Weber GM. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. Br Med J 2018 Apr 30;361:k1479 [FREE Full text] [Medline: 29712648]
10. ten Berg MJ, Huisman A, van den Bemt PM, Schobben AF, Egberts AC, van Solinge WW. Linking laboratory and medication data: new opportunities for pharmacoepidemiological research. Clin Chem Lab Med 2007;45(1):13-19. [doi: 10.1515/CCLM.2007.009] [Medline: 17243908]
11. Graham KC, Cvach M. Monitor alarm fatigue: standardizing use of physiological monitors and decreasing nuisance alarms. Am J Crit Care 2010 Jan;19(1):28-34 [FREE Full text] [doi: 10.4037/ajcc2010651] [Medline: 20045845]
12. Hardin JM, Chhieng JC, Ozaydin B. Data Mining and Clinical Decision Support Systems. In: Berner E, editor. Clinical Decision Support System. New York: Springer; 2007.
13. Berner ES, editor. Clinical Decision Support Systems: Theory and Practice 3rd. New York: Springer; 2016.
14. Mohri M, Rostamizadeh A, Talwalkar A. Foundations of Machine Learning. Cambridge: The MIT Press; 2012.
15. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521(7553):436-444. [doi: 10.1038/nature14539] [Medline: 26017442]
16. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Sci Rep 2016 Dec 17;6:26094. [doi: 10.1038/srep26094] [Medline: 27185194]
17. Kappen TH, Vergouwe Y, van Klei WA, van Wolfswinkel L, Kalkman CJ, Moons KG. Adaptation of clinical prediction models for application in local settings. Med Decis Making 2012;32(3):E1-10 [FREE Full text] [doi: 10.1177/0272989X12439755] [Medline: 22427369]
18. Kappen TH, Peelen LM. Prediction models: the right tool for the right problem. Curr Opin Anaesthesiol 2016 Dec;29(6):717-726. [doi: 10.1097/ACO.0000000000000386] [Medline: 27536997]
19. Kappen TH, van Loon K, Kappen MA, van Wolfswinkel L, Vergouwe Y, van Klei WA, et al. Barriers and facilitators perceived by physicians when using prediction models in practice. J Clin Epidemiol 2016 Feb;70:136-145 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.09.008] [Medline: 26399905]
20. Kuhn M, Johnson K. Applied Predictive Modeling. New York: Springer; 2013:600.
21. The European Parliament and the Council of the European Union. GDPR URL: https://eugdpr.org/ [accessed 2019-01-29] [WebCite Cache ID 75n8Qm2h9]
22. US Food & Drug Administration. 2016. Medical Device Reporting for Manufacturers (MDR) URL: http://cat.ocw.uci.edu/media/AR01/06011/mdr.pdf [WebCite Cache ID 71DtXDpNP]
23. Varghese J, Kleine M, Gessner SI, Sandmann S, Dugas M. Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review. J Am Med Inform Assoc 2018 May 1;25(5):593-602. [doi: 10.1093/jamia/ocx100] [Medline: 29036406]

XSL•FO
RenderX

24. Lobach DF. The road to effective clinical decision support: are we there yet? B Med J 2013 Mar 13;346:f1616. [doi: 10.1136/bmj.f1616] [Medline: 23487164]
25. Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003;10(6):523-530 [FREE Full text] [doi: 10.1197/jamia.M1370] [Medline: 12925543]
26. Scheepers-Hoeks AM, Grouls RJ, Neef C, Ackerman EW, Korsten EH. Physicians' responses to clinical decision support on an intensive care unit--comparison of four different alerting methods. Artif Intell Med 2013 Sep;59(1):33-38. [doi: 10.1016/j.artmed.2013.05.002] [Medline: 23746663]
27. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Effect of clinical decision-support systems: a systematic review. Ann Intern Med 2012 Jul 3;157(1):29-43. [doi: 10.7326/0003-4819-157-1-201207030-00450] [Medline: 22751758]
28. Macnish K. Unblinking eyes: the ethics of automating surveillance. Ethics Inf Technol 2012 Mar 7;14(2):151-167. [doi: 10.1007/s10676-012-9291-0]
29. Morek R. The regulatory framework for online dispute resolution. U Tol L Rev 2006:163-192 [FREE Full text]

Abbreviations

AI: artificial intelligence
CDS: clinical decision support
ECG: electrocardiogram
EHR: electronic health record

Edited by G Eysenbach; submitted 30.07.18; peer-reviewed by G Weber, D Hanson; comments to author 11.10.18; revised version received 05.11.18; accepted 26.11.18; published 19.03.19.

Please cite as:
Bezemer T, de Groot MCH, Blasse E, ten Berg MJ, Kappen TH, Bredenoord AL, van Solinge WW, Hoefer IE, Haitjema S
A Human(e) Factor in Clinical Decision Support Systems
J Med Internet Res 2019;21(3):e11732
URL: http://www.jmir.org/2019/3/e11732/
doi:10.2196/11732
PMID:30888324

©Tim Bezemer, Mark CH de Groot, Enja Blasse, Maarten J ten Berg, Teus H Kappen, Annelien L Bredenoord, Wouter W van Solinge, Imo E Hoefer, Saskia Haitjema. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Social Media Intervention Design: Applying an Affordances Framework

Megan A Moreno[1], MD, MSEd, MPH; Jonathan D'Angelo[1], PhD

Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin, Madison, WI, United States

Corresponding Author:
Megan A Moreno, MD, MSEd, MPH
Department of Pediatrics
School of Medicine and Public Health
University of Wisconsin
2870 University Ave
Suite 200
Madison, WI, 53705
United States
Phone: 1 6082632846
Fax: 1 608 263 0503
Email: mamoreno@pediatrics.wisc.edu

Abstract

Social media interventions are a growing area of internet research, particularly for adolescent health. Researchers developing social media intervention approaches face the task of selecting a social media platform for their intervention. In this paper, we present the theoretical framework of affordances to help guide social media platform selection for intervention research. Affordances are a concept often used in fields associated with design and by those systematically studying the impact of a design of an object. Thus, the affordances approach is often used by those considering the impact of information technology and the design of social media platforms. Affordances are often described as properties of artifacts that can be recognized by users and contribute to their function or items that present an action possibility. We describe common affordances that can be applied to intervention design as well as current evidence and an intervention case example for each affordance. A scientific approach for the selection of the appropriate social media platform for a given intervention is an important research priority to advance the field of internet research.

(J Med Internet Res 2019;21(3):e11014) doi:10.2196/11014

KEYWORDS

social media; health; adolescent; research

Introduction

Background

Social media interventions are a growing area of internet research. This trend is illustrated by the growing number of systematic reviews examining social media interventions across different health topics including chronic illness and behavioral risk [1-3]. Social media as a platform for interventions targeting adolescents has technological benefits including ease of scaling up an intervention to reach large numbers of participants. Furthermore, social media has developmental salience for the adolescent population. Over 90% of adolescents report going on the Web every day, and almost a quarter of teens report that they go on the Web "almost constantly [4]." Adolescents have been dubbed *digital natives* given that they have grown up with access to computers and the internet from an early age [5]. Adolescents typically maintain a "social media portfolio" by using a number of different platforms including Facebook, Instagram, and Tumblr [4].

Researchers interested in social media intervention approaches are faced with many tasks in the design of an intervention including theoretical or clinical grounding, selecting an appropriate target population, and defining appropriate outcome measures. For a social media intervention, an important task is selecting a social media platform for the intervention. If investigators rely on familiarity with or popularity of a platform, these selection criteria lack scientific grounding and may introduce bias in study design. In this paper, we present a theoretical framework to help guide social media platform selection for intervention research. The goal of this framework is to increase the likelihood that the success of interventions hinges on replicable technological qualities as opposed to personal preferences or other spurious considerations. This

XSL•FO
RenderX

paper is aligned with recent calls for more availability of information on processes and mechanisms to understand what makes Web intervention approaches feasible and effective [6]. We will first define our approach using affordances and describe a methodological approach for incorporating 4 common affordances into intervention design. Within each affordance section, we will include relevant evidence and a hypothetical case scenario. An evidence-based approach for the selection of the appropriate social media platform for a given intervention is thus an important research priority to advance the field of internet research.

Affordances Defined

Affordances are a concept often used in fields associated with design and by those systematically studying the impact of a design of an object. Thus, the affordances approach is often used by those considering the impact of information technology and the design of social media platforms. Affordances are often described as properties of artifacts that can be recognized by users and contribute to their function [7] or items that present an action possibility [8]. Perceived affordances can also be conceptualized as "design aspects of objects that suggest to the user how the object should be used [7]." These definitions can feel somewhat opaque; however, a simple example of an affordance is that the design elements of a chair suggest to a person that the object could be used for sitting. A chair has 4 legs, a flat platform, and a straight back; these design affordances support the idea of a chair being an object designed for sitting. Importantly, although an individual's goal or desire may change, an affordance does not evolve [9].

Given the rapid evolution of computer-mediated environments, many fields that study outcomes associated with this context have pushed researchers to utilize this framework for the sake of developing lasting and generalizable theory. Several scientific fields and disciplines have moved to discuss or encouraged the move to discuss technology not in terms of specific communication platforms but rather in terms of the affordances. These fields include management science [10], communication theory [11], education research [12], design research [13], information technology [14], organizational research [15], and tourism and marketing research [16]. The initiative is also present in clinical research as previous work has incorporated affordances related to utilizing social media to manage chronic disease [17]. Just as observational research benefits from enhanced clarity and meaning from the affordance framework [18], it is likely that intervention research will benefit as well.

Benefits to an Affordance Approach in Social Media Intervention Research

There are several benefits to the application of the affordance approach for intervention design. First, an affordance approach provides a scientific framework for the selection of a social media platform. With this approach, a researcher can identify the necessary functionalities of an intervention such as key functions the platform should have or critical constructs linked to the behavioral or health-related theory that inform the intervention. These functions and constructs can then be used to match to the affordances of a social media platform. Second, this approach goes beyond the selection of interventions on the basis of the popularity of platforms. Given that the popularity of platforms can change, a public perception that social media is always changing is a concern that scientists may encounter. With an affordance approach, a scientist can describe the affordances needed for intervention and thus present potential platforms that would meet these criteria rather than relying on a single platform. If a successful intervention sees decline because a platform is losing popularity or becoming obsolete, an intervention designed around affordances can identify another platform that fits (or even design one). Finally, an affordance approach takes a step beyond the brand name approach of selection of social media platforms and changes the dialogue to be more closely linked to theory and technical function.

Affordances to Consider in Intervention Design

Among the many affordances described in the literature, we will focus on 5 affordances that apply to social media. In what follows, we will describe the affordance category, provide evidence of how that particular technological affordance has been linked to psychological or behavioral change, and illustrate an intervention design targeted at an adolescent population.

Identity Affordances

Definition

The first category of affordances of social media is identity affordances. Identity affordances include opportunities on social media platforms for *identity development and portrayal*. An example of a high-identity affordance platform is Facebook, on which users can upload a profile picture to identify themselves, list their "likes," and share "life events." Identity is further emphasized by the expectation on Facebook to use one's real name (or a version of it). In contrast, Reddit is a low-identity social media platform on which users are typically identified by a chosen username. Furthermore, the emphasis on Reddit is more on what content a user contributes to group conversations rather than the personal identity of who posts.

It has been proposed that social media platforms that require fewer identity clues can allow users to take on new identities within different Web-based conversations, sometimes called *generative role taking* [19]. For example, an overweight teen working to become more physically active could create a new Twitter account called @TeenWhoLovesToRun dedicated to that identity. It is important to note that an extreme of this generative role-taking is represented by social media platforms on which users are anonymous, such as YikYak (shut down in 2017), and on which users could post with complete anonymity. A risk of fewer identity clues can lead to users taking on roles that are more dangerous, such as being aggressive, trolling, or bullying others.

Platforms that allow users to develop their identity may present benefits to adolescents in being able to explore and experiment with their identities, a critical developmental task of adolescence. However, by enhancing one's own identity presence on a site, there are also risks of being identified by strangers. Thus, identity should be balanced with whether the

XSL•FO
RenderX

site allows users to establish *privacy* settings. This attention to privacy may be more salient for interventions that involve illegal or stigmatizing behaviors such as substance use or mental health conditions. A balance of identity development and privacy protection is essential.

Evidence

The effectiveness of utilizing social media to help individuals shift identity or self-concept has been well established in experimental and observational settings. For example, when individuals were asked to present themselves as more extraverted on the Web, they, in turn, perceived themselves as more extraverted [20]. Furthermore, individuals asked to present themselves on the Web as loyal to a particular brand positively increased attitude toward that brand [21]. Finally, couples who post more relationship-related material on Facebook (thus cultivating a certain relationship identity on the Web) were more likely to still be together after 6 months [22]. Hence, there is empirical evidence to suggest that designing interventions on the basis of the ability of a platform to allow users to portray different aspects of identity can lead to positive change.

A Hypothetical Intervention

A social media intervention was being designed to promote physical activity among overweight adolescents. The intervention was intended to deliver positive messages to adolescents. The messages were designed to promote reflection on participants' own skills and strengths to shift their view of themselves as active and fit, as well as providing prompts to engage in physical activity. The primary behavioral theory informing the intervention, which served as the root of the messages, was Self-Determination Theory [23]. In addition, findings of identity shift [20,24], the notion that individuals can internalize qualities that they present on the Web, supported this line of intervention. Thus, this intervention relied on the use of a social media platform that allowed identity affordances.

The researchers determined that low-identity affordance platforms such as Reddit and Twitter may not be the best fit for this intervention and considered the higher-identity affordance platforms of Instagram or Facebook. The intervention was then designed to take place on Facebook and include Facebook Badges and content that teens could incorporate into their own profiles. The rationale for the Facebook Badges approach was to allow teens to take salient intervention messages and integrate them into their own digital identities on Facebook.

Social Affordances

Definition

Given the interactive nature of social media, it is no surprise that most social media platforms offer many social affordances. Social affordances include a sense of *belongin* g to a group such as a group focused on a particular interest, experience, social group, or religion [25]. Some social media platforms provide specific tools that allow the user to identify members of their group and enhance a feeling of belonging, such as tagging [26]. Hashtags (ie, content labeled with a # sign) are commonly used on Twitter and Instagram and can enhance belonging. When content is labeled with a hashtag, it is connected to all other content on that site that has also been labeled in that way. Thus, a hashtag can allow the creation of a community of users across different social networks who are connected by use of a particular topic or term. Any user on a platform can search by a hashtag and be connected to all content by others using that hashtag; thus, a community of those using that particular word or phrase can be connected. A positive example of belonging is teen cancer survivors who use a particular hashtag to connect to other teen survivors nationally to provide support and share similar experiences. Furthermore, social media may promote *network-informed associations* such as when Facebook suggests friends for a user on the basis of the user's friends' friends. This allows users to see how friends are connected to other people and see their interests [19].

An important task for adolescents to learn within the social media setting is how much information to share. A previous study found that intimate disclosures shared publicly can be judged as inappropriate [27]. Users must thus engage in *audience management* by monitoring and checking what audience is receiving their messages. This can impact how much personal information a teen is willing to share within a group-based intervention such as one using private Facebook groups. At the same time, users must also come to realize that their audience may not be exactly who they imagine it to be [28]. Assuming an audience different than one actually finds on social media may have unique psychological or social consequences. Finally, social media allows teens to experience and participate in *metavoicing* [19]. This term refers to how a social media user who posts anything in social media is engaging in a larger context including other people's presence, profiles, content, and activities. Thus, an adolescent posting on Twitter is not merely voicing a single opinion but contributing to content that is already in that space and connected to others.

Evidence

The use of social affordances in health interventions specifically designed around this notion has already been shown to be effective in several areas. In a previous intervention, patients with non–small cell lung cancer who partook in an online support group, as opposed to those who only had access to internet articles, experienced fewer symptoms of distress [29]. Another intervention found that newly diagnosed breast cancer patients who received support from breast cancer survivors in an online forum reported improved quality of life and decreased depression [30]. Study trials are being designed to create Web-based portals for older adults who often suffer from isolation; early findings have demonstrated potential benefits [31]. Although not all researchers can design their own Web-based portal, this affordance can be utilized as it exists on already popular social media websites.

A Hypothetical Intervention

A social media intervention was designed to promote social support among teens with depression. The intervention was designed to promote peer-to-peer support among teens with depression, as peer support has received strong evidence as an effective intervention [32]. Thus, social affordances that promote peer-to-peer communication were a priority. Furthermore, pilot testing with teens led to the feedback that the social

XSL•FO
RenderX

communication needed to be in private settings. The rationale for the platform selected was to utilize private groups in a platform that teens already visited.

Given that the teens already utilized Reddit, researchers decided to implement this intervention as a private group on Reddit. In addition to already being utilized by the participants, the platform had added affordances. There was no identity requirement for this platform, enabling participants to comfortably disclose more of their thoughts and feelings. The functional affordances of this message board allowed for the construction of messages at a time and rate that were comfortable to the participants.

In carefully analyzing the platform according to the affordance framework, the researchers realized that Reddit offers 1 emotional affordance that may be less beneficial: the ability to upvote and downvote messages. However, moderators of a given message board, the position that the intervention leaders will act in, have the ability to remove the ability of board members to downvote messages. By taking this action, the intervention platform was adjusted so that posts could only receive positive feedback.

Cognitive Affordances

Definition

Cognitive affordances include using social media tools *to expand one's learning*. Examples may include increasing awareness of global news events using Twitter or learning a new strength training exercise on YouTube. These affordances may be particularly useful for adolescents who may not traditionally get exposed to this information offline [7]. Cognitive affordances may also include the opportunity for enhanced creativity, such as on platforms such as Tumblr that allow customizable content.

Social media can also allow *reallocation of cognitive resources* by allowing a user to focus on 1 aspect of a person's presentation at a time versus the many cognitive cues that emerge during a face-to-face conversation with a person. Face-to-face communication requires an adolescent to be cognizant of the complex interplay of words, tone, facial expression, and body language. In contrast, social media tools allow more simplified communication venues, such as Facebook messaging, which is text-only against a background of personal information. This type of communication requires fewer cognitive inputs and may be easier for some teens, but it can also lead to the potential for misinterpretation of information. For example, interventions that include humor or sarcasm in messaging may be at risk for misinterpretation without voice or visual cues. As a developmental task of adolescence is learning how to develop skills in communication, these tools may both help and hinder an adolescent's journey.

Another cognitive affordance of many social media platforms is *triggered attending*. Triggered attending involves rejoining a Web-based conversation or responding to content when an automated alert informs the user to do so [19]. Examples include setting mobile phone alerts to trigger when specific users or friends post content. This triggered attending may be helpful to reengage adolescents at different stages of an intervention, but it could also be disruptive to adolescents who are engaged in other activities such as homework or driving.

Evidence

Similar to utilizing the social aspect of the internet and social media, the cognitive affordances of social media have begun to show promise. At the most fundamental level, social media has been shown to support informal learning at home [33], as well as be a source for those who seek health information on the Web [34,35]. More dynamically, aspects such as the reallocation of cognitive resources have been shown to allow users to reduce social anxiety associated with interpersonal interaction [36,37]. This reallocation can allow users to potentially present their best possible self, leading to more reciprocal self-disclosure and more intimate Web-based relationships [38]. Triggered attending was well documented in cell phone–based interventions, where the evidence can be simply expressed in the form of text message (short message service, SMS) reminders regarding patient appointments [39]. Given that mobile phones now produce notifications for social media in just as noticeable a fashion as SMS text messages, social media interventions are likely to benefit from such an affordance.

A Hypothetical Intervention

A social media intervention was designed to promote teaching of diabetes management skills among newly diagnosed adolescents. A rich media platform for teaching and a high cognitive affordances approach was a priority for the research team. The team determined that video was the best medium to teach basic skills, as evidence suggests that video is more efficient than text-based e-learning for practical or procedural skills sets [40]. Researchers worked with teens to create a private YouTube channel and new YouTube accounts for participants in the study as identity affordances were not critical to the study design. The research team was able to share videos to teach diabetes management skills, and participants were able to ask questions within the YouTube channel. Participants were also invited to exercise creativity and create their own teaching content to share on the private channel.

Even if participants were not inclined to create their own content, the functional affordances of this domain offered other routes to participation. Specifically, YouTube allows for the cultivation of video lists on profiles. Hence, participants were encouraged to cultivate a list of nutrition-related videos on each profile. This enabled 2 positive outcomes. First, participants and moderators were able to discuss the merits (or flaws) of each video to enhance the media and nutritional literacy of participants. Second, libraries of helpful videos were created for the participants to reference over time.

Emotional Affordances

Definition

Emotional affordances include attributes of social media that can trigger or stimulate users' emotional reactions [7]. For example, many platforms allow users to express emotion via the capacity to like or dislike content such as the "favorite" function on Twitter or "upvoting" or "downvoting" on Reddit.

XSL•FO
RenderX

Emotional affordances can also include *generating empathy* by seeing personal photographs or names alongside messages or news stories. It is common to see fund-raising efforts on social media that capitalize on emotional affordances by providing photos and personal stories. Research groups may seek to engage participants by sharing personal information about the researchers or providing photos of the research team. In contrast, another affordance of social media is *comparison*. Previous studies have illustrated that by examining other people's social media profiles, users are at risk of feeling inadequate or envious [41].

Evidence

Relatively simple emotional affordances can prove to have distinct emotional outcomes. For instance, receiving Facebook likes on a post can make users feel socially supported [42], and photographs can affect judgments of personality more so than text in social media [43]. Moreover, experiments suggest that individuals already utilize social media to manage their emotions: when individuals were induced into a negative mood, they viewed social media profiles of individuals they perceived as less physically attractive or successful (downward social comparison) [44]. The effects of social comparison have also been evidenced in a lab-based health intervention experiment: individuals who received an appearance-based sun-protection intervention that included downward social comparison (ie, viewed photographs of individuals who had already experienced sun damage) were associated with less sun-protective behaviors compared with viewing the control condition. However, upward social comparison leads to a slight increase in sun-protective behaviors [45].

A Hypothetical Intervention

An intervention was designed for smoking cessation for older adolescents. The investigators wanted to share stories of older adolescents who had successfully quit to utilize emotional affordances to motivate and connect to participants. The investigators built a blog using a blogging website and promoted content via Instagram. A critical component of how content was shared with participants on Instagram was to use personal photos of each new blogger.

This intervention was grounded in notions of homophile, as people are more likely to adopt health advice offered by similar others [46]. Photographs were crucial for generating the emotional connection as photographs can have a greater impact on the judgment of a target than textual information [43]. Hence, it was uniquely catching for participants to see a picture of someone just like them who had successfully quit smoking.

Functional Affordances

Definition

Functional affordances are principal functions that affect how social media messages are transmitted or saved. Although many of these affordances describe most social media platforms, an understanding of these affordances may be helpful for researchers to frame the key functional design elements of the tools needed for intervention.

Most social media platforms allow for *replicability* of messages, meaning that content can be reused by others. An example is the ability to "retweet" content on Twitter. This function can allow interventions to deliver "booster" doses of intervention messages by resharing previous content. Social media also allows messages to be highly *scalable*; there are numerous examples in the popular media of social media posts going "viral" and reaching far beyond the initial intended audience. For researchers, this can be an important consideration for ensuring that control group participants are not exposed to messages specific to an intervention if these are present on public social media.

Social media is also *searchable*; users can find specific content or people using a variety of Web-based means such as hashtags. Social media also provides *permanence* to messages by placing them in Web-based conversations or on profiles. Even on platforms that are intended to be ephemeral, such as Snapchat, content is generally stored or could be captured and saved by others. Social media allows for *unlimited composition time* before posting a message, in contrast with the back and forth of a typical face-to-face conversation. However, not all users utilize this affordance with every post. Finally, social media has varying levels of the user's ability to edit once something is posted. Some sites allow for removal or editing of content easily; Facebook allows for revising or deleting posts. However, the speed of Twitter can lead to challenges with removing or editing content once it has been retweeted by others. These functions may be important to particular intervention approaches and be incorporated as explicit tools that are relevant to a given intervention.

Evidence

The replicability of messages in social media in the form of repeats, shares, or videos going viral is a goal of many health interventions, but it is not always achieved. For example, Twitter has been shown to be an effective means of recruitment for health research when a message is successfully retweeted by others [47]. At the same time, there is often no control over replicability: health intervention designers cannot always get videos to go viral as they might intend [48]. Hence, many functional affordances need to be considered for potential outcomes, both intended and unintended, but they are not always the best focus for a mechanism of change.

A Hypothetical Intervention

An awareness-raising intervention message focused on sexually transmitted infection testing was developed by a group of researchers. They successfully pilot-tested the approach and found that it increased awareness of and willingness to seek testing in a sample of adolescents. They now want to bring the intervention into practice and launch a social media campaign using their messages. They leverage Twitter for the intervention, knowing that retweets are a less-effort-intensive way to engage participants, and ask for retweeting of the message to increase its reach. They measure the number of likes, retweets, and impressions for their message over a 7-day period.

XSL•FO
RenderX

Discussion

Summary

The ever-changing landscape of social media sites can lead to challenges for researchers to apply scientific scaffolding to justify the selection of platforms for interventions. We present a research priority of applying an affordance framework for mindful and evidence-based selection of social media platforms to match intervention requirements and approaches. As indicated in these examples, thoughtful articulation of intervention outcomes and assessment of platform affordances can lead to an evidence-based selection of an ideal intervention platform, considering the affordances of a platform can help identify potential strengths as well as undesirable outcomes in an intervention and offer solutions. For instance, in an example intervention above, we identified the potential negative outcomes of a downvote on Reddit and thus removed that option within that particular intervention. There also exists a second potential pathway to more effective interventions: the design of an intervention that requires a set of affordances not yet present in current technology. Rather than a limitation, this represents opportunity in the area of technology. Previous computer-mediated communication research has offered suggestions to website designers to increase the happiness of Web-based daters and increase the instructional efficiency of educational technology [49], and new app developers need to be cognizant of the affordances of the products they develop. Identifying a grouping of affordances that can be beneficial but does not yet exist represents an opportunity for new platform development.

Limitations

The affordances that we present here represent early efforts to define and apply affordances to adolescent health interventions. Conceptualization of affordances is an ongoing, iterative process across many research disciplines. Even in fields long interested in conceptualizing affordances, there is still debate around the definition and boundaries of each affordance [11]. However, the fact that this debate continues is evidence of the utility of this framework. Even if not perfect, it allows for purposeful selection and engagement across a broad and ever-evolving technological environment. We have presented several example interventions and named a few specific platforms as examples, but this neither implies that we endorse those platforms nor implies that all social media platforms are included in this paper.

Furthermore, in focusing this paper on an affordance approach, we recognize that intervention development takes into account many other factors beyond this theoretical approach. Appropriate targeting of a given platform to the participant group is critical. If your target demographic does not use a particular platform, this platform is not likely the one for your intervention. Platforms may also need to be considered on the basis of what behavior is targeted in the intervention and whether that behavior is one that is appropriate to discuss on that platform. Considering the flexibility or malleability of the platform selected is important; platforms with shifting features could change over the course of a study and impact results. Researchers reading this paper will want to consider the other feasibility and acceptability factors that they typically would in any given intervention planning, alongside their theoretical approach.

Conclusions

Despite these limitations, the affordances approach provides a theoretical framework for selecting intervention platforms on the basis of specific criteria and functions. We present a research priority of applying an affordance framework for mindful selection of social media platforms to match intervention requirements and approaches. As indicated in these examples, thoughtful articulation of intervention outcomes and assessment of platform affordances can yield a complementary partnership that leads to health outcomes.

Acknowledgments

The authors would like to thank Christine Richards for her assistance with this submission.

Conflicts of Interest

None declared.

References

1. Chou WY, Prestin A, Lyons C, Wen KY. Web 2.0 for health promotion: reviewing the current evidence. Am J Public Health 2013 Jan;103(1):e9-18. [doi: 10.2105/AJPH.2012.301071] [Medline: 23153164]
2. Park BK, Calamaro C. A systematic review of social networking sites: innovative platforms for health research targeting adolescents and young adults. J Nurs Scholarsh 2013 Sep;45(3):256-264. [doi: 10.1111/jnu.12032] [Medline: 23676115]
3. Elaheebocus SM, Weal M, Morrison L, Yardley L. Peer-based social media features in behavior change interventions: systematic review. J Med Internet Res 2018 Feb 22;20(2):e20 [FREE Full text] [doi: 10.2196/jmir.8342] [Medline: 29472174]
4. Lenhart A. Pew Research Center. Washington, DC: Pew Internet and American life Project; 2015. Teens, social media & technology overview 2015: URL: http://www.pewresearch.org/wp-content/uploads/sites/9/2015/04/PI_TeensandTech_Update2015_0409151.pdf [accessed 2019-02-10] [WebCite Cache ID 765zEAS0L]
5. Palfrey J, Gasser U. Born Digital: Understanding the First Generation of Digital Natives. Philadelphia, USA: Basic Books; 2012:78--132.

XSL•FO
RenderX

6. Rogers M, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Dec 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
7. Zhao Y, Liu J, Tang J, Zhu Q. Conceptualizing perceived affordances in social media interaction design. Aslib Proc 2013 Mar;65(3):289-303. [doi: 10.1108/00012531311330656]
8. Gibson J. The Ecological Approach to Visual Perception. Boston, MA: Houghton Mifflin; 1979.
9. Norman D. The Psychology of Everyday Things. New York, NY: Basic Books; 1988.
10. Leonardi P, Vaast E. Social media and their affordances for organizing: a review and agenda for research. Acad Manag Ann 2017 Jan;11(1):150-188 [FREE Full text] [doi: 10.5465/annals.2015.0144]
11. Evans S, Pearce KE, Vitak J, Treem JW. Explicating affordances: a conceptual framework for understanding affordances in communication research. J Comput-Mediat Comm 2016 Dec 26;22(1):35-52. [doi: 10.1111/jcc4.12180]
12. Kannengiesser U, Gero JS. A process framework of affordances in design. Des Issues 2012;28(1):50-62 [FREE Full text] [doi: 10.2307/41427809]
13. Webb M. Affordances of ICT in science learning: implications for an integrated pedagogy. Int J Sci Edu 2005;27(6):705-735. [doi: 10.1080/09500690500038520]
14. Bygstad B, Munkvold BE, Volkoff O. Identifying generative mechanisms through affordances: a framework for critical realist data analysis. J Info Tech 2016 Mar;31(1):83-96. [doi: 10.1057/jit.2015.13]
15. Ellison N, Gibbs JL, Weber MS. The use of enterprise social network sites for knowledge sharing in distributed organizations. Am Behav Sci 2014 Jul 17;59(1):103-123. [doi: 10.1177/0002764214540510]
16. Cabiddu F, Carlo MD, Piccoli G. Social media affordances: enabling customer engagement. Ann Tour Res 2014 Sep;48:175-192. [doi: 10.1016/j.annals.2014.06.003]
17. Merolli M, Gray K, Martin-Sanchez F. Health outcomes and related effects of using social media in chronic disease management: a literature review and analysis of affordances. J Biomed Inform 2013 Dec;46(6):957-969 [FREE Full text] [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]
18. McLeod J, Pan Z. Concept explication and theory construction. In: Dunwoody S, Becker BL, McLeod DM, Kosicki G, editors. The Evolution of Key Mass Communication Concepts: Honoring Jack M McLeod. Cresskill, USA: Hampton Press; 2005:13-76.
19. Majchrzak A, Faraj S, Kane GC, Azad B. The contradictory influence of social media affordances on online communal knowledge sharing. J Comput-Mediat Comm 2013 Oct 18;19(1):38-55. [doi: 10.1111/jcc4.12030]
20. Gonzales A, Hancock J. Identity shift in computer-mediated environments. Media Psychol 2008 Jun 25;11(2):167-185 [FREE Full text] [doi: 10.1080/15213260802023433]
21. Carr C, Hayes R. Identity shift effects of self-presentation and confirmatory and disconfirmatory feedback on self-perceptions of brand identification. Media Psychol 2017 Nov 13;1:1-27. [doi: 10.1080/15213269.2017.1396228]
22. Toma C, Choi M. The couple who Facebooks together, stays together: Facebook self-presentation and relationship longevity among college-aged dating couples. Cyberpsychol Behav Soc Netw 2015 Jul;18(7):367-372. [doi: 10.1089/cyber.2015.0060] [Medline: 26167834]
23. Fortier M, Duda JL, Guerin E, Teixeira PJ. Promoting physical activity: development and testing of self-determination theory-based interventions. Int J Behav Nutr Phys Act 2012 Mar 02;9:20 [FREE Full text] [doi: 10.1186/1479-5868-9-20] [Medline: 22385751]
24. D'Angelo J, Kerr B, Moreno MA. Facebook displays as predictors of binge drinking: from the virtual to the visceral. Bull Sci Technol Soc 2014;34(5-6):159-169 [FREE Full text] [doi: 10.1177/0270467615584044] [Medline: 26412923]
25. Lundby K. Patterns of belonging in ofline/offline interfaces of religion. Inf Commun Soc 2011 Dec;14(8):1219-1235. [doi: 10.1080/1369118X.2011.594077]
26. Cook J, Pachler N. Online people tagging: social (mobile) network(ing) services and work-based learning. Br J Educ Technol 2012 Aug 24;43(5):711-725. [doi: 10.1111/j.1467-8535.2012.01346.x]
27. Bazarova NN. Public intimacy: disclosure interpretation and social judgments on Facebook. J Commun 2012 Aug 27;62(5):815-832. [doi: 10.1111/j.1460-2466.2012.01664.x]
28. Litt E. Knock, knock. Who's there? The imagined audience. J Broadcast Electron Media 2012 Jul;56(3):330-345. [doi: 10.1080/08838151.2012.705195]
29. Gustafson D, DuBenske LL, Namkoong K, Hawkins R, Chih MY, Atwood AK, et al. An eHealth system supporting palliative care for patients with non-small cell lung cancer: a randomized trial. Cancer 2013 May 1;119(9):1744-1751 [FREE Full text] [doi: 10.1002/cncr.27939] [Medline: 23355273]
30. Moon T, Chih M, Shah DV, Yoo W, Gustafson DH. Breast cancer survivors' contribution to psychosocial adjustment of newly diagnosed breast cancer patients in a computer-mediated social support group. Journal Mass Commun Q 2017 Jan 19;94(2):486-514. [doi: 10.1177/1077699016687724]
31. Gustafson D, McTavish F, Gustafson DH, Mahoney JE, Johnson RA, Lee JD, et al. The effect of an information and communication technology (ICT) on older adults' quality of life: study protocol for a randomized control trial. Trials 2015 Apr 25;16:191 [FREE Full text] [doi: 10.1186/s13063-015-0713-2] [Medline: 25909465]

XSL•FO
RenderX

32. Pfeiffer P, Heisler M, Piette JD, Rogers MA, Valenstein M. Efficacy of peer support interventions for depression: a meta-analysis. Gen Hosp Psychiatry 2011;33(1):29-36 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.10.002] [Medline: 21353125]
33. Selwyn N. The use of computer technology in university teaching and learning: a critical perspective. J Comput Assist Learn 2007;23(2):83-94. [doi: 10.1111/j.1365-2729.2006.00204.x]
34. Greenhow C, Robelia B. Informal learning and identity formation in online social networks. Learn Media Tech 2009 Jun 23;34(2):119-140. [doi: 10.1080/17439880902923580]
35. Fox S. Pew Research Center. Washington, DC; 2014 Jan 15. The social life of health information URL: http://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/ [accessed 2019-02-10] [WebCite Cache ID 7667d4Ka1]
36. High A, Caplan SE. Social anxiety and computer-mediated communication during initial interactions: implications for the hyperpersonal perspective. Comput Hum Behav 2009 Mar;25(2):475-482. [doi: 10.1016/j.chb.2008.10.011]
37. Walther J. Selective self-presentation in computer-mediated communication: hyperpersonal dimensions of technology, language, and cognition. Comput Hum Behav 2007 Sep;23(5):2538-2557. [doi: 10.1016/j.chb.2006.05.002]
38. Trepte S, Reinecke L. The reciprocal effects of social network site use and the disposition for self-disclosure: a longitudinal study. Comput Hum Behav 2013 May;29(3):1102-1112. [doi: 10.1016/j.chb.2012.10.002]
39. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
40. Buch S, Treschow FP, Svendsen JB, Worm BS. Video- or text-based e-learning when teaching clinical procedures? A randomized controlled trial. Adv Med Educ Pract 2014;5:257-262 [FREE Full text] [doi: 10.2147/AMEP.S62473] [Medline: 25152638]
41. Chou H, Edge N. "They are happier and having better lives than I am": the impact of using Facebook on perceptions of others' lives. Cyberpsychol Behav Soc Netw 2012 Feb;15(2):117-121 [FREE Full text] [doi: 10.1089/cyber.2011.0324] [Medline: 22165917]
42. Hayes R, Carr CT, Wohn DY. One click, many meanings: interpreting paralinguistic digital affordances in social media. J Broadcast Electron Media 2016 Mar;60(1):171-187. [doi: 10.1080/08838151.2015.1127248]
43. van der Helde B, D'Angelo JD, Schumaker EM. The effects of verbal versus photographic self-presentation on impression formation in Facebook. J Comm 2012;62(1):98-116. [doi: 10.1111/j.1460-2466.2011.01617.x]
44. Johnson B, Knobloch-Westerwick S. Glancing up or down: mood management and selective social comparisons on social networking sites. Comput Hum Behav 2014 Dec;41(2):33-39. [doi: 10.1016/j.chb.2014.09.009]
45. Mahler H, Kulik JA, Gerrard M, Gibbons FX. Effects of upward and downward social comparison information on the efficacy of an appearance-based sun protection intervention: a randomized, controlled experiment. J Behav Med 2010 Dec;33(6):496-507 [FREE Full text] [doi: 10.1007/s10865-010-9279-3] [Medline: 20652391]
46. Wang Z, Walther JB, Pingree S, Hawkins RP. Health information, credibility, homophily, and influence via the internet: web sites versus discussion groups. Health Commun 2008 Jul;23(4):358-368. [doi: 10.1080/10410230802229738] [Medline: 18702000]
47. O'Connor A, Jackson L, Goldsmith L, Skirton H. Can I get a retweet please? Health research recruitment and the Twittersphere. J Adv Nurs 2014 Mar;70(3):599-609. [doi: 10.1111/jan.12222] [Medline: 23909740]
48. Nguyen P, Gold J, Pedrana A, Chang S, Howard S, Ilic O, et al. Sexual health promotion on social networking sites: a process evaluation of The FaceSpace Project. J Adolesc Health 2013 Jul;53(1):98-104. [doi: 10.1016/j.jadohealth.2013.02.007] [Medline: 23583509]
49. Antonenko P, Dawson DK, Sahay A. A framework for aligning needs, abilities and affordances to inform design and practice of educational technologies. Br J Educ Technol 2016 Jun 09;48(4):916-927 [FREE Full text] [doi: 10.1111/bjet.12466]

Abbreviations

SMS: short message service

Edited by G Eysenbach; submitted 20.06.18; peer-reviewed by K Aguirre, E Fisser; comments to author 21.09.18; revised version received 15.12.18; accepted 28.12.18; published 26.03.19.

Please cite as:
Moreno MA, D'Angelo J
Social Media Intervention Design: Applying an Affordances Framework
J Med Internet Res 2019;21(3):e11014
URL: http://www.jmir.org/2019/3/e11014/
doi:10.2196/11014
PMID:30912754

XSL•FO
RenderX

©Megan A Moreno, Jonathan D'Angelo. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

The Rise of the Empowered Physician in the Digital Health Era: Viewpoint

Bertalan Mesko[1,2*], MD, PhD; Zsuzsa Győrffy[1,2*], PhD

[1]The Medical Futurist Institute, Budapest, Hungary

[2]Department of Behavioral Sciences, Semmelweis University, Budapest, Hungary

[*]all authors contributed equally

Corresponding Author:
Bertalan Mesko, MD, PhD
The Medical Futurist Institute
Povl Bang-Jensen u 2/B1 4/1
Budapest, 1118
Hungary
Phone: 36 +36703807260
Email: berci@medicalfuturist.com

Abstract

Being a 21st-century health care provider is extremely demanding. The growing number of chronic diseases, lack of medical workforce, increasing amounts of administrative tasks, the cost of medical treatment, and rising life expectancy result in an immense challenge for medical professionals. This transformation has been triggered by the growing presence of digital health. Digital health does not only refer to technological transformation; it also fundamentally reshapes the physician-patient relationship and treatment circumstances. We argue that patient empowerment, the spread of digital health, the biopsychosocial-digital approach, and the disappearance of the ivory tower of medicine lead to a new role for physicians. Digital health allows the job of being a medical professional to become more rewarding and creative. The characteristics of a physician-as-idol could shift from self-confident to curious, from rule follower to creative, and from lone hero to team worker. Empowered physicians (e-physicians) can be described as "electronic," where they use digital technologies in their practice with ease; "enabled," where they are enabled by regulations and guidelines; and "empowered," where they are empowered by technologies that support their job and their empowered patients (e-patients). They can be described as "experts" in the use of technologies in their practice or in knowing the best, most reliable, and trustworthy digital health sources and technologies. They can also be described as "engaged," when understanding the feelings and points of view of their patients, giving relevant feedback, and involving them throughout the whole healing process. The skills and approaches that characterize this era of e-physicians, such as face-to-face communication skills, digital literacy, interdisciplinarity, knowing where to find information, translating large amounts of data into insights for patients, among others, should always have been at the core of practicing medicine. However, the economical, technological, and administrative burden of the profession has not made it possible for most physicians to enjoy the benefits of their training, individual capabilities, and creativity. By understanding how digital health technologies can support or augment their capabilities, physicians would have the chance to practice the art of medicine like never before.

(J Med Internet Res 2019;21(3):e12490) doi:10.2196/12490

KEYWORDS

e-physician; e-patient; empowerment; doctor-patient relationship; digital health; technology

Introduction

The 21st century has brought unprecedented challenges to medical professionals. The growing number of chronic diseases, global doctor shortages, increasing costs of medical treatment, and rising life expectancy all together result in an immense challenge for medical professionals [1]. Managing and treating the increasing dominancy of chronic conditions started to depend more on cooperation between physicians and patients than on individual decisions. Physicians are pressed to integrate health IT into their jobs, while also trying to stay up-to-date with emerging technologies. They often have to deal with low-quality, bugged, or inefficient software and technologies that further decrease the amount of time they can spend with their patients [2].

XSL•FO
RenderX

Every second physician suffers from burnout according to a study by the American Medical Association; several other studies have highlighted the same observation worldwide [3,4]. The four major causes of burnout are bureaucratic tasks, spending too many hours at work, feeling like a cog in the wheel, and increasing computerization of practice. Additionally, having to balance between medically justified, economically affordable, and morally acceptable solutions is the trilemma of modern medicine. Therefore, being a 21st-century health care provider is extremely demanding.

The era of digital health, a cultural transformation that brings disruptive technologies to both patients and health care providers, might so far have benefited patients more than physicians. Physicians are trained to act as demigods who should not say, "I don't know," and should have all the information at their disposal, even when there are over 28 million medical papers in the PubMed database. In the meantime, information, knowledge, and technologies within the ivory tower have started to become accessible for anyone through new digital tools, social media, or crowdsourcing.

As such fundamental changes appeared only in a matter of years, while neither regulations, medical education, nor guidelines followed them, a fight-or-flight reaction has become common among physicians. Thus, many of them are either reluctant to adopt digital health or they protect the power they are used to having [5]. Not only has the sheer amount of information grown, it has also become crucial to know and be able to use even the latest technologies from apps and telemedicine to health sensors and portable diagnostic devices.

The new phenomenon we call *digital health* has initiated changes in providing care and in practicing medicine. Digital health is defined as "the cultural transformation of how disruptive technologies that provide digital and objective data accessible to both health care providers and patients leads to an equal-level doctor-patient relationship with shared decision-making and the democratization of care." As technological innovations become inseparable from health care, and as health care systems worldwide are becoming financially unsustainable, a paradigm shift is imminent [1]. The cultural component of this transformation implies that how the stakeholders of health care adopt or reject new technologies is more dependent on the outcomes than how the technologies progress.

In summary, using digital technology has become unavoidable in practicing medicine, and empowered patients, also known as e-patients, have needs that are different than what medical education prepares physicians for. Today's physicians are looking for their place, authority, and function in this status quo.

Patient Empowerment Has Been Booming

In the second half of the 20th century, the biomedical model of medicine has been replaced by the biopsychosocial paradigm [6]. The biomedical model states that the biological determinants are the main causes of diseases. On the other hand, in the third and fourth epidemiological periods, it has become obvious that psychological and social elements of disease are equally important in development of diseases. We should not only seek one cause, but also the complex interaction of predictors, triggers, and maintaining factors.

Later, this approach has been complemented by a digital component, thus making it the biopsychosocial-digital model. The digital component means that the digital expansion of the biological self, the engagement of technology, and the use of online networks are as notable as the other biopsychosocial factors [7].

The digital component could affect health outcomes in many ways. For example, portable devices support management of health and enable affordable access to people with low socioeconomic status and/or in remote environments [8-12]. It was shown in a systematic review that technology could also be used to reduce the disparity in melanoma incidence, mortality rates, and accessibility to posttreatment care management between urban and rural or remote populations [13]. Online social networking has a potential effect on health, for example, through social support; also, interactive information-sharing has an influence on patient health and health behavior.

The story of how Dave deBronkart—otherwise known as e-Patient Dave—used technologies to help in the treatment of his cancer shows the contribution patients can make to the complexities of medicine [14].

Moreover, in the 21st century, personalized medicine has become unavoidable in treating certain conditions, such as several types of cancer or diabetes [15,16]. The challenge is that physicians are required to come up with solutions tailored to each patient's needs instead of using treatment pathways of mass production.

Not only were physicians affected by the advent of the Internet and new technologies, but these technologies have also reshaped the lives and disease management of patients from the ground up. The e-patient movement came to life by raising issues and challenges that medical curricula do not address. The patient's reaction to changes in access to information is understandably to participate in the healing process [17].

E-patients are active in their care and demonstrate the power of the participatory medicine model. The "e" can stand for "electronic" (ie, uses digital technologies in their disease or health management), "equipped" (ie, has digital health technologies at their disposal), "enabled" (ie, has newly acquired access to information), "empowered" (ie, by the loss of the ivory tower), "engaged" (ie, taking an active part in their care), and "expert" (ie, in using technologies in their care or health management) [14,18].

We argue that patient empowerment, the spread of digital health, the biopsychosocial-digital approach, and the disappearance of the ivory tower of medicine lead to a new role for physicians. Instead of key holders to the ivory tower of medicine, they are slowly transforming into guides for their patients in the jungle of health care and digital information.

XSL•FO
RenderX

We also need to emphasize that knowledge can only potentially mean power. Therefore, it will definitely be challenging for medical professionals to adapt to not being an intermediary (ie, someone who consumes information and passes it on), but to become an apomediary (ie, someone who directs the patient to high-quality information and services) and, thus, stop being a prerequisite to obtaining information. This new approach also means that patients will not be labelled as such, but will be labelled as consumers, users, citizens, or persons who may already use the public resources of digital health [19,20].

We also propose that it is time to empower physicians in the same way patients have been empowered and to let them use their unique vision, knowledge, and insights to help make the best decisions for patients aided, not replaced, by advanced technologies. Thus, the era of digital health not only means to equip e-patients with information, tools, and technologies, but also to equip empowered physicians (e-physicians) with time, opportunity, and technologies to fulfil the modern vision of a practicing physician. Here we discuss the potential ways of facilitating this transition.

E-patients have become experts about their illnesses, while their chosen health care providers help them to be able to help themselves. This has sparked new expectations from patients, from monitoring and recording to sharing their data. As practicing medicine becomes a collaborative process, not only among health care professionals but also involving patients, the features of both e-patients and e-physicians become comparable.

The “e” in e-physicians can also stand for “electronic“ (ie, use digital technologies in their practice with ease), ”equipped“ (ie, have digital health technologies at their disposal), ”enabled“ (ie, by regulations and guidelines) [21,22], ”empowered“ (ie, by technologies that support their jobs and their e-patients), “engaged” (ie, need compassion and empathy to understand the feelings and points of view of patients, give relevant feedback, and involve them throughout the whole healing process), and ”expert“ (ie, in using technologies in their practice or know the best, most reliable, and trustworthy digital health sources and technologies) (see Table 1).

The doctor-patient relationship has been changing due to digital technology and the shared access to information. Insights about medical issues and the use of technology can now come from both sides. They are moving toward shared decision-making, communicating extensively, and managing health and disease through teamwork.

Digital health further offers the opportunity to make the job of being a medical professional rewarding and creative. While advanced technologies such as narrow and general artificial intelligence might seem to threaten replacement of physicians, they are more likely to support them and reduce the repetitive elements of their job that do not require the attention of a human mind. Thus, by adapting to the cultural changes initiated by digital health technologies, the characteristics of a physician-as-idol could shift from self-confident to curious, from rule follower to creative, and from lone hero to team worker (see Figure 1).

Certain skills are therefore crucial for e-physicians of the 21st century. Since there are more and more elderly patients struggling with chronic and polymorbid diseases, health care providers should be able to form an appropriate relationship with patients. Sufficient communication skills and the involvement of patients in prevention and treatment will become more important than ever. Adapting to constantly developing technologies is necessary and clinical skills should be improved with that in mind. The location of care has moved to the patient’s personal space (ie, home), which has been made possible by monitoring from afar with wearable sensors and portable diagnostic devices, among other technologies.

With the headway of telemedicine, new skills are needed regarding how to diagnose a patient and communicate with them without first a personal contact. Health care providers will need to be trained in such a way that they can diagnose, treat, educate, and monitor patients who are far away. A further improvement of this is the *hospital at home program*, which can allow more complicated treatments (eg, dialysis) to be available in the patient’s home, thus lowering the costs of hospital care [23].

Efficient teamwork is indispensable, since the development of science and technology makes it practically impossible for a healer to solve all challenges of a case on their own. As knowledge and treatment become more globalized, international research teams and the ability to work with clinical teams will be necessary [24].

Technology-focused professionals are also becoming a part of the team. New health care-related professions are going to emerge, such as clinical data scientists, medical software engineers, or digital medicine specialists [25].

The skills of managing, protecting, and orienting within datasets will also become irreplaceable. An e-physician will need to handle the information at hand in a critical and selective manner. The 21st-century healer will have to realize the ethical challenges created by digital health. For example, the way health-related data is collected, stored, accessed, and shared is an enormous privacy issue [24].

There are major factors that facilitate the transition of physicians from demigods to guides who enjoy their jobs. Examples include meaningful incentives proposed by hospitals, policy makers, and payers; a well-designed medical curriculum, including postgraduate education skills relevant to teaching; the wider availability of technologies; useful recommendations from peers; a rising number of evidence-based papers and guidelines; technologies that help save time and effort; and, generally, a good experience with e-patients [26-34] (see Textbox 1).

Among the many skills mentioned above, there are three cornerstones to this phenomenon that each e-physician should take into consideration: (1) the e-physician phenomenon means knowledge of, and positive attitude toward, digital technologies; (2) the e-physician phenomenon means the doctor-patient relationship will transition into a partnership; and (3) the e-physician phenomenon means that compassionate healing must remain the fundamental basis of health care.

XSL•FO
RenderX

Table 1. Summary of features of patient and medical professional empowerment.

Feature	Patient	Medical professional
Electronic	Uses digital technologies in their disease or health management	Uses digital technologies in their practice with ease
Equipped	Has digital health technologies at their disposal	Has digital health technologies at their disposal
Enabled	Enabled by their newly acquired access to information	Enabled by regulations and guidelines
Empowered	Empowered by the loss of the ivory tower	Empowered by technologies that support their job and e-patients
Engaged	Taking an active part in their care	Needs compassion and empathy to understand the feelings and points of view of patients, involving them throughout the whole healing process
Expert	Expert in the use of technologies in their care or health management	Expert in the use of technologies in their practice

Figure 1. Schematic view of the approaches, skills, and features of an empowered physician (e-physician).



XSL•FO
RenderX

Textbox 1. Summary of potential benefits and barriers of digital health adoption in medical practice, as well as potential actionable steps to address those barriers.

Potential benefits for physicians of digital health adoption:

- Improves treatment efficiency
- Saves time
- Increases patient satisfaction
- Increases patient safety
- Improves diagnostic ability
- Improves daily productivity
- Improves physician-patient relationship and communication
- Improves interpersonal communication with colleagues
- Can ease the burden of routine check-ups for chronic patients
- Cost-savings for the medical practice

External and internal barriers to digital health adoption:

- Lack of time
- Increasing workload
- Lack of resources and financial difficulties
- Lack of reimbursements
- Lack of knowledge about digital health technologies
- Lack of trained medical staff
- Increase in misuse and misunderstanding of digital health technologies by patients
- A rise of overdiagnosis
- An increase in health disparities due to limited health literacy
- Increased administrative tasks
- Troubled patient data privacy and security
- Resistance from physicians (eg, losing control)
- A work culture refusing innovation

Potential actionable steps to address barriers to digital health adoption:

- Evidence-based digital health solutions
- Practice guidelines
- Availability of special training
- Support by colleagues and work environment
- Improving the quality, safety, and effectiveness of digital health technologies
- Facilitating laws and regulations for proper usage
- Recommendations from peers
- Perceived usefulness, or relative advantage, and compatibility (ie, with work process)
- Ease of use and user-friendly interfaces
- Incentive structures
- Patients' positive attitudes and preferences regarding digital health solutions
- An innovation-oriented work culture

XSL•FO
RenderX

The Reality Behind the Rise of E-Physicians

At the moment, factors that prevent the transition to digital health adoption in medical practice seem to be outweighing the positive elements. There are only a handful of examples where hospitals, policy makers, and payers offer good incentives to improve the use of meaningful technologies. There is a serious lack of guidelines and policies. Health care is a complex system; many disruptive technologies are still too expensive to become widely available. In addition, there is a general reluctance of peers to adopt digital health, and there are even patients who do not wish to become empowered. For the above-mentioned changes to occur, the following factors are indispensable: strengthening professional competence and reshaping the medical curriculum.

As KR Sethuraman stated, "The physicians of tomorrow are taught by the teachers of today using the curriculum of the past" [35]. Obviously, medical education must include preparation for the digital era with evidence-based examples of curricula, such as the course *Lessons in Digital Health* at Semmelweis Medical School [36]. This is an open-access course available worldwide and shows examples of physicians who are masters of using digital health technologies, but only to allow themselves more time to listen to patients discuss their health issues with undisturbed empathy.

There are already positive examples available, as illustrated below, about practicing physicians who embody the image of the e-physician that this paper describes:

1. Dr Wendy Sue Swanson advocates for the use of social media to strengthen communication between health care providers and patients. She supports the idea that technologies can assist patients and their families in becoming stewards of their own health. She also launched a company to help other physicians learn to use online tools more effectively in helping patients make informed decisions based on scientific evidence [37].
2. Dr Jay Parkinson is the founder a primary care practice that also uses online tools and platforms for remote care. He has been building services that explore what the Internet means to health care delivery [38].
3. Dr Bryan Vartabedian is considered one of health care's most influential voices on social technology and medicine. He regularly expresses his views on patient-centricity, while also understanding medicine's emerging digital culture and how new media can be leveraged by organizations and individual stakeholders [39].
4. Dr Bas Bloem, a Dutch professor of neurology and Director of the Parkinson Center in Nijmegen, advocates for placing patients at the center of disease and health management and is a popular voice advocating for the use of new technologies [40].

Such e-physicians could serve as role models for young students who aspire to practice medicine but are afraid of the burden of IT issues, time management, and a huge workload. The skills and approaches that characterize this era of e-physicians, such as face-to-face communication skills, digital literacy, interdisciplinarity, knowing where to find information, translating large amounts of data into insights for patients, among others, should always have been at the core of practicing medicine. However, the economical, technological, and administrative burden of the profession has not made it possible for most physicians to enjoy the benefits of their training, individual capabilities, and creativity. By understanding how digital health technologies can support or augment their capabilities, physicians would have the chance to practice the art of medicine like never before.

Conflicts of Interest

None declared.

References

1. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]
2. Mesko B. Health IT and digital health: The future of health technology is diverse. J Clin Transl Res 2017;3(S3):431-434. [doi: 10.18053/jctres.03.2017S3.006]
3. Lee RT, Seo B, Hladkyj S, Lovell BL, Schwartzmann L. Correlates of physician burnout across regions and specialties: A meta-analysis. Hum Resour Health 2013 Sep 28;11:48 [FREE Full text] [doi: 10.1186/1478-4491-11-48] [Medline: 24074053]
4. Mulero A. Healthcare Dive. 2016 Jan 15. Bureaucracy tops causes of burnout among physicians URL: https://www.healthcaredive.com/news/bureaucracy-tops-causes-of-burnout-among-physicians/412227/ [accessed 2019-03-04] [WebCite Cache ID 76cklIt30]
5. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. mHealth adoption by healthcare professionals: A systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
6. Engel GL. The need for a new medical model: A challenge for biomedicine. Science 1977 Apr 08;196(4286):129-136. [Medline: 847460]
7. Ahmadvand A, Gatchel R, Brownstein J, Nissen L. The biopsychosocial-digital approach to health and disease: Call for a paradigm expansion. J Med Internet Res 2018 May 18;20(5):e189 [FREE Full text] [doi: 10.2196/jmir.9732] [Medline: 29776900]

XSL•FO
RenderX

8. Goodridge D, Marciniuk D. Rural and remote care: Overcoming the challenges of distance. Chron Respir Dis 2016 May;13(2):192-203 [FREE Full text] [doi: 10.1177/1479972316633414] [Medline: 26902542]
9. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The role of social network technologies in online health promotion: A narrative review of theoretical and empirical factors influencing intervention effectiveness. J Med Internet Res 2015 Jun 11;17(6):e141 [FREE Full text] [doi: 10.2196/jmir.3662] [Medline: 26068087]
10. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
11. 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 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
12. Ranney ML, Genes N. Social media and healthcare quality improvement: A nascent field. BMJ Qual Saf 2016 Dec;25(6):389-391. [doi: 10.1136/bmjqs-2015-004827] [Medline: 26658773]
13. Rollin A, Ridout B, Campbell A. Digital health in melanoma posttreatment care in rural and remote Australia: Systematic review. J Med Internet Res 2018 Sep 24;20(9):e11547 [FREE Full text] [doi: 10.2196/11547] [Medline: 30249578]
14. deBronkart D. How the e-patient community helped save my life: An essay by Dave deBronkart. BMJ 2013 Apr 02;346:f1990. [doi: 10.1136/bmj.f1990] [Medline: 23550048]
15. Mesko B. The role of artificial intelligence in precision medicine. Expert Rev Precis Med Drug Dev 2017 Sep 20;2(5):239-241. [doi: 10.1080/23808993.2017.1380516]
16. Dassau E, Zisser H, Harvey RA, Percival MW, Grosman B, Bevier W, et al. Clinical evaluation of a personalized artificial pancreas. Diabetes Care 2013 Apr;36(4):801-809 [FREE Full text] [doi: 10.2337/dc12-0948] [Medline: 23193210]
17. Riggare S. E-patients hold key to the future of healthcare. BMJ 2018 Dec 26;360:k846. [doi: 10.1136/bmj.k846] [Medline: 29483151]
18. Fox S. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2008 Aug 26. The engaged e-patient population URL: http://www.pewinternet.org/2008/08/26/the-engaged-e-patient-population/#people-turn-to-the-internet-for-health-information-when-the-stakes-are-high-and-the-connection-fast [accessed 2019-03-04] [WebCite Cache ID 76clJXZFy]
19. Eysenbach G. Medicine 2.0: Social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008 Aug 25;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
20. Dowie J, Kaltoft MK. The future of health is self-production and co-creation based on apomediative decision support. Med Sci (Basel) 2018 Sep;6(3):66 [FREE Full text] [doi: 10.3390/medsci6030066] [Medline: 30135365]
21. American Medical Association. 2018. Digital Health Implementation Playbook URL: https://www.ama-assn.org/amaone/ama-digital-health-implementation-playbook [accessed 2019-03-04] [WebCite Cache ID 76clXXejk]
22. A Coherent and Trustworthy Health Network for All: Digital Health Strategy 2018–2022. Copenhagen, Denmark: Ministry of Health, Denmark; 2018 Jan 21. URL: https://www.sum.dk/Aktuelt/Publikationer/~/media/Filer%20-%20Publikationer_i_pdf/English/2018/A-coherent-and-trustworthy-health-network-for-all-jan-2108/A-coherent-and-trustworthy-health-network-jan-2018.pdf [accessed 2019-03-04] [WebCite Cache ID 76clpkxqh]
23. Wallace EL, Rosner MH, Alscher MD, Schmitt CP, Jain A, Tentori F, et al. Remote patient management for home dialysis patients. Kidney Int Rep 2017 Nov;2(6):1009-1017 [FREE Full text] [doi: 10.1016/j.ekir.2017.07.010] [Medline: 29634048]
24. Lerner I, Veil R, Nguyen DP, Luu VP, Jantzen R. Revolution in health care: How will data science impact doctor-patient relationships? Front Public Health 2018;6:99 [FREE Full text] [doi: 10.3389/fpubh.2018.00099] [Medline: 29666789]
25. Toole B. Book review. The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care. J Technol Hum Serv 2013;31(3):294-296. [doi: 10.1080/15228835.2013.796303]
26. Gagnon MP, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. J Med Syst 2012 Feb;36(1):241-277 [FREE Full text] [doi: 10.1007/s10916-010-9473-4] [Medline: 20703721]
27. Bourla A, Ferreri F, Ogorzelec L, Peretti CS, Guinchard C, Mouchabac S. Psychiatrists' attitudes toward disruptive new technologies: Mixed-methods study. JMIR Ment Health 2018 Dec 14;5(4):e10240 [FREE Full text] [doi: 10.2196/10240] [Medline: 30552086]
28. Hazzam J, Lahrech A. Health care professionals' social media behavior and the underlying factors of social media adoption and use: Quantitative study. J Med Internet Res 2018 Nov 07;20(11):e12035 [FREE Full text] [doi: 10.2196/12035] [Medline: 30404773]
29. Duhm J, Fleischmann R, Schmidt S, Hupperts H, Brandt SA. Mobile electronic medical records promote workflow: Physicians' perspective from a survey. JMIR Mhealth Uhealth 2016 Jun 06;4(2):e70 [FREE Full text] [doi: 10.2196/mhealth.5464] [Medline: 27268720]
30. Saigi-Rubió F, Jiménez-Zarco A, Torrent-Sellens J. Determinants of the intention to use telemedicine: Evidence from primary care physicians. Int J Technol Assess Health Care 2016 Jan;32(1-2):29-36. [doi: 10.1017/S0266462316000015] [Medline: 27472158]

XSL•FO
RenderX

31. Grünloh C, Cajander Å, Myreteg G. "The record is our work tool!": Physicians' framing of a patient portal in Sweden. J Med Internet Res 2016 Dec 27;18(6):e167 [FREE Full text] [doi: 10.2196/jmir.5705] [Medline: 27349531]
32. Kooij L, Groen WG, van Harten WH. Barriers and facilitators affecting patient portal implementation from an organizational perspective: Qualitative study. J Med Internet Res 2018 May 11;20(5):e183 [FREE Full text] [doi: 10.2196/jmir.8989] [Medline: 29752253]
33. Miller DP, Latulipe C, Melius KA, Quandt SA, Arcury TA. Primary care providers' views of patient portals: Interview study of perceived benefits and consequences. J Med Internet Res 2016 Jan 15;18(1):e8 [FREE Full text] [doi: 10.2196/jmir.4953] [Medline: 26772771]
34. Digital Health Study: Physicians' Motivations and Requirements for Adopting Digital Clinical Tools. Chicago, IL: American Medical Association; 2016. URL: https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/specialty%20group/washington/ama-digital-health-report923.pdf [accessed 2019-03-05] [WebCite Cache ID 76eyLdMXN]
35. Bhopal A. Attitude and the 21st century doctor. Lancet Glob Health 2015 Mar;3(3):e126-e127 [FREE Full text] [doi: 10.1016/S2214-109X(14)70376-6] [Medline: 25701989]
36. The Medical Futurist. 2017 Dec 05. Let medical students become 21st century physicians URL: https://medicalfuturist.com/let-medical-students-become-21st-century-physicians [accessed 2019-03-04] [WebCite Cache ID 76cnfmWdN]
37. Swanson WS. Seattle Mama Doc. 2019 Jan 29. Washington State measles outbreak: What to do for infants and children URL: https://seattlemamadoc.seattlechildrens.org/ [accessed 2019-03-04] [WebCite Cache ID 76cnjokYr]
38. Jay Parkinson MD, MPH. URL: https://blog.jayparkinsonmd.com/ [accessed 2019-03-04] [WebCite Cache ID 76cnn1MRl]
39. 33 charts. Bryan Vartabedian, M.D. FAAP URL: https://33charts.com/bryan-vartabedian-md/ [accessed 2019-03-04] [WebCite Cache ID 76cnpdzJs]
40. Radboudumc. Bas Bloem MD PhD URL: https://www.radboudumc.nl/en/people/bas-bloem [accessed 2019-03-04] [WebCite Cache ID 76cnseoS6]

Abbreviations

e-patient: empowered patient
e-physician: empowered physician

Edited by G Eysenbach; submitted 13.10.18; peer-reviewed by J Dowie, L Lenzo; comments to author 23.12.18; revised version received 11.01.19; accepted 20.01.19; published 26.03.19.

Please cite as:
Mesko B, Győrffy Z
The Rise of the Empowered Physician in the Digital Health Era: Viewpoint
J Med Internet Res 2019;21(3):e12490
URL: http://www.jmir.org/2019/3/e12490/
doi:10.2196/12490
PMID:30912758

©Bertalan Mesko, Zsuzsa Győrffy. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Internet-Based Cognitive Behavioral Therapy in Stepped Care for Chronic Fatigue Syndrome: Randomized Noninferiority Trial

Margreet Worm-Smeitink[1,2], MSc; Anthonie Janse[3], MSc; Arno van Dam[4], PhD; Andrea Evers[5,6], Prof Dr; Rosalie van der Vaart[5], PhD; Michel Wensing[7,8], Prof Dr; Hans Knoop[3,9], Prof Dr

[1]Department of Medical Psychology, Radboud University Medical Center, Nijmegen, Netherlands

[2]Specialist Center for Complex Medically Unexplained Symptoms and Somatic Symptom Disorders, Dimence, Deventer, Netherlands

[3]Department of Medical Psychology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

[4]School of Social and Behavioural Sciences, Tranzo, Tilburg University, Tilburg, Netherlands

[5]Health, Medical and Neuropsychology Unit, Institute of Psychology, Leiden University, Leiden, Netherlands

[6]Department of Psychiatry, Leiden University Medical Center, Leiden, Netherlands

[7]Department of General Practice and Health Services Research, Heidelberg University Hospital, Heidelberg, Germany

[8]Radboud Institute of Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands

[9]Expert Center for Chronic Fatigue, Department of Medical Psychology, Amsterdam University Medical Centers, Vrije Universiteit, Amsterdam, Netherlands

Corresponding Author:
Hans Knoop, Prof Dr
Department of Medical Psychology
Amsterdam University Medical Centers
University of Amsterdam
PO Box 22660
Amsterdam, 1100DD
Netherlands
Phone: 31 205668506
Email: hans.knoop@amc.uva.nl

Abstract

Background: Internet-based cognitive behavioral therapy (I-CBT) leads to a reduction of fatigue severity and disability in adults with chronic fatigue syndrome (CFS). However, not all patients profit and it remains unclear how I-CBT is best embedded in the care of CFS patients.

Objective: This study aimed to compare the efficacy of stepped care, using therapist-assisted I-CBT, followed by face-to-face (f2f) cognitive behavioral therapy (CBT) when needed, with f2f CBT (treatment as usual [TAU]) on fatigue severity. The secondary aim was to investigate treatment efficiency.

Methods: A total of 363 CFS patients were randomized to 1 of the 3 treatment arms (n=121). There were 2 stepped care conditions that differed in the therapists' feedback during I-CBT: *prescheduled* or *on-demand*. When still severely fatigued or disabled after I-CBT, the patients were offered f2f CBT. Noninferiority of both stepped care conditions to TAU was tested using analysis of covariance. The primary outcome was fatigue severity (Checklist Individual Strength). Disabilities (Sickness Impact Profile -8), physical functioning (Medical Outcomes Survey Short Form-36), psychological distress (Symptom Checklist-90), and proportion of patients with clinically significant improvement in fatigue were the secondary outcomes. The amount of invested therapist time was compared between stepped care and TAU. Exploratory comparisons were made between the stepped care conditions of invested therapist time and proportion of patients who continued with f2f CBT.

Results: Noninferiority was indicated, as the upper boundary of the one-sided 98.75% CI of the difference in the change in fatigue severity between both forms of stepped care and TAU were below the noninferiority margin of 5.2 (4.25 and 3.81, respectively). The between-group differences on the secondary outcomes were also not significant (P=.11 to P=.79). Both stepped care formats required less therapist time than TAU (median 8 hours, 9 minutes and 7 hours, 25 minutes in stepped care vs 12 hours in TAU; P<.001). The difference in therapist time between both stepped care formats was not significant. Approximately half of the patients meeting step-up criteria for f2f CBT after I-CBT did not continue.

XSL•FO
RenderX

Conclusions: Stepped care, including I-CBT followed by f2f CBT when indicated, is noninferior to TAU of f2f CBT and requires less therapist time. I-CBT for CFS can be used as a first step in stepped care.

Trial Registration: Nederlands Trial Register NTR4809; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=4809 (Archived by WebCite at http://www.webcitation.org/74SWkw1V5)

(J Med Internet Res 2019;21(3):e11276) doi:10.2196/11276

KEYWORDS

eHealth; chronic fatigue syndrome; cognitive behavioral therapy; randomized controlled trial

Introduction

Chronic fatigue syndrome (CFS), sometimes referred to as myalgic encephalomyelitis (ME), is characterized by medically unexplained, severe, ongoing, and disabling fatigue. It is not alleviated by rest, and according to the revised US Centers for Disease Control (CDC) criteria from 2003, it is accompanied by at least 4 out of the following 8 additional symptoms: postexertional malaise, headache, unrefreshing sleep, muscle pain, joint pain, sore throat, tender lymph nodes, and impaired concentration or memory [1,2]. The prevalence of CFS is estimated to be around 1% [3] and the prognosis is unfavorable without treatment; the median spontaneous recovery rate is 5% [4]. It is unknown what causes CFS, and it is commonly assumed to be multifactorially determined.

It is helpful to distinguish between precipitating and perpetuating factors of CFS [5]. Precipitating factors might be a virus infection or a stressful life event that triggers severe fatigue. Perpetuating factors are thought to maintain the fatigue, even when the precipitating factor is no longer present. According to the cognitive behavioral model of CFS, these perpetuating factors are behavior and beliefs [6,7]. Cognitive behavioral therapy (CBT) aims to change these beliefs and behaviors and is found to lead to a significant reduction of fatigue and disability [8-10].

CBT for CFS is a time-intensive treatment, requiring 12 hours of therapist contact on average [10]. This is a problem, as treatment capacity is limited [11]. A possible solution is using internet-based CBT (I-CBT). I-CBT often takes less therapist time to deliver, increasing the number of patients that can be treated [12-14]. I-CBT has a number of other advantages over face-to-face (f2f) CBT. It reduces the traveling time and the need to schedule appointments, which relieves some of the burden of treatment for patients [13]. The treatment is accessible whenever patients want [12,13], which may be empowering [14]. Furthermore, I-CBT might increase motivation, as it offers a wide range of attractive audiovisual information elements and the possibility to receive feedback on the progress made [13].

In the past years, I-CBT was developed for several mental disorders and chronic medical conditions, and its efficacy has been tested in randomized controlled trials (RCTs) [12-14]. Systematic reviews seem to indicate that guided I-CBT can reach effect sizes equivalent to those found in f2f CBT [15,16]. However, I-CBT was not often directly compared with a more traditional, high-intensity f2f CBT. Additionally, I-CBT is not yet available for all medical and psychiatric conditions.

We developed I-CBT for adults with CFS and recently tested its efficacy in an RCT. I-CBT led to a reduction of fatigue and disabilities compared with a waiting list [17]. Approximately 40% of the patients had a clinically significant and reliable change in fatigue severity and were no longer severely fatigued after I-CBT. To gain more insight into the role of therapist feedback in I-CBT, 2 delivery formats of I-CBT were tested, which differed in therapist guidance. In the *protocol-driven feedback* format, the therapists' feedback was scheduled with preset time intervals. The therapist sent reminders if the schedule was not adhered to. In the *feedback on demand* format, feedback was only given when the patient asked for it. Both the I-CBT formats were equally effective, but the feedback-on-demand version required significantly less therapist time [17].

It is important to uncover how I-CBT can best be embedded in clinical care for CFS. The effect size of I-CBT on fatigue severity was smaller than the effect size previously found in a study investigating the efficacy of f2f CBT delivered in groups, using the same treatment principles (effect sizes 0.6 and 1.1, respectively) [17,18]. This suggests that I-CBT is less effective than f2f CBT. A straightforward solution would be to blend I-CBT with f2f CBT. Unfortunately, there is a lack of data to conclude that combinations of internet-based and f2f CBT are as effective as established f2f treatments. In addition, it is not known what an optimal format of blending of both treatment formats would be (eg, nonsequential vs sequential) [19]. We chose to embed I-CBT in a stepped care approach, as a first step. Patients who did not profit from I-CBT, as indicated by severe fatigue or disability still present following I-CBT, would step up to f2f CBT. If stepped care is as effective as treatment as usual (TAU; ie, f2f CBT), this would imply that I-CBT can be used in routine clinical care for CFS. Treatment intensity would then be decreased for the substantial group of patients for which I-CBT suffices. Furthermore, stepped care may be more time-efficient than TAU, in that less therapist time would be needed to deliver stepped care [20]. This could increase cost-effectiveness.

The primary objective of this study was to determine whether stepped care for CFS, with I-CBT as the first and f2f CBT as the second step, would be as effective as TAU. This randomized noninferiority trial was a follow-up study of the RCT investigating the efficacy of I-CBT for adults with CFS. All patients who were still severely fatigued or functionally impaired after I-CBT or the waiting list were offered f2f CBT, resulting in 2 stepped care conditions and 1 TAU arm. The efficacy in reduction of fatigue and efficiency in reduction of therapist time needed for stepped care were compared with TAU.

XSL•FO
RenderX

Moreover, we explored the role of therapist guidance in I-CBT, as it is not well understood how therapist feedback influences I-CBT. Guided I-CBT was found to be more effective than unguided I-CBT [21-23], but it is not known how much guidance is needed, who needs to provide the guidance [23,24], and what aspect of the guidance is (most) helpful [22]. In our previous study, we found that whether therapist feedback was *prescheduled* or *on-demand* did not influence the treatment effect [17]. We explored in this study how therapist guidance during I-CBT influenced the outcome of stepped care. More specifically, (1) whether the difference in therapist time needed between the I-CBT arms persisted in stepped care and (2) whether the proportion of patients stepping up to f2f CBT after I-CBT differed in both formats, as the format of therapist feedback may influence the willingness to step up to f2f CBT.

Methods

Trial Design

This study was registered in the Dutch trial register (NTR4809) and was approved by the medical ethical committee of the Radboud University Medical Center (reference NL42543.091.12). It was a follow-up study of an RCT testing the efficacy of I-CBT for CFS (NTR4013) and was designed as a three-arm, parallel, randomized, noninferiority trial. In 2 arms, the patients received stepped care (SC) consisting of I-CBT, either with protocol-driven feedback (SC-protocol-driven feedback) or with feedback on demand (SC-feedback-on-demand), followed by f2f CBT when necessary, that is, still severely fatigued (CIS fatigue severity >35) or disabled (SIP >700) after I-CBT. The third arm was f2f CBT after a variable waiting period (TAU).

Before randomization, all patients completed a baseline assessment (T0). Directly after randomization, the patients in the stepped care arms started with I-CBT for a duration of 6 months. After 6 months, they completed a second assessment (T1). If they started additional f2f CBT after T1, they were assessed again 6 months later (T2). Patients from the TAU group were placed on the waiting list directly after T0, for a duration of maximally 6 months (refer to the section Interventions), after which they completed their T1 assessment. Directly after T1, the f2f CBT was started. Furthermore, 6 months after T1, the patients were assessed again (T2). Assessments were web-based questionnaires.

Participants

Participants were all recruited from adult CFS patients consecutively referred to a tertiary CFS treatment center at a university hospital. They comprised the 240 participants of the RCT testing the efficacy of I-CBT [17] and 123 participants additionally randomized to gain sufficient power for the primary research question of this study. It was a closed study; only patients referred to the treatment center could participate. All patients were first seen by a consultant of the outpatient clinic of the department of Internal Medicine of the hospital. The consultant checked the medical status to confirm that patients were sufficiently examined and if not, they examined the patients according to the national CFS guidelines [25]. Patients meeting CDC criteria for CFS [1,2] underwent a clinical assessment at the treatment center, including a structured interview (the Mini-International Neuropsychiatric Interview [26]), to rule out psychiatric disorders that could explain the presence of fatigue. Comorbidities that could not explain fatigue were not exclusion criteria.

Eligible patients were informed about the study and were included after giving written informed consent. They were asked to refrain from seeking treatment for CFS elsewhere for the duration of the study. Inclusion criteria were being aged 18 years or above; ability to speak, read, and write Dutch; meeting the CDC criteria, revised in 2003, for CFS [1,2]; severely fatigued (checklist Individual strength subscale, fatigue severity score of ≥35 [27]); severely disabled (Sickness Impact Profile 8, total score of ≥700 [28]); able to use a computer and have access to the internet; and given written informed consent. Exclusion criteria were being involved in a legal procedure concerning disability/benefit claims and/or participating in other CFS research [29].

Interventions

Treatment as Usual (Direct Face-to-Face Cognitive Behavioral Therapy)

For patients in the TAU group, the therapy started after a waiting period. During the first part of the study (first 240 randomized patients), the waiting list was 6 months for all patients. In the second part of the study (last 123 patients), the waiting period fluctuated depending on treatment capacity but was not longer than 6 months. For ethical reasons, patients could start sooner if the waiting period for routine clinical care was less than 6 months.

CBT was delivered according to a treatment protocol [30] that was also used in previous RCTs [18,20]. The treatment aims at changing fatigue-related beliefs and behavior. CBT starts with educating the patients on the cognitive-behavioral model of CFS and formulating treatment goals which, when attained, imply recovery from CFS. Following this, the patient learns to adopt a regular sleep-wake cycle, with fixed bed times and without sleep during the day. The patient is taught to shift attention away from fatigue and to challenge unhelpful beliefs regarding fatigue and disability. This is followed by a graded activity program in which a systematic increase in physical activity, regardless of symptoms, usually by walking or cycling, is introduced. Patients are encouraged to challenge dysfunctional beliefs about symptoms and activity during the program. The graded activity program is tailored to the activity pattern of the patient: relatively active patients learn to spread activity evenly first, whereas less active patients start directly with graded activity. An actometer, a motion-sensing device, was used to determine the activity pattern [29,31]. After the increase in physical activity, the patients learn to use these principles to increase mental and social activity. After the graded activity program, the patient is taught to reach his personal goals step by step. Finally, the patient is encouraged to experiment with fluctuating bedtimes and activity levels and to adopt a healthy view on normal fatigue.

XSL•FO
RenderX

Stepped Care

The I-CBT was accessible via a website, with a username and password that the patient received by email. Patients did not have to pay to use the intervention, although internet access was not provided. Patients were not trained to use the platform and could ask for support by email. During the f2f diagnostic sessions, the patient had met the therapist who would deliver the I-CBT. Both versions of I-CBT were based on the treatment protocol for f2f CBT for CFS [17,30]. The conditions differed in when the patient received feedback. In SC protocol-driven feedback, during I-CBT, the therapist asked the patient to report on the progress made, on fixed time points: weekly in the first 4 weeks and fortnightly in the following 8 weeks. From week 13, the frequency could be lowered to once every 3 weeks, if enough progress was being made. The therapist gave feedback via email and sent a reminder if no update was received. The feedback was aimed at motivating the patient to put the instructions of the intervention into practice. In the *feedback-on-demand* format, patients received feedback only when they indicated a need for advice. No reminders were sent.

The I-CBT consisted of 7 modules: (1) In "getting started and goal setting" psychoeducation is given, a treatment contract is signed, and goals are set. When goals are submitted, the next 5 modules are accessible: (2) "regulate sleep-wake cycle," (3) "helpful beliefs about fatigue," (4) "how to communicate with others about CFS," (5) "gradually increasing my activity," and (6) "reaching my goals step by step." When this module is completed, the last module becomes accessible: (7) "evaluation and the future." The modules are described in more detail elsewhere [29].

Within 2 weeks after T1, patients had an f2f evaluation session with their therapist. During this session, it was evaluated whether treatment goals were reached or additional f2f CBT was indicated. Patients were offered f2f CBT after I-CBT if they were still severely fatigued (CIS fatigue severity ≥35) and/or severely disabled (SIP8 ≥700) and/or if not all therapy goals were attained and the therapist expected that additional treatment was necessary to attain them. This decision could only be made after consultation of an experienced supervising CBT therapist (HK). The reason for continuation was, in that case, recorded and reported.

The additional CBT was delivered according to the treatment protocol [30]. It evaluated which beliefs and behaviors were already changed and consequently, which interventions were still needed to make further progress. Therapists were trained to encourage patients to step up when still fatigued or disabled after I-CBT. They were instructed to avoid that patients conclude that the I-CBT failed and additional f2f CBT will not lead to a (further) reduction of symptoms. The therapists helped patients appreciate what was already achieved and clarify that the f2f CBT is tailored to what is needed to make further gains.

Therapists in all the 3 arms were psychologists, trained in CBT for CFS, both f2f and internet-based. Therapists received weekly group supervision during the study [29].

Measures

Baseline Characteristics

The following patient characteristics were recorded: age, sex, duration of symptoms, the presence of each CDC symptom, work status, years of education followed, and the presence of depressive symptoms above a clinical significant cutoff of 4 on the Beck Depression Inventory, primary care version [32]. Furthermore, it was recorded whether patients met the SEID (systemic exertion intolerance disease) criteria [33]. SEID was recently proposed by the U.S. Institute of Medicine as an alternative to ME/CFS. SEID was met when postexertional malaise, unrefreshing sleep, and memory/concentration problems were all reported [33]. SEID can also be confirmed when a patient has orthostatic intolerance instead of memory/concentration problems, aside from severe and persistent fatigue. However, orthostatic intolerance was not assessed.

Primary Outcome Measure: Fatigue Severity

Fatigue severity (the primary, noninferiority outcome) was measured with the Checklist Individual Strength (CIS) fatigue severity subscale [27,34]. The CIS contains 20 items, Likert scaled (from 1-7) assessing 4 aspects of fatigue. The fatigue severity subscale is often used as a measure of fatigue in studies on CBT for CFS and is reliable and valid [27]. It contains 8 items, scores range from 8 (not fatigued) to 56 (severely fatigued). A validated cutoff of 35 was used to indicate severe fatigue [35].

Secondary Outcome Measures

Disabilities

A total score of Sickness Impact Profile 8 (SIP8; [28]) measures disabilities at 8 domains of daily functioning. Patients can indicate which out of 86 statements apply to their functioning as a result of their health status. Each statement has a weighting factor indicating severity [28]. The weighed total score was used. Higher scores indicated more severe disabilities. In this study, a total score of ≥700 was used as a cutoff to indicate significant disability.

Physical Functioning

The Medical Outcome Survey Short Form-36 (SF-36; [36,37]) physical functioning subscale was used to measure self-reported physical functioning and ranged from 0 (maximum limitations) to 100 (no limitations).

Psychological Distress

The Symptom Checklist-90 (SCL-90 [38]) total scale score measures psychological distress with 90 items, answered on a 5-point Likert scale (range of 90-450). Higher scores indicate more distress.

Invested Therapist Time

Therapists recorded the time needed per patient. Time spent on I-CBT (in stepped care), on additional telephone calls or email contact (all conditions) were recorded in minutes. Each f2f CBT session counted for 60 min (all conditions). In all conditions, 120 min was counted for the diagnostic assessment, consisting of 2 sessions.

XSL•FO
RenderX

Sample Size

Assuming a power of 0.80, a one-sided alpha of .0125 (correcting for 2 comparisons), an SD of 13.6 [20], and a noninferiority limit of 5.2 units on the CIS, 108 patients were needed per arm. To account for an expected dropout rate of 10.7%, the sample size needed for each arm was 121, making the total number of patients needed to 363. Therefore, 123 patients were to be randomized in addition to the 240 who were already randomized. The dropout rate of 10.7% was found in a study investigating stepped care for CFS with a self-help booklet as the first step [20]. A one-sided interval was used, as we tested for noninferiority. The 5.2 noninferiority margin is the estimated average decrease on the CIS fatigue severity subscale that occurs during waiting list, which is assumed to be clinically nonsignificant [39].

Randomization and Blinding

Patients were randomly allocated to 1 of the 3 conditions. Randomization was computer-generated, in blocks of 12 patients. The randomization program was programmed by a statistician, not involved in this study. Randomization was performed by an administrative assistant, in the presence of the patient and the therapist. Participants were partly blinded: they were unaware of the existence of 2 I-CBT formats and were told that they could either receive I-CBT followed by CBT if needed or f2f CBT. After randomization, therapists and patients could read the result from the computer screen: "(1) internet therapy" and "(2) internet therapy" or "wait list." The therapist knew that condition "1" of I-CBT was protocol-driven feedback and "2" was feedback on demand.

Statistical analysis was performed on a data file, which blinded the researcher performing the analysis to patient and allocation condition. Post analysis, allocation to condition was unmasked to enable the authors to interpret the results.

Statistical Analyses

Analyses were done on the basis of intention to treat after imputation of missing primary and secondary outcome measures at postassessment. Postassessment was T1 for patients who stopped after waiting list or after I-CBT and was T2 for patients who received (additional) f2f CBT. When T2 data were needed but only T1 was present, postassessment scores were imputed as well. We used multiple imputation (20 imputed data sets), assuming that data were missing at random. All outcome variables at postassessment were included in the imputation model and were imputed. Baseline variables of all outcomes were only entered as predictors for the imputation model and were not imputed. Imputation was done in IBM SPSS version 22, as were all statistical analyses.

Outcomes of Both Versions of Stepped Care Compared With Direct Face-to-Face Cognitive Behavioral Therapy

To answer the primary research question, an analysis of covariance (ANCOVA) was performed with postassessment CIS fatigue severity score as the dependent variable, baseline fatigue scores as covariate, and treatment condition as the fixed factor. Noninferiority was assumed when post-treatment CIS fatigue severity was maximally 5.2 points higher for either form of stepped care in comparison with TAU [20]. Hence, the upper bound of the one-sided 98.75% CI of the difference between the formats should be no larger than 5.2. To compare the effect on disabilities, physical functioning, and psychological distress of stepped care and TAU, ANCOVAs were performed for SF-36, SIP8, and SCL-90, with baseline of the dependent variable as covariate and condition as fixed factor.

Differences in the proportion of patients with clinically significant improvement in fatigue severity were compared with chi-Square tests. Each stepped care group was compared with TAU. Clinically significant improvement in fatigue severity was assumed when there was a statistically reliable change of >1.96 SD in combination with a CIS fatigue scores of <35 on postassessment. The reliable change index (RCI) of the CIS fatigue severity was calculated for each person following Jacobson and Truax [40]. For each condition, the SD used in the formula was the SD of the baseline CIS fatigue severity score in the specific condition. The reliability of the CIS used in the calculation was 0.88 [34]. An RCI larger than 1.96 indicates that with a CI of 95%, it can be assumed that the improvement in CIS fatigue severity represents a true change and is not the consequence of the unreliability of the measure.

In addition, 3 sensitivity analyses were performed. First, the noninferiority was tested, whereas missing observations on fatigue severity were not imputed with multiple imputations but replaced in the following manner: (1) it was hypothesized that patients in the TAU group improved, missing scores were replaced by the CIS fatigue severity score at T0 minus the mean change in fatigue of the TAU group and (2) it was hypothesized that patients of the stepped care groups deteriorated, scores were replaced by the maximum CIS fatigue severity score. Second, the main analysis was repeated excluding patients who received CBT by telephone or email, instead of f2f. Third, the main analysis was repeated selecting only patients who met the SEID criteria.

Therapist Time Invested

Therapist time of both stepped care conditions were compared with TAU (either means with *t* tests or medians with Mann-Whitney *U* tests, when not normally distributed). The analysis was done with 2 selections: (1) including all patients: patients who did not start therapy only accounted for 120 min for the diagnostic assessment sessions, (2) including only "starters": for the stepped care conditions, this was defined as having logged in the I-CBT program for at least three times and having submitted treatment goals [29] or having started f2f CBT. For TAU, it was defined as having received at least one therapy session. For patients who were still in treatment at the end of the study, the therapist time and therapy duration until that point were used. This was at least 1 year after randomization and at least 6 months after the start of f2f CBT. In addition, therapist time spent in both stepped care formats was compared.

Additional Effect Face-to-Face Cognitive Behavioral Therapy After Internet-Based Cognitive Behavioral Therapy

For patients in the stepped care groups who received additional f2f CBT, whether treatment outcomes at T2 (after f2f CBT)

XSL•FO
RenderX

differed significantly from T1 (after I-CBT) was investigated using paired-samples *t* tests.

Outcome Differences Between First 240 and Last 123 Randomized Patients

This analysis compared 2 groups of patients. The first 240 randomized patients were included between April 2013 and June 2015 [17] and the second 123 subsequently, between June 2015 and December 2016. Both took place in different circumstances; during the second part of the study, the treatment center was unexpectedly moved to another university medical center, which resulted in therapist changes and increased travel time for patients. Treatment outcome for all measures were compared with *t* tests.

Results

Overview

Patients were recruited between April 2013 and December 2016 and data collection was completed in December 2017. As shown in the flowchart (Figure 1), out of 766 patients screened for eligibility, 363 patients were randomly assigned to 3 treatment arms. Of these, 7 randomized patients did not meet the inclusion criteria, as the number of additional CDC criteria was less than 4.

In the TAU condition, the waiting list duration was 29 weeks (SD 4) for the first part of the study and 18 weeks (SD 12) for the second part of the study. The Spearman rank correlation between waiting list duration and reduction of fatigue was r_s=0.075, *P*=.50. After waiting list, 5 patients were lost to follow-up. Moreover, 12 patients did no longer meet the step-up criteria (severely fatigued as indicated by CIS fatigue severity >35 and/or severely disabled indicated by SIP8>700). In addition, 3 of them started individual CBT. Out of the remaining 104 patients, 77 started f2f CBT and 3 received CBT via telephone/email.

In the SC-protocol-driven feedback condition, 116 out of the 121 patients started I-CBT (95.8%). At T1, 87 patients met the step-up criteria (71.9%). Of them, 37 (43%) started f2f CBT. Furthermore, 3 patients, who did not meet the step-up criteria, received f2f CBT because several goals were not reached. In addition, 1 patient received the additional CBT by telephone and email.

In the SC feedback-on-demand condition, 113 patients started I-CBT (93.3%). Of the 85 patients who met the step-up criteria (70.2%), 48 (56%) started f2f CBT. Moreover, 6 patients who did not meet the step-up criteria received f2f CBT because not all treatment goals were attained.

The difference in the proportion of patients meeting the step-up criteria that did actually step up to f2f CBT between both stepped care conditions was not significant (χ^2_1=3.3; *P*=.07). Reasons for not stepping up were registered (see Figure 1): in the SC-protocol-driven feedback format, 29 (57%) did not want CBT anymore (15 because of the nature of the intervention, eg, did not want to travel, 14 could not fit f2f CBT in their lives at that moment eg, because of pregnancy), 10 (20%) were satisfied with the result of I-CBT, 8 (16%) no longer viewed CFS as the main problem, and 4 (8%) started treatment elsewhere. In SC-feedback-on-demand, 22 patients (56%) did not want CBT (17 because of the nature of the intervention and for 5 it did not fit into their lives), 9 (23%) were satisfied with the result of I-CBT, 6 (15%) no longer viewed CFS as the main problem, and 2 (5%) started treatment elsewhere.

At the end of the study, 1 patient was still in therapy and completed the postassessment 6 months after the start of f2f CBT. None of the patients were excluded from the main analysis. Table 1 shows the baseline characteristics.

Outcomes of Stepped Care Compared With Treatment As Usual

Noninferiority With Respect to Effect on Fatigue Severity

Data on the primary outcome were missing for 6 patients (all in the TAU condition). The upper boundary of the one-sided 98.75% CI of both forms of stepped care was below 5.2, indicating noninferiority. The mean difference in fatigue severity for the SC-protocol-driven feedback condition with TAU was −0.04 points; the CI upper bound was 3.81. For the SC-feedback-on-demand condition, the difference was 0.41 points; the upper bound of the CI was 4.25.

In all sensitivity analyses, both stepped care conditions remained noninferior to TAU (upper boundary CI for SC-protocol-driven feedback ranging from 3.34-3.95; upper bound CI SC-feedback-on-demand ranging from 4.12-4.50).

Secondary Outcome Measures

There were no significant differences between either stepped care condition or TAU for all secondary outcomes, that is, level of disabilities (SIP8), physical functioning (SF-36 physical functioning), and psychological distress (SCL-90). Within-group effect sizes are provided in Table 2.

Proportion of Patients With Clinically Significant Improvement in Fatigue Severity

Post stepped care, 49 out of 121 (40%) patients in SC-Protocol-driven feedback during I-CBT showed clinically significant improvement in fatigue severity. In the TAU group, 53 out of 115 patients improved. Assuming that patients with missing data did not improve, this was 44%. The difference between both groups was not significant; χ^2_1(N=236)=0.7; *P*=.39.

In SC-feedback-on-demand, 61 patients improved (50%). The difference between SC-feedback-on-demand and TAU was also not significant: χ^2_1(N=236)=0.4; *P*=.51.

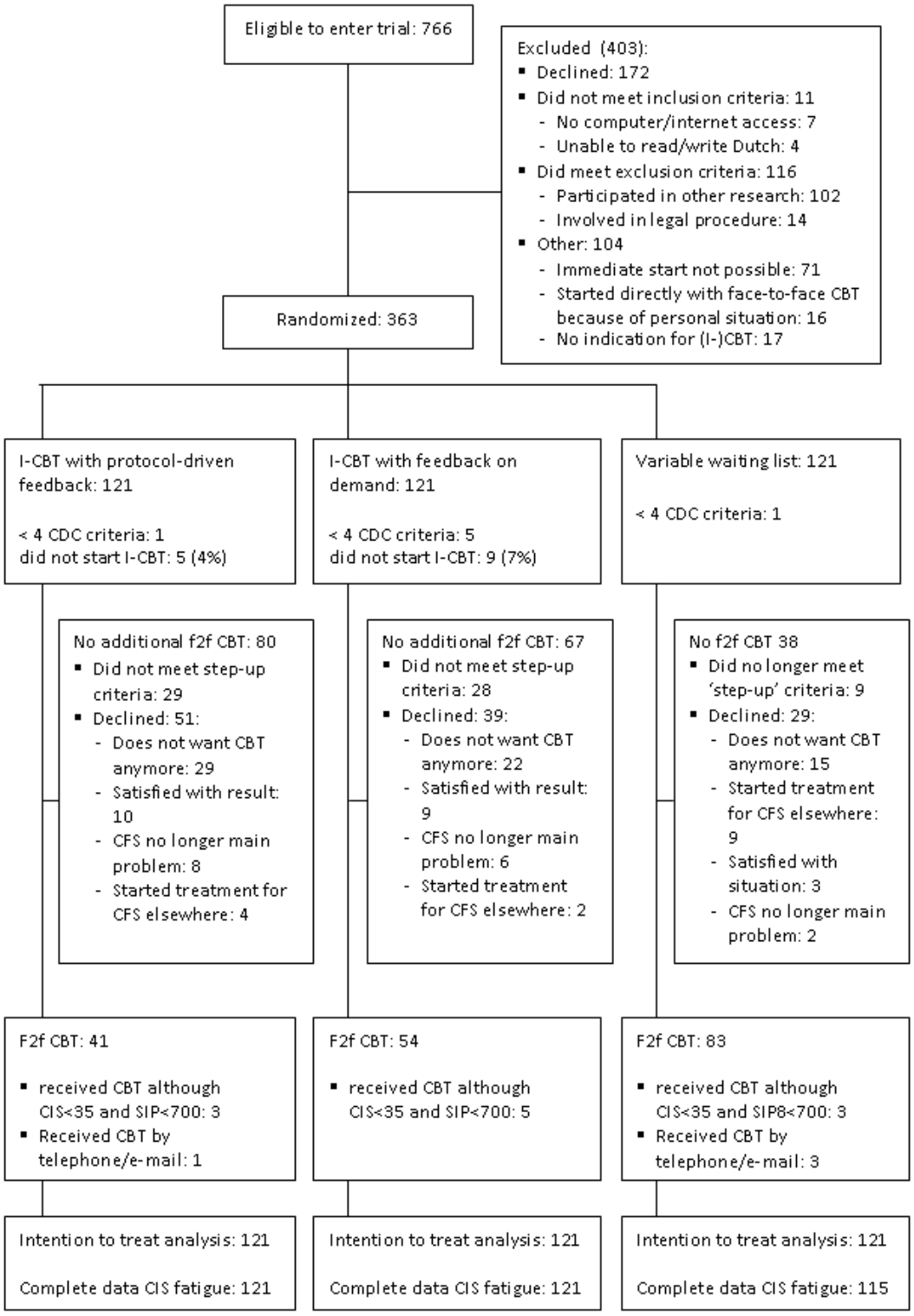
Therapist Time Invested

As the data on therapist time were not normally distributed, medians were compared with Mann-Whitney *U* tests (see Table 3). The intention-to-treat analysis showed that the median therapist time invested did not differ significantly between SC-protocol-driven feedback and TAU (*U*=7068.0; *P*=.64) and for SC-feedback-on-demand and TAU (*U*=7272.5; *P*=.93).

XSL•FO
RenderX

The analysis including only patients who started treatment showed different results. Median therapist time of both the stepped care formats differed significantly from TAU (U=6819.5; P<.001 for SC-protocol-driven feedback and U=6883.5; P<.001 for SC-feedback-on-demand).

Figure 1. CONSORT flow diagram of eligibility criteria. CBT: cognitive behavioral therapy; CDC: Centers for Disease Control and Prevention; CFS: chronic fatigue syndrome; f2f: face-to-face; I-CBT: internet-based cognitive behavioral therapy.



XSL•FO
RenderX

Table 1. Baseline characteristics.

Baseline characteristic	Stepped care		Treatment as usual
	Protocol-driven feedback	Feedback on demand	
Age in years, mean (SD)	36.6 (12.8)	37.2 (12.3)	38.7 (12.5)
Female (N=121), n (%)	78 (64.5)	69 (57.0)	74 (61.2)
Paid job (N=119), n (%)	82 (68.9)	79 (65.8)	77 (64.7)
Education level in years, mean (SD)	15.4 (1.9)	14.8 (2.3)	15.7 (1.5)
Duration of complaints in years, median (IQR[a])	4.0 (8.0)	5.0 (12.0)	6.0 (9.0)
Number of CDC[b] symptoms, median (IQR)[c]	6 (2)	6 (3)	6 (2)
Memory and/or concentration problems (N=121), n (%)	114 (94.2)	112 (92.6)	116 (95.9)
Sore throat (N=121), n (%)	53 (43.8)	57 (47.1)	56 (46.3)
Tender lymph nodes (N=121), n (%)	49 (40.5)	61 (50.4)	52 (43.0)
Muscle pain (N=121), n (%)	91 (75.2)	97 (80.2)	99 (81.8)
Multi-joint pain (N=121), n (%)	83 (68.6)	86 (71.7)	93 (76.9)
Headaches (N=121), n (%)	91 (75.2)	94 (77.7)	85 (70.2)
Unrefreshing sleep (N=121), n (%)	119 (98.3)	114 (94.2)	119 (98.3)
Postexertional malaise (N=121), n (%)	113 (93.4)	109 (90.1)	113 (93.4)
Meeting SEID[d] criteria (N=121), n (%)	89 (73.6)	92 (76.0)	93 (76.9)
Fatigue severity[e], mean (SD)	50.8 (5.0)	50.2 (4.8)	49.7 (5.3)
Overall impairment[f], mean (SD)	1488.6 (550.1)	1534.7 (562.0)	1599.2 (589.6)
Physical functioning[g], mean (SD)	62.3 (20.1)	60.5 (19.4)	61.0 (20.4)
Psychological distress[h], mean (SD)	154.4 (31.8)	160.2 (37.7)	161.2 (38.0)
Clinically relevant depressive symptoms[i] (N=120), n (%)	42 (35.0)	39 (32.5)	44 (37.0)[j]
No current psychiatric diagnosis[k] (N=121), n (%)	100 (83)	103 (85)	99 (82)
Any depressive disorder[k] (N=121), n (%)	11 (9)	9 (7)	14 (12)
Any anxiety disorder[k] (N=121), n (%)	11 (9)	11 (9)	8 (7)
Other psychiatric disorder[k] (N=121), n (%)	1 (1)	1 (1)	4 (3)
Pain[g], mean (SD)	59.3 (25.5)	59.1 (25.5)	57.7 (25.2)

[a]IQR: interquartile range.

[b]CDC: Centers for Disease Control and Prevention.

[c]Memory and/or concentration problems were scored together, so a maximum of 8 symptoms was scored.

[d]SEID: systemic exertion intolerance disease.

[e]CIS: Checklist Individual Strength.

[f]SIP8: Sickness Impact Profile.

[g]SF-36: Medical Outcomes Survey Short Form-36.

[h]SCL-90: Symptom Checklist-90.

[i]BDI-PC: Beck Depression Inventory-PC; total score ≥4.

[j]N=119.

[k]MINI: The Mini-International Neuropsychiatric Interview.

XSL•FO
RenderX

Table 2. Treatment effects.

Outcome measure	Stepped care		Treatment as usual
	Protocol-driven feedback	Feedback-on-demand	
CIS[a] fatigue severity			
Pre	50.78	50.20	49.69
Post	35.60	35.68	34.94
Cohen *d* (95% CI)[b]	1.44 (1.16-1.73)	1.50 (1.21-1.79)	1.41 (1.12-1.69)
SIP8[c]			
Pre	1488.56	1534.74	1593.20
Post	822.09	797.10	961.32
Cohen *d* (95% CI)[b]	1.09 (0.82-1.36)	1.22 (0.94-1.50)	0.91 (0.65-1.18)
SF-36[d] physical functioning			
Pre	62.27	60.54	60.95
Post	75.34	77.82	76.54
Cohen *d* (95% CI)[b]	0.58 (0.33-0.84)	0.86 (0.59-1.12)	0.72 (0.46-0.98)
SCL-90[e]			
Pre	154.36	160.20	161.22
Post	137.69	140.79	143.65
Cohen *d* (95% CI)[b]	0.42 (0.17-0.68)	0.46 (0.21-0.72)	0.40 (0.14-0.65)

[a]CIS: Checklist Individual Strength.

[b]Uncontrolled effect size: within-group effect. Cohen d=($Mean_{pre}$−$Mean_{post}$/pooled SD).

[c]SIP8: Sickness Impact Profile 8.

[d]SF-36: Medical Outcomes Survey Short Form-36.

[e]SCL-90: Symptom Checklist-90.

Table 3. Therapist time invested in total treatment in hours.

Treatment arm	Intention to treat				Starters only			
	n	Mean (hours)	Median (hours)	Minimum-maximum (hours)	n	Mean (hours)	Median (hours)	Minimum-maximum (hours)
Stepped care–protocol-driven feedback	121	09:10	08:00	2:55-22:20	118	09:19	08:09	2:55-22:20
Stepped care–feedback-on-demand	121	08:30	06:55	2:00-21:45	117	08:42	07:25	2:00-21:45
Treatment as usual	121	08:54	09:00	2:00-27:00	83	12:03	12:00	4:00-27:00

Exploratory Comparison of Both Stepped Care Formats

The difference in time invested between both the stepped care versions was significant when analyzing data of all patients (U=6237.0; P=.047), but the difference failed to reach significance when selecting only starters (U=5918.0; P=.06). More detailed information on the therapist time is provided in Table 4. On average, in SC-protocol-driven feedback, 4:04 hours (SD 2:20 hours) were spent on I-CBT. Moreover, 40 patients received f2f CBT, which took an average 6:18 hours (SD 3:37 hours) per person who received it. In SC-feedback-on-demand, 2:29 hours (SD 2:28 hours) was spent on I-CBT. In addition, 54 patients received f2f CBT, which took on average 6:30 hours (SD 4:10 hours) per patient.

Subgroup Analyses

Outcome of Face-to-Face Cognitive Behavioral Therapy After Internet-Based Cognitive Behavioral Therapy

In total, 95 patients received f2f CBT after I-CBT. As shown in Table 5, CIS fatigue was on average 5.6 points lower after f2f CBT, in comparison to after I-CBT, which was a significant change. Improvement on the SIP8 and SF-36 physical functioning was also statistically significant, whereas the SCL-90 score showed no statistically significant further improvement.

XSL•FO
RenderX

Table 4. Therapist time in stepped care.

Selected group	Stepped care–protocol-driven feedback						Stepped care–feedback-on-demand					
	n	Mean total time (hours)	Mean I-CBT[a] time (hours)	Percentage of total time	Mean f2f CBT[b] time (hours)	Percentage of total time	n	Mean total time (hours)	Mean I-CBT time (hours)	Percentage of total time (hours)	Mean f2f CBT time (hours)	Percentage of total time (hours)
Total group	121	9:10	4:04	44	2:06	23	121	8:30	2:29	29	2:54	34
No f2f CBT	81	7:00	4:06	59	0:00	0	67	5:11	2:19	45	0:00	0
Received f2f CBT	40	13:32	3:59	29	6:18	47	54	12:37	2:43	21	6:30	52

[a]I-CBT: internet-based cognitive behavioral therapy.

[b]f2f CBT: face-to-face cognitive behavioral therapy.

Table 5. Treatment effect of cognitive behavioral therapy after internet-based cognitive behavioral therapy.

Outcome measure	At T1 (after I-CBT[a]), mean (SD)	At T2 (after CBT), mean (SD)	*t* test (*df*)	*P* value
CIS[b] fatigue severity (N=95)	42.99 (9.35)	37.39 (12.06)	4.901 (94)	<.001
SIP8[c] (N=91)	1151.86 (660.76)	851.18 (673.74)	4.569 (90)	<.001
SF-36[d] physical functioning (N=91)	71.54 (21.89)	77.03 (21.36)	-2.866 (90)	.005
SCL-90[e] (N=86)	146.56 (33.29)	144.52 (46.01)	0.517 (85)	.607

[a]CBT: cognitive behavioral therapy.

[b]CIS: Checklist Individual Strength.

[c]SIP8: Sickness Impact Profile.

[d]SF-36: Medical Outcomes Survey Short Form-36.

[e]SCL-90: Symptom Checklist-90.

Differences Between First 240 and Last 123 Randomized Patients

There was a significant difference between the 2 groups regarding reduction of fatigue severity. In the first 240 patients, the change score was on average 16.4 points on the CIS fatigue severity subscale (SD 13.9). In the last 123 patients, the change score was 11.8 points (SD 11.3), t=3.374, P<.001. The differences between cohort 1 and 2 in change score for limitations (700.5 and 636.1, respectively), physical functioning (−15.7 and −14.5, respectively), and psychological distress (18.3 and 16.7, respectively) were not significant (P=.38, P=.59, and P=.74, respectively).

Discussion

Principal Findings

This study showed that I-CBT embedded in stepped care for chronic fatigue syndrome is noninferior to f2f CBT (TAU) in reducing fatigue severity. Treatment outcome of stepped care did not differ from TAU with respect to the level of disability, physical functioning, and psychological distress. The proportions of patients with clinically significant improvement of fatigue severity were equal for stepped care as well as TAU. Interestingly, this was despite the fact that approximately 50% of the patients who met the step-up criteria for f2f CBT after I-CBT did not step up.

For patients who did step up, it was found that f2f CBT after I-CBT led to a significant further improvement in fatigue severity and impairment. This suggests that stepped care with I-CBT as a first step is a viable treatment model for CFS. It was more time-efficient than usual care, as approximately 25% less therapist time was needed to deliver it. This is an important finding from a cost-effectiveness perspective, since therapist time accounts for a large proportion of the treatment costs in mental health care. Previous studies found stepped care, not including I-CBT, to be effective for CFS [20,41]. The findings of this study extend this observation and provide more insights into how stepped care can be offered in clinical practice. We found only 1 RCT that also investigated I-CBT embedded in stepped care. In patients with panic and social anxiety disorder, stepped care containing psychoeducation, I-CBT, and f2f CBT was compared with f2f CBT. Our findings are in line with this study: stepped care was noninferior to f2f CBT and was less time-intensive [42].

How therapist feedback during I-CBT influenced stepped care was explored. As in our previous study [17], we found that during I-CBT, less therapist time was needed in the feedback-on-demand than in the protocol-driven feedback format of I-CBT. However, when patients stepped up to f2f

XSL•FO
RenderX

CBT, therapists spent similar time in both conditions. Although SC-feedback-on-demand remained more time-efficient, the difference between both became smaller and failed to reach significance. This might be explained by the fact that relatively more patients in the SC-feedback-on-demand format received f2f CBT, which led to an increase in invested time. Furthermore, since we did not power for a direct comparison between both conditions, we may need more patients to draw firm conclusions on this.

We also explored if the proportion of patients willing to step up after I-CBT differed in both the feedback formats. It was found that of the patients who needed to step up after I-CBT, the proportion of patients that received f2f CBT differed in favor of the feedback-on-demand format. However, this difference was not significant. It is important to know what prevented patients from stepping up. One plausible explanation could be that patients became less motivated after unsuccessful I-CBT [43,44]. In our study, we found that approximately 60% of the patients who declined did this because they did not want further therapy, although CFS remained a problem (eg, preoccupied with other matters in life, did not want to travel, or have no faith in further recovery). Interestingly, approximately 20% were satisfied with the result of I-CBT, although they were still severely fatigued and/or disabled. It is important to know why they were satisfied despite having severe complaints. Exit interviews could be used to investigate these matters further.

Although this study shows that I-CBT fits well into a stepped care model, it is problematic that many patients do not step up when this is needed. Although this did not lead to a lower proportion of improved patients (than in TAU), and it is uncertain whether these patients would have otherwise started and completed TAU, it is possible that some of these patients would have profited more if they had received f2f CBT. To improve the integration of I-CBT in clinical care, there are some options. An aim could be to increase the number of patients that step up for f2f CBT. For example, by stopping earlier with I-CBT when it appears to be ineffective, demoralization can be prevented. However, what is a reliable indicator of the need to step up is not known. In a process study on f2f CBT for CFS, it was found that in a substantial proportion of patients, symptoms did not decrease until after 3 to 4 months [45]. The absence of a change in fatigue-perpetuating cognitions and behavior may, therefore, be a more suitable indicator early on in therapy. Further research could focus on the predicting value of these perpetuators on treatment effect in I-CBT.

It is also possible to further improve I-CBT, so that fewer patients need to step up. An option would be to improve I-CBT by developing a more flexible version, in which the intensity of therapist guidance can be further varied. The integration of videoconferencing in I-CBT makes it possible to combine the advantages of f2f sessions and I-CBT. Perhaps this also can be *on demand*, as our study indicated that patients are able to determine how much guidance they need. It is also an option to predefine specific moments in therapy, in which video consults may have an added value, because some interventions are more difficult to deliver via the internet. For example, supporting reformulation of dysfunctional beliefs is probably easier in direct interaction than by email.

Strengths and Limitations

A significant limitation of our study was that in TAU, one-third of the patients did not start the therapy. This may have reduced the treatment results in this reference group. The within-group effect size for fatigue severity in the TAU condition in our study fell outside the 95% CIs of 2 other CFS studies that had lower proportions of nonstarters [18,46]. Another possible explanation for the somewhat lower effect size in the TAU condition could be that during the study, the treatment center was moved to another university medical center. This led to organizational problems, such as uncertainty for patients about when treatment would start, changes in therapists, and substantially increased travel times for patients. There are indications that organizational features of a mental health center can influence the treatment outcome of behavioral interventions [47]. In all 3 conditions, patients who started their treatment during the last part of our study had a significantly lower reduction of fatigue.

Furthermore, a possible limitation is that the waiting list policy had changed during the study. However, retaining an unnecessarily long waiting list duration would not have been ethical. It was found that the duration of the waiting list had not influenced the treatment effect on reduction of fatigue.

Another limitation is that we did not provide data on dropout. It was difficult to produce a dropout definition that allowed for a comparison between all treatment arms. For TAU, termination of treatment before a certain number of sessions is often used as a criterion for dropout [48]. Dropout during I-CBT is difficult to measure reliably. Opening modules or logging in does not necessarily represent receiving treatment. The assessment of the therapist may also be unreliable if a patient does not request feedback or does not respond to an email; this does not have to imply that the patient dropped out of treatment as they still can follow the treatment via the internet. Unless patients actively report discontinuation of treatment, it is difficult to be certain whether someone dropped out. Furthermore, for stepped care, the number of f2f CBT sessions needed after I-CBT differs between patients, as the exposure to the content of I-CBT varies. This makes it difficult to define the minimal number of sessions that a patient has to follow before he or she is considered to have completed the treatment. More specific information on dropout, for example, by interviewing patients after they have terminated the treatment, would further aid comparison of the treatment formats.

Recently, new diagnostic criteria were proposed for CFS, by the US Institute of Medicine [33]. To aid comparison with other studies using this definition, we explored the proportion of our patients meeting SEID criteria and repeated our primary analysis with this subgroup of patients. We found the same pattern of results, suggesting that findings can probably be generalized to a significant subgroup of patients meeting the SEID criteria. It should, however, be noted that we did not assess orthostatic intolerance, which could have led to an underestimation of the proportion of patients meeting SEID in our sample.

There were important strengths of the study: it is one of the few RCTs that compared stepped care with I-CBT as a first-step treatment with usual care [19]. Our comparator was a high-intensity treatment that has proven to be effective in several

XSL•FO
RenderX

RCTs [46]. Often, the efficacy of I-CBT or stepped care is compared with a no-treatment or low-intensity control group, which seriously limits the conclusions than can be drawn from these studies [15,43,49]. Another strength is that almost all patients completed postassessment on the primary outcome measure. Finally, the study included a large sample of patients, who underwent an extensive diagnostic procedure and were treated in a specialized tertiary treatment center by experienced, well-trained, and supervised therapists.

Conclusion

In conclusion, this RCT showed that stepped care with I-CBT as the first and f2f CBT as the second step for chronic fatigue syndrome is noninferior to TAU in the reduction of fatigue severity. A substantial part of the patients did not need to step up after stepped care, which made treatment for them less intensive. For patients who needed to step up, f2f CBT led to an additional treatment effect. Stepped care took less therapist time to deliver. However, a substantial proportion of patients did not step up after I-CBT with limited results. Treatment efficacy can probably be improved when relatively more patients step up and by further developing the I-CBT.

Acknowledgments

The authors thank all the patients for their participation and all therapists involved for delivering CBT. They also thank T Fasotti-Dumont, S van Hees, J Smeets, and L Vermeeren for their help with data collection and administration as well as R Donders and P Nieuwkerk for statistical advice and H Abrahams for performing the statistical analyses. The authors also thank J Rosmalen for her comments on a draft of the manuscript.

Authors' Contributions

HK, AJ, and MW-S designed the study. MW-S and AJ were involved with data acquisition. MW-S conducted the analysis. Interpretation of data was performed by HK and MW-S. All authors were involved in the interpretation of results as well as reviewing and revising the manuscript.

Conflicts of Interest

HK receives royalties for the treatment protocol of f2f-CBT, which was published.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 7MB - jmir_v21i3e11276_app1.pdf]

References

1. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. Ann Intern Med 1994 Dec 15;121(12):953-959. [Medline: 7978722]
2. Reeves WC, Lloyd A, Vernon SD, Klimas N, Jason LA, Bleijenberg G, International Chronic Fatigue Syndrome Study Group. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. BMC Health Serv Res 2003 Dec 31;3(1):25 [FREE Full text] [doi: 10.1186/1472-6963-3-25] [Medline: 14702202]
3. Johnston S, Brenu EW, Staines D, Marshall-Gradisnik S. The prevalence of chronic fatigue syndrome/ myalgic encephalomyelitis: a meta-analysis. Clin Epidemiol 2013;5:105-110 [FREE Full text] [doi: 10.2147/CLEP.S39876] [Medline: 23576883]
4. Cairns R, Hotopf M. A systematic review describing the prognosis of chronic fatigue syndrome. Occup Med (Lond) 2005 Jan;55(1):20-31. [doi: 10.1093/occmed/kqi013] [Medline: 15699087]
5. Prins JB, van der Meer JW, Bleijenberg G. Chronic fatigue syndrome. Lancet 2006 Jan 28;367(9507):346-355. [doi: 10.1016/S0140-6736(06)68073-2] [Medline: 16443043]
6. Vercoulen JH, Swanink CM, Galama JM, Fennis JF, Jongen PJ, Hommes OR, et al. The persistence of fatigue in chronic fatigue syndrome and multiple sclerosis: development of a model. J Psychosom Res 1998 Dec;45(6):507-517. [Medline: 9859853]
7. Keijmel SP, Saxe J, van der Meer JW, Nikolaus S, Netea MG, Bleijenberg G, et al. A comparison of patients with Q fever fatigue syndrome and patients with chronic fatigue syndrome with a focus on inflammatory markers and possible fatigue perpetuating cognitions and behaviour. J Psychosom Res 2015 Oct;79(4):295-302. [doi: 10.1016/j.jpsychores.2015.07.005] [Medline: 26272528]

XSL•FO
RenderX

8. Malouff JM, Thorsteinsson EB, Rooke SE, Bhullar N, Schutte NS. Efficacy of cognitive behavioral therapy for chronic fatigue syndrome: a meta-analysis. Clin Psychol Rev 2008 Jun;28(5):736-745. [doi: 10.1016/j.cpr.2007.10.004] [Medline: 18060672]
9. Price JR, Mitchell E, Tidy E, Hunot V. Cognitive behaviour therapy for chronic fatigue syndrome in adults. Cochrane Database Syst Rev 2008 Jul 16(3):CD001027. [doi: 10.1002/14651858.CD001027.pub2] [Medline: 18646067]
10. Castell BD, Kazantzis N, Moss-Morris RE. Cognitive behavioral therapy and graded exercise for chronic fatigue syndrome: a meta-analysis. Clin Psychol Sci Pract 2011;18(4):311-324. [doi: 10.1111/j.1468-2850.2011.01262.x]
11. Gezondheidsraad. Het chronische-vermoeidheidssyndroom. Ned Tijdschr Geneeskd 2005;Den Haag(149):739-741.
12. 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]
13. 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]
14. Gratzer D, Khalid-Khan F. Internet-delivered cognitive behavioural therapy in the treatment of psychiatric illness. CMAJ 2016 Mar 01;188(4):263-272 [FREE Full text] [doi: 10.1503/cmaj.150007] [Medline: 26527829]
15. 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]
16. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
17. Janse A, Worm-Smeitink M, Bleijenberg G, Donders R, Knoop H. Efficacy of web-based cognitive-behavioural therapy for chronic fatigue syndrome: randomised controlled trial. Br J Psychiatry 2018 Feb;212(2):112-118. [doi: 10.1192/bjp.2017.22] [Medline: 29436329]
18. Wiborg JF, van Bussel J, van Dijk A, Bleijenberg G, Knoop H. Randomised controlled trial of cognitive behaviour therapy delivered in groups of patients with chronic fatigue syndrome. Psychother Psychosom 2015;84(6):368-376. [doi: 10.1159/000438867] [Medline: 26402868]
19. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Dec 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
20. Tummers M, Knoop H, Bleijenberg G. Effectiveness of stepped care for chronic fatigue syndrome: a randomized noninferiority trial. J Consult Clin Psychol 2010 Oct;78(5):724-731. [doi: 10.1037/a0020052] [Medline: 20873907]
21. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-9; quiz 870. [doi: 10.1586/ern.12.63] [Medline: 22853793]
22. 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. [doi: 10.1080/16506070903318960] [Medline: 20183695]
23. 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]
24. 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 Jun 08;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]
25. CBO. 2013. Richtlijn Diagnose, Behandeling, Begeleiding en Beoordeling van Patiënten met het Chronisch Vermoeidheidssyndroom (CVS) URL: http://nkcv.nl/app/uploads/2015/10/richtlijn-chronisch-vermoeidheidssyndroom-cvs.pdf [accessed 2018-06-08] [WebCite Cache ID 701FmE2wP]
26. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59(Suppl 20):22-33;quiz 34-57 [FREE Full text] [Medline: 9881538]
27. Worm-Smeitink M, Gielissen M, Bloot L, van Laarhoven HW, van Engelen BG, van Riel P, et al. The assessment of fatigue: psychometric qualities and norms for the checklist individual strength. J Psychosom Res 2017 Jul;98:40-46. [doi: 10.1016/j.jpsychores.2017.05.007] [Medline: 28554371]
28. Jacobs HM, Luttik A, Touw-Otten FW, de Melker RA. [The sickness impact profile; results of an evaluation study of the Dutch version]. Ned Tijdschr Geneeskd 1990 Oct 06;134(40):1950-1954. [Medline: 2234151]
29. Janse A, Worm-Smeitink M, Bussel-Lagarde J, Bleijenberg G, Nikolaus S, Knoop H. Testing the efficacy of web-based cognitive behavioural therapy for adult patients with chronic fatigue syndrome (CBIT): study protocol for a randomized controlled trial. BMC Neurol 2015 Aug 12;15:137 [FREE Full text] [doi: 10.1186/s12883-015-0392-3] [Medline: 26264735]
30. Knoop H, Bleijenberg G. Het Chronisch Vermoeidheidssyndroom: Behandelprotocol cognitieve gedragstherapie voor CVS. Houten: Bohn Stafleu van Loghum; Jun 12, 2010.

XSL•FO
RenderX

31. Stulemeijer M, de Jong LW, Fiselier TJ, Hoogveld SW, Bleijenberg G. Cognitive behaviour therapy for adolescents with chronic fatigue syndrome: randomised controlled trial. Br Med J 2005 Jan 01;330(7481):14 [FREE Full text] [doi: 10.1136/bmj.38301.587106.63] [Medline: 15585538]
32. Beck AT, Guth D, Steer RA, Ball R. Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. Behav Res Ther 1997 Aug;35(8):785-791. [Medline: 9256522]
33. Institute of Medicine of the National Academies. Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining An Illness. Washington, DC: The National Academies Press; 2015.
34. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res 1994 Jul;38(5):383-392. [Medline: 7965927]
35. Knoop H, Bleijenberg G, Gielissen MF, van der Meer JW, White PD. Is a full recovery possible after cognitive behavioural therapy for chronic fatigue syndrome? Psychother Psychosom 2007;76(3):171-176. [doi: 10.1159/000099844] [Medline: 17426416]
36. van der Zee KI. In: Sanderman R, editor. Het meten van de algemene gezondheidstoestand met de RAND-36: een handleiding. Rijksuniversiteit Groningen: Research Institute SHARE UMCG; 2012.
37. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
38. Arrindell WA. In: Ettema JH, editor. SCL-90 Symptom Checklist. Lisse: Swets Test Publishers; 2003.
39. Knoop H, van der Meer JW, Bleijenberg G. Guided self-instructions for people with chronic fatigue syndrome: randomised controlled trial. Br J Psychiatry 2008 Oct;193(4):340-341. [doi: 10.1192/bjp.bp.108.051292] [Medline: 18827302]
40. 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. [Medline: 2002127]
41. Ali S, Goldsmith K, Burgess M, Chalder T. Guided self-help for patients with chronic fatigue syndrome prior to starting cognitive behavioural therapy: a cohort study. Behav Cogn Psychother 2017 Sep;45(5):448-466. [doi: 10.1017/S135246581700025X] [Medline: 28473005]
42. Nordgreen T, Haug T, Öst LG, Andersson G, Carlbring P, Kvale G, et al. Stepped care versus direct face-to-face cognitive behavior therapy for social anxiety disorder and panic disorder: a randomized effectiveness trial. Behav Ther 2016 Mar;47(2):166-183. [doi: 10.1016/j.beth.2015.10.004] [Medline: 26956650]
43. van Straten A, Hill J, Richards DA, Cuijpers P. Stepped care treatment delivery for depression: a systematic review and meta-analysis. Psychol Med 2015 Jan;45(2):231-246. [doi: 10.1017/S0033291714000701] [Medline: 25065653]
44. Jaehne A, Loessl B, Frick K, Berner M, Hulse G, Balmford J. The efficacy of stepped care models involving psychosocial treatment of alcohol use disorders and nicotine dependence: a systematic review of the literature. Curr Drug Abuse Rev 2012 Mar;5(1):41-51. [Medline: 22280331]
45. Heins MJ, Knoop H, Burk WJ, Bleijenberg G. The process of cognitive behaviour therapy for chronic fatigue syndrome: which changes in perpetuating cognitions and behaviour are related to a reduction in fatigue? J Psychosom Res 2013 Sep;75(3):235-241 [FREE Full text] [doi: 10.1016/j.jpsychores.2013.06.034] [Medline: 23972412]
46. Worm-Smeitink M, Nikolaus S, Goldsmith K, Wiborg J, Ali S, Knoop H, et al. Cognitive behaviour therapy for chronic fatigue syndrome: differences in treatment outcome between a tertiary treatment centre in the United Kingdom and the Netherlands. J Psychosom Res 2016 Aug;87:43-49. [doi: 10.1016/j.jpsychores.2016.06.006] [Medline: 27411751]
47. Clark DM, Canvin L, Green J, Layard R, Pilling S, Janecka M. Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. Lancet 2018 Feb 17;391(10121):679-686 [FREE Full text] [doi: 10.1016/S0140-6736(17)32133-5] [Medline: 29224931]
48. Sharf J, Primavera LH, Diener MJ. Dropout and therapeutic alliance: a meta-analysis of adult individual psychotherapy. Psychotherapy (Chic) 2010 Dec;47(4):637-645. [doi: 10.1037/a0021175] [Medline: 21198249]
49. Ho FY, Yeung WF, Ng TH, Chan CS. The efficacy and cost-effectiveness of stepped care prevention and treatment for depressive and/or anxiety disorders: a systematic review and meta-analysis. Sci Rep 2016 Jul 05;6:29281 [FREE Full text] [doi: 10.1038/srep29281] [Medline: 27377429]

Abbreviations

ANCOVA: analysis of covariance
CBT: cognitive behavioral therapy
CDC: Centers for Disease Control and Prevention
CFS: chronic fatigue syndrome
CIS: Checklist Individual Strength
f2f: face-to-face
I-CBT: internet-based cognitive behavioral therapy
ME: myalgic encephalomyelitis
RCI: reliable change index
RCTs: randomized controlled trials

XSL•FO
RenderX

SC: stepped care
SCL-90: Symptom Checklist-90
SEID: systemic exertion intolerance disease
SF-36: Medical Outcomes Survey Short Form-36
SIP8: Sickness Impact Profile
TAU: treatment as usual

Edited by G Eysenbach; submitted 12.06.18; peer-reviewed by D Gratzer, F Khalid-Khan; comments to author 05.09.18; revised version received 17.09.18; accepted 19.09.18; published 14.03.19.

Please cite as:
Worm-Smeitink M, Janse A, van Dam A, Evers A, van der Vaart R, Wensing M, Knoop H
Internet-Based Cognitive Behavioral Therapy in Stepped Care for Chronic Fatigue Syndrome: Randomized Noninferiority Trial
J Med Internet Res 2019;21(3):e11276
URL: http://www.jmir.org/2019/3/e11276/
doi:10.2196/11276
PMID:30869642

©Margreet Worm-Smeitink, Anthonie Janse, Arno van Dam, Andrea Evers, Rosalie van der Vaart, Michel Wensing, Hans Knoop. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User-Centered Design of a Web-Based Tool to Support Management of Chemotherapy-Related Toxicities in Cancer Patients

Rebecca M Prince[1,2], MBBS, FRACP, MSc; Anthony Soung Yee[1], PhD; Laura Parente[1], BDes; Katherine A Enright[3], MD, MPH; Eva Grunfeld[2,4], MD, DPhil, FCFP; Melanie Powis[1], MSc; Amna Husain[2], MD, CCFP, MP; Sonal Gandhi[5], MD, MSc, FRCPC; Monika K Krzyzanowska[1,2], MD, MPH, FRCPC

[1]University Health Network, Toronto, ON, Canada

[2]University of Toronto, Toronto, ON, Canada

[3]Trillium Health Partners, Mississauga, ON, Canada

[4]Ontario Institute for Cancer Research, Toronto, ON, Canada

[5]Sunnybrook Odette Cancer Centre, Toronto, ON, Canada

Corresponding Author:
Monika K Krzyzanowska, MD, MPH, FRCPC
University Health Network
610 University Ave.
Toronto, ON,
Canada
Phone: 1 416 946 6542
Email: monika.krzyzanowska@uhn.ca

Abstract

Background: Cancer patients receiving chemotherapy have high symptom needs that can negatively impact quality of life and result in high rates of unplanned acute care visits. Remote monitoring tools may improve symptom management in this patient population.

Objective: This study aimed to design a prototype tool to facilitate remote management of chemotherapy-related toxicities.

Methods: User needs were assessed using a participatory, user-centered design methodology that included field observation, interviews, and focus groups, and then analyzed using affinity diagramming. Participants included oncology patients, caregivers, and health care providers (HCPs) including medical oncologists, oncology nurses, primary care physicians, and pharmacists in Ontario, Canada. Overarching themes informed development of a Web-based prototype, which was further refined over 2 rounds of usability testing with end users.

Results: Overarching themes were derived from needs assessments, which included 14 patients, 1 caregiver, and 12 HCPs. Themes common to both patients and HCPs included gaps and barriers in current systems, need for decision aids, improved communication and options in care delivery, secure access to credible and timely information, and integration into existing systems. In addition, patients identified missed opportunities, care not meeting their needs, feeling overwhelmed and anxious, and wanting to be more empowered. HCPs identified accountability for patient management as an issue. These themes informed development of a Web-based prototype (*bridges*), which included toxicity tracking, self-management advice, and HCP communication functionalities. Usability testing with 11 patients and 11 HCPs was generally positive; however, identified challenges included tool integration into existing workflows, need for standardized toxicity self-management advice, issues of privacy and consent, and patient-tailored information.

Conclusions: Web-based tools integrating just-in-time self-management advice and HCP support into routine care may address gaps in systems for managing chemotherapy-related toxicities. Attention to the integration of new electronic tools into self-care by patients and practice was a strong theme for both patients and HCP participants and is a key issue that needs to be addressed for wide-scale adoption.

(J Med Internet Res 2019;21(3):e9958) doi:10.2196/jmir.9958

XSL•FO
RenderX

KEYWORDS

prototype; Web-based tool; toxicity management; chemotherapy; self-management

Introduction

Cancer patients receiving chemotherapy have high symptom burden, which is reflected in their frequent utilization of the emergency department (ED) and high rates of hospitalization during treatment [1-4]. Chemotherapy-related toxicities usually occur between ambulatory visits to the oncology clinic. Unplanned acute care visits among patients receiving systemic therapy in routine practice are likely a reflection of suboptimal management of these toxicities due to inadequate use of preventative strategies or lack of timely access to advice and assessment by the health care team. Some ED visits and hospitalizations may be potentially avoidable with proactive monitoring between clinic visits [5]. The widespread diffusion of health information technology (HIT) represents an opportunity to address gaps in current health care systems with access to Web-enabled devices becoming increasingly common.

Technological solutions have been found to be acceptable to a wide range of populations including older individuals and those with little experience using Web-based technologies [5-8]. Web-based solutions allow an immediacy of access to information and feedback that paper-based systems are not able to provide [7,9] and with a greater degree of accuracy [10]. The design of an effective Web-based chemotherapy toxicity management tool requires an understanding of factors associated with the interaction between humans, technology, and care context to ensure uptake by end users and integration into existing clinical workflows [11]. Engaging end users from the outset also increases external validity and results in highly accurate and relevant solutions while avoiding features and functionalities that were not relevant or useful in existing tools.

To design a prototype Web-based tool to facilitate remote management of chemotherapy-related toxicities, we used an iterative, participatory design methodology informed by human factors principles. As the majority of side effects and subsequent ED visits and hospitalizations occur between clinic visits, we focused on the needs of patients receiving chemotherapy in the outpatient setting. We initiated this study at a time when the evidence base regarding the development and effectiveness of such tools was limited; hence, we felt a local solution was needed [12-14]. A recently published single institution randomized trial of electronic symptom tracking between clinic visits in patients with advanced cancer receiving chemotherapy has shown improvement in patient outcomes including fewer ED visits and hospitalizations [15,16]. However, HIT is an area that is evolving rapidly and, as such, tools which began development over a decade ago [12] may no longer be as relevant. Furthermore, existing tools have mostly undergone academic development in a limited number of settings and have not yet moved into routine practice in the cancer context so there were no *off the shelf tools* that could be easily adopted. In addition, none of the existing tools have been validated in Canada, so it is unclear whether they would be relevant to the Canadian context of a universal health care system that includes provincially organized but locally delivered cancer care. Not all publicly funded health systems have a similar organization, availability of resources, and result in patient populations with different needs. Successful implementation of an electronic tool requires deep understanding of the required features and functionalities and of the health system into which it will be deployed but can also provide insights applicable beyond local context, thus adding to the growing body of knowledge.

Methods

Study Overview

To design a prototype of an electronic tool to address gaps in chemotherapy-related symptom management, we used a user-centered participatory design methodology [17]. A Web-based tool was preferred to allow use on any internet-enabled device to mitigate logistical and cost concerns. Potential end users, including patients, their caregivers, and health care providers (HCPs; medical oncologists, primary care physicians, and oncology nurses and pharmacists) were involved in all steps of prototype conceptualization and development. The study was approved by the University Health Network and Trillium Health Partners ethics boards. Informed consent was obtained from all participants; all participants received an honorarium for their time.

Recruitment and Questionnaire Administration

A convenience sampling approach was utilized whereby HCPs were invited to participate directly by a study team member. Patients and their caregivers were invited to participate by their treating medical oncologist or through an email sent out by the Cancer Care Ontario Patient and Family Advisory Council. Patient participants were required to have received chemotherapy for any cancer type with any intent within the previous 2 years. To ensure a wide sampling of views, each participant undertook only one study activity (ethnographic field study, focus group, or prototype testing). All participants completed a baseline questionnaire to assess their level of interaction with information technology. The patient questionnaire consisted of 11 demographic questions and 10 information technology questions; the HCP questionnaire consisted of 11 demographic questions and 11 information technology questions. For questions estimating hours of computer and internet usage per day, the lower range number was used for analysis.

User Needs Assessment

To understand and gather insights into the context in which care is provided, we utilized the event-focused ethnographic field study methodology of Bloomberg et al [17] consisting of field observation and interviews and focus groups with potential end users. To understand the tasks, workflows, information requirements and usage patterns, communication and decision-making mechanisms, and use of supporting systems involved in toxicity management, field observations were undertaken. Field observation consisted of silent first-hand observation [18] of interactions between HCPs and patients

XSL•FO
RenderX

during routine clinic visits in the ambulatory clinical setting with contemporaneous notes taken by 2 2 human factors specialists (ASY and LP). Following each field observation session, separate semistructured interviews were held with providers and patients to gather further information on gaps and facilitators. Interviews were guided by a script but left room to clarify any observed issues to gain deeper insights. Contemporaneous notes were taken, and interviews were audio recorded. Saturation of patient insights, defined as exhaustion of new feedback and themes [19], was not reached following the first round of ethnography; so additional patients were recruited and studied until saturation was reached, as confirmed by sampling of 2 additional patients. A total of 5 study team members confirmed saturation by consensus.

To encourage open discussion of views on issues and to minimize the risk of power imbalance, separate HCP and patient and caregiver focus groups were held and moderated by 2 human factors specialists (ASY and LP). Each participant attended a 1- to 2-hour focus group. Focus groups utilized open-ended questions and probes to generate discussion about gaps in the current health system related to symptom management as well as the content and functionality of a toxicity management tool. Focus groups were audio recorded and contemporaneous notes were taken.

Analysis

Findings from demographics and preferences questionnaires were summarized using descriptive statistics. Qualitative data gathered during the ethnographic field studies and focus groups were thematically analyzed using the affinity diagramming method of Holtzblatt and Beyer [20] during ideation sessions.

Preparation for the affinity diagramming method consisted of a coding exercise on all the gathered data. Moreover, each of the 2 human factors specialists (ASY and LP) reviewed all data from all sessions generating codes denoting relevant keywords, phrases, and quotes from participant data. Each source of data (field observation and interviews and focus groups) was analyzed separately, which afforded methodological triangulation [21].

These codes were used in the ideation sessions, which were attended by 2 human factors study members (ASY and LP) and 3 HCP study members (RP, MK, and MP). On the basis of the principles of a Team Interpretation session [20], the 5 study members discussed each piece of coded data, sharing insights from the clinical and human factors perspectives. Through discussion of coded data, natural groupings began to form, which developed into themes that encapsulated each natural grouping. For example, one quote from an HCP, “Patients are overwhelmed. My guess is that 95% of the information is instantly forgotten.” was coded as *Overwhelmed* and grouped under the patient-specific theme of *Anxiety and feeling overwhelmed*.

Consensus on the themes was reached by all 5 members, and theme saturation was achieved when no new themes emerged for any presented piece of coded data [19]. This method allowed for effective identification of interrelated factors and existing issues related to remote management of chemotherapy-related toxicities; these themes informed the design of the prototype Web-based toxicity management tool.

Design and Usability Testing

On the basis of the different functional and information requirements gathered during the needs assessment, distinct interactive prototype interfaces were designed for both the patient and HCP roles. Usability testing was facilitated by human factors specialists and comprised hour-long cognitive walk-throughs, where participants performed realistic tasks while *thinking aloud* [22]. Depending on participant preference, usability testing was conducted either in a private room in the clinical area or at the human factors usability laboratories. The human factors specialists (ASY and LP) explained the purpose and objectives of the usability testing to participants and presented the low fidelity prototype (screen shots without any of the interactive functionality) with scenarios for the participant to complete using the prototype. Sessions were observed, videotaped, and audio taped, and contemporaneous notes were taken. Participants were debriefed at the conclusion of testing to ascertain if any further issues arose during testing.

Target recruitment for usability testing was 4 to 5 participants in the patient and caregiver and clinician groups, respectively, to identify up to 80% of usability problems as recommended in the literature [11]. The tool, comprising both a patient-facing and clinician-facing interface, was refined over 2 rounds of usability testing, resulting in a higher fidelity interactive prototype.

The data from the patient and caregiver and clinician groups were analyzed separately. The data from the video recordings from usability testing also underwent thematic analysis using affinity diagramming methodology [20] with 2 human factors specialists (ASY and LP) reaching consensus on the issues on both interfaces discovered through usability testing. The 2 human factors specialists jointly rank ordered the issues using a weighted decision matrix [23], assigning weights and ratings to various criteria including *Ease of design change*, *Severity of usability issue*, *Benefit to patient*, and *Benefit to HCP*, for each of the identified issues. The issues were iterated upon in the prototype starting from the highest overall rank until all issues were addressed.

Results

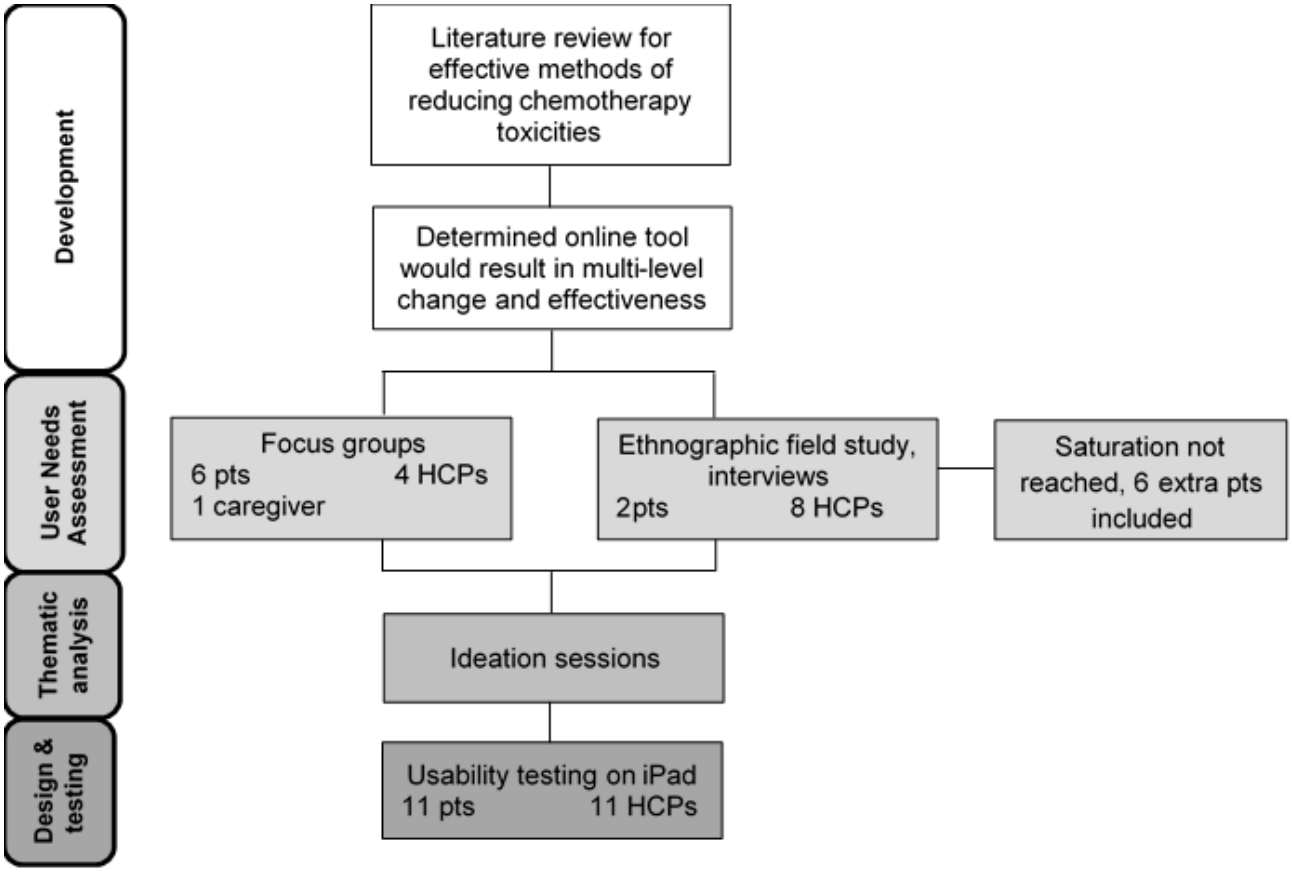
Participant Demographics

Between December 2014 and November 2015, 49 patients, caregivers, and HCPs participated in the study. Overall, 8 patients and 8 HCPs participated in ethnographic field studies; 6 patients, 1 caregiver, and 4 HCPs participated in focus groups; and 11 patients and 11 HCPs participated in usability testing (Figure 1). Baseline characteristics of study participants are shown in Table 1; as only 1 caregiver participated, findings were combined with patient responses. Patient and caregiver participants were primarily female (58% [15/26]) with a median age of 55 years, had a diagnosis or cared for someone with a diagnosis of breast cancer (31% [8/26]), had at least a college education (81% [21/26]), and spoke English as their first language (92% [24/26]). HCPs were primarily female (80%

XSL•FO
RenderX

[16/20]) with a median age of 50 years, had been in the medical profession on average for 20 years, spoke English as their first language (85% [17/20]), and had a hospital-based practice (90% [18/20]).

Figure 1. Flow diagram of study design. HCP: health care providers; pts: points.



XSL•FO
RenderX

Table 1. Summary of participant demographics and clinical characteristics.

Characteristics	All (n=46)	Patients and caregivers (n=26)	Health care providers (n=20[a])
Gender, n (%)			
Male	13 (28)	10 (38)	3 (15)
Female	31 (67)	15 (58)	16 (80)
Missing	2 (4)	1 (4)	1 (5)
Age (years)			
Median (range)	—[b]	55 (29-75)	50 (31-65)
Missing	—	1	1
Profession, n (%)			
Family physician	—	—	2 (10)
Medical oncologist	—	—	5 (25)
Oncology nurse	—	—	12 (60)
Pharmacist	—	—	1 (5)
Years in health care			
Median (range)	—	—	20 (1.5-41)
Missing	—	—	1
Diagnosis, n (%)			
Gastrointestinal cancer	—	5 (19)	—
Breast cancer	—	8 (31)	—
Lung cancer	—	2 (8)	—
Lymphoma	—	6 (23)	—
Other	—	2 (8)	—
Missing	—	2 (8)	—
Treatment, n (%)			
Chemotherapy	—	25 (96)	—
Radiation	—	13 (50)	—
Surgery	—	13 (50)	—
Education, n (%)			
Professional/ graduate degree	—	8 (31)	—
College/university	—	13 (50)	—
High school	—	3 (12)	—
Primary/middle school	—	2 (8)	—
Income, n (%)			
CAN $30-59k	—	3 (12)	—
CAN $60-89k	—	5 (19)	—
>CAN $90k	—	12 (46)	—
Prefer not to say	—	6 (23)	—
First language, n (%)			
English	41 (89)	24 (92)	17 (85)
Other	5 (11)	2 (8)	3 (5)
Practice setting, n (%)			
Community-based clinic	—	—	2 (10)

XSL•FO
RenderX

Characteristics	All (n=46)	Patients and caregivers (n=26)	Health care providers (n=20[a])
Hospital	—	—	7 (35)
Hospital clinic	—	—	11 (55)

[a]Data missing for 3 health care providers.

[b]Not applicable.

Baseline Technology Use

Computer and internet use were very common among HCPs, patients, and caregivers (Table 2) with the majority of participants using computers at work (80% [37/46]) and at home (93% [43/46]) and having access to internet at home (98% [45/46]). HCPs reported using electronic devices approximately twice as much as patients (6.5 vs 3.5 hours per day). Most participants reported being *comfortable* or *very comfortable* using computers, smartphones and tablets, internet, email, and instant messaging. Only 50% (23/46) of patients and HCPs felt *comfortable* using social media with 50% (13/26) of patients and 15% (3/20) of HCPs either *not using* or feeling *not at all comfortable* with social media platforms.

User Needs

Thematic analysis of data collected from the ethnographic field studies and focus groups revealed themes of the recognition of gaps in the current health system, the existence of barriers to accessing care, and the need for timely support in decision making, which were common to both patients and HCPs (Figure 2). Additional themes were specific to either patients and caregivers or to HCPs. Patients and caregivers recognized that health systems often do not match their needs and that many psychological symptoms are not well addressed, especially feelings of being overwhelmed and anxious. HCPs identified the need for clear lines of accountability for any decisions and advice given to patients through a Web-based tool.

Design of Interactive Prototype

The design of an interactive prototype with separate patient and HCP interfaces, *bridges*, was informed by the findings of the user needs assessment. The patient interface includes functionality for toxicity reporting, self-management advice based on the reported toxicities, an appointment calendar, educational materials, and options for HCP communication (Figure 3). The HCP interface includes functionality for viewing patient visit history and voicemails left by the patient, trending of toxicity reporting data, and a standardized treatment guide for handling common chemotherapy-related toxicities.

Usability Testing

Iterative rounds of usability testing were undertaken with 22 participants (11 patients and 11 HCPs) to further refine the prototype and evaluate end-user needs. During usability testing, patients noted that it is often left to them to self-organize their care including integration of complex scheduling information and medication administration (Textbox 1). Many patients desired access to an up-to-date Web-based calendar from the hospital or a calendar they could synchronize with their own device. Patients were very clear that they wanted access to information specific to their cancer type and treatment regimen rather than general information which may not apply to them. They had concerns about whether they would actually use the tool if they were feeling unwell and highlighted that the ability for a caregiver to fill it out for them would be useful. Patients also provided insight into the fact that care teams are frequently comprised of many members that change regularly.

Issues noted by HCPs included concerns of how a Web-based tool would be integrated into existing workflows and procedures, especially existing electronic medical records, to avoid entering the same information into multiple HIT systems. The context within which patient-reported toxicities occurred was felt to be very important, so the ability to add notes within the app was desired. Issues of confidentiality and privacy were highlighted as critical and requiring clarification, particularly with regard to patient consent around communication between HCPs and caregivers. The ability to prioritize patient symptoms and other issues in order of urgency was also felt to be important to ensure that the most serious issues were addressed first. The use of a treatment guide for symptom management was seen as a useful tool to ensure HCPs give consistent advice to patients.

XSL•FO
RenderX

Table 2. Baseline computer and information technology use.

Category and Response	All (n=46)	Patients and caregivers (n=26)	Health care providers (n=20[a])
Computer use at work, n (%)			
Yes	37 (80)	17 (65)	20 (100)
No	7 (15)	7 (27)	0 (0)
Missing	2 (4)	2 (8)	0 (0)
Computer use at home, n (%)			
Yes	43 (93)	24 (92)	19 (95)
No	3 (7)	2 (8)	1 (5)
Internet access at home, n (%)			
Yes	45 (98)	25 (96)	20 (100)
No	0 (0)	0 (0)	0 (0)
Missing	1 (2)	1 (4)	0 (0)
Hours on computer, smartphone, and tablet, per day, median (range)	—[b]	3.5 (<1-9)	6.5 (1-12)
Comfort using a computer, n (%)			
Not at all	1 (2)	1 (4)	0 (0)
A little	2 (4)	1 (4)	1 (5)
Comfortable	15 (33)	6 (23)	9 (45)
Very comfortable	27 (59)	17 (65)	10 (50)
Missing	1 (2)	1 (4)	0 (0)
Comfort using a smartphone or tablet, n (%)			
Do not use	1 (2)	1 (4)	0 (0)
Not at all	3 (7)	2 (8)	1 (5)
A little	4 (9)	1 (4)	3 (15)
Comfortable	15 (33)	6 (23)	9 (45)
Very comfortable	23 (50)	16 (62)	7 (35)
Hours on internet per day, median (range)	—	2 (<1-12)	3 (1-12)
Comfort using internet, n (%)			
Do not use	1 (2)	1 (4)	0 (0)
Not at all	0 (0)	0 (0)	0 (0)
A little	3 (7)	3 (12)	0 (0)
Comfortable	15 (33)	5 (19)	10 (50)
Very comfortable	27 (59)	17 (65)	10 (50)
Comfort using email, n (%)			
Do not use	0 (0)	0 (0)	0 (0)
Not at all	2 (4)	2 (8)	0 (0)
A little	1 (2)	1 (4)	0 (0)
Comfortable	14 (30)	5 (19)	9 (45)
Very comfortable	29 (63)	18 (69)	11 (55)
Comfort using instant messaging, n (%)			
Do not use	3 (7)	3 (12)	0 (0)
Not at all	1 (2)	1 (4)	0 (0)

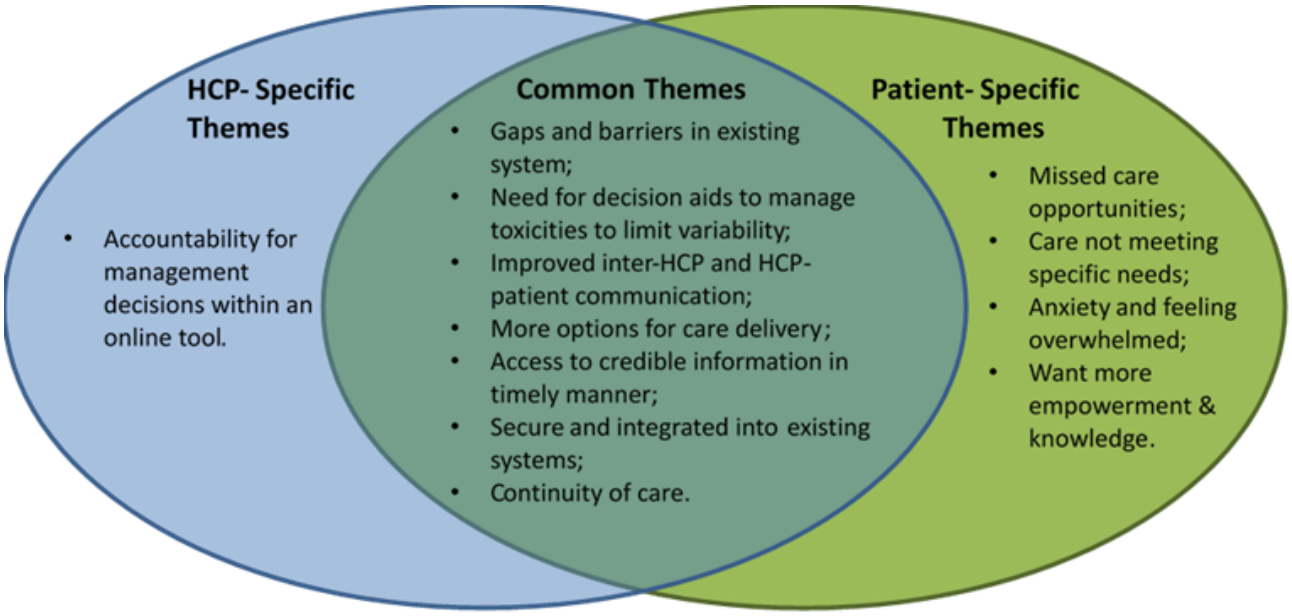
XSL•FO
RenderX

Category and Response	All (n=46)	Patients and caregivers (n=26)	Health care providers (n=20[a])
A little	4 (9)	1 (4)	3 (15)
Comfortable	12 (26)	4 (15)	8 (40)
Very comfortable	26 (57)	17 (65)	9 (45)
Comfort using social media, n (%)			
Do not use	11 (24)	9 (35)	2 (10)
Not at all	5 (11)	4 (15)	1 (5)
A little	7 (15)	0 (0)	7 (35)
Comfortable	12 (26)	7 (27)	5 (25)
Very comfortable	11 (24)	6 (23)	5 (25)
Technology used in normal workday, n (%)			
Computer	—	—	20 (100)
Smartphone	—	—	11 (55)
Tablet	—	—	3 (15)

[a]Data missing for 3 health care providers.

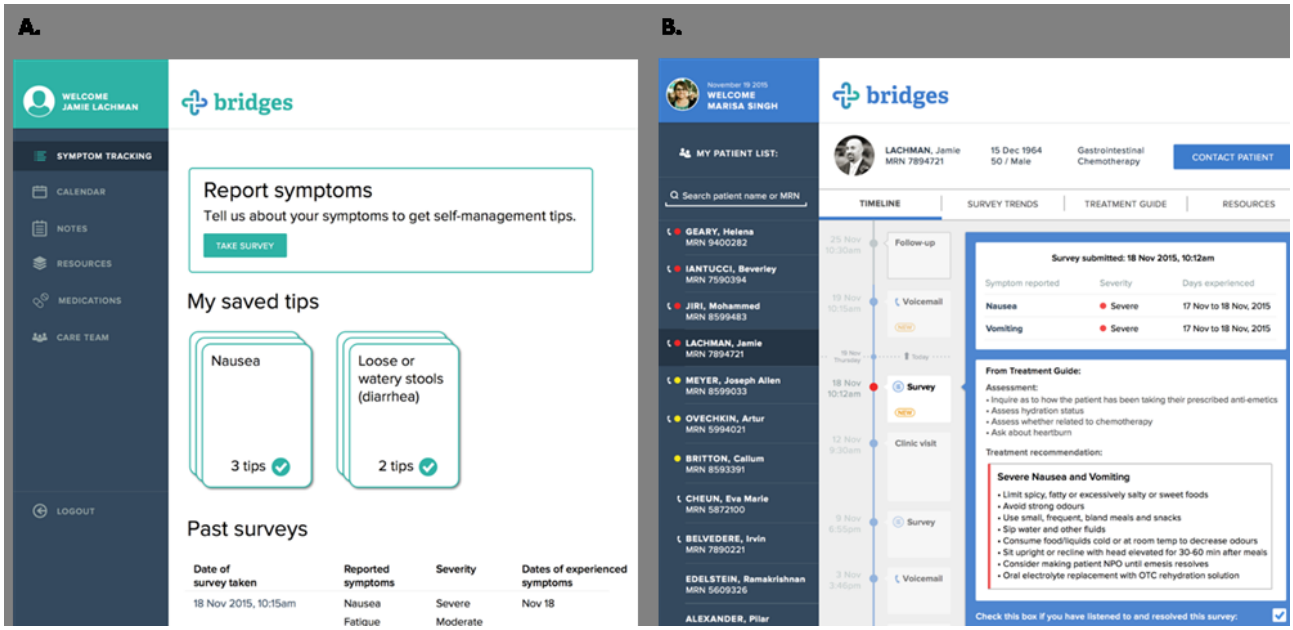
[b]Not applicable.

Figure 2. Summary of findings from thematic analysis of user needs data. HCP: health care providers.



XSL•FO
RenderX

Figure 3. Select screenshots of bridges from the patient/caregiver (panel A) and health care provider (panel B) user interfaces.



Textbox 1. Patient and health care provider (HCP) themes from usability testing.

Patients

- Responsible for self-organizing care
- Want appointment calendar synced with hospital
- Want information specific to their cancer and treatment
- Practicalities of using online application—if unwell, want caregivers to fill out
- Heath care is dynamic—care teams often large and ever-changing

Health care providers

- Integration into existing work flows and practices
- Provision of context to patient feedback as notes
- Privacy and consent issues—sharing patient information, especially with caregivers
- Prioritizing incoming information—deal with most serious problems first
- Consistent assessments and provision of standardized management information (ie, treatment guidelines)

Discussion

Principal Findings

In our study, patients, caregivers, and HCPs engaged in designing an interactive Web-based prototype of a chemotherapy-related toxicity management tool. Issues within the existing systems identified by participants included the need for HCP-supported decision making and self-management strategies for patients, access to credible and timely information, improved communication (both between patients and HCPs, and among HCPs), and integration of the tool within existing workflows to prevent redundancy and confidentiality issues. Understanding local context including the required features and functionalities and potential implementation issues particular to the local context is essential for long-term success. Our study identified a number of additional functionalities and implementation challenges that add to this growing body of literature such as concerns from providers regarding accountability within the tool.

Comparison With Previous Work

Functionalities of existing tools for patients undergoing cancer treatment that have been developed and pilot tested range from symptom tracking alone [24-30], symptom tracking in conjunction with self-care support [31-36], and symptom tracking with active symptom monitoring [12,15,37-47]. Our findings confirm previous work, some of which was published before and some during the conduct of our study, such as high levels of patient satisfaction with remote symptom monitoring systems and the need for ease of use of the tool. When using any tool designed to improve the quality of patient care, the local context is a critical factor that must be addressed, as it will determine the success of implementation. Our study also identified additional functionalities and implementation challenges that add to this growing body of literature including

XSL•FO
RenderX

the need to streamline etool integration into patient self-care and current clinical practice workflows before wide-scale adoption and accountability for provider actions within the etool necessitating clearly defined roles and responsibilities of users. This complements findings from a previous study by Mooney et al [48] utilizing a telephone-based remote monitoring system for cancer patients that highlighted the need to explicitly address who is going to address alerts, as assuming providers will add these duties on to their existing responsibilities does not work.

The evidence for positive impact of etools for symptom monitoring during cancer treatment is growing. A recent study by Denis et al [46] of Web-based follow-up of self-reported symptoms in lung cancer patients following initial treatment reported improved overall survival. Basch et al [15,16] found that remote monitoring of patient symptoms during chemotherapy resulted in significantly improved quality of life, fewer ED visits and hospitalizations, and was associated with prolonged survival. Both interventions included health team alerting functionality for severe symptoms. Likewise, Chumbler et al [49] found a significant reduction in clinic visits and chemotherapy-related hospitalizations using a Cancer Care Coordination Home Telehealth approach.

Limitations

Our findings need to be evaluated in the context of study limitations. Our participants were from 2 large urban centers in southern Ontario and, thus, may not represent end-user needs from other jurisdictions. Selection bias may be present as HCPs who volunteered to participate may have been inherently more interested in using HIT solutions. Likewise, patients who contacted the study coordinator directly may represent a more motivated population than a random sample and may not be representative of the views of the overall population. HCP and patient participants were mostly female (HCPs: 16 vs 3; patients/caregivers: 15 vs 10), so findings may not adequately represent the views of their male counterparts. Our patient and caregiver participants had a median age of 55 and reported having at least a college education (81% [21/26]), English as their first language (92% [24/26]), and comfort using the internet and technology. As such, additional work is needed to better understand the needs of end users who are older, non-English speaking, less educated, or less familiar with technology. The design of our interactive prototype is a first step in building an effective electronic tool for patient care. Additional studies are needed to explore whether the tool impacts patient outcomes; although, the recently published trials from the United States [15] and France [46] provide important information regarding the benefits of such systems for patients in controlled settings. Data on wide-scale adoption of etools in this setting are not yet available.

The method of affinity diagramming aims to bring together issues and insights from various stakeholders, from which overarching themes emerge. Although the data were collected using a number of approaches to understand the complexities and subtleties of toxicity management within the Canadian context, it could be argued that the resulting artifact from affinity diagraming affords only a thin description of toxicity management, that is, a limitation of the method is that the resulting affinity diagram is only a brief summary of the themes that emerged, lacking rich context. However, it should be noted that the human factors study members who conducted the ethnographic field study and participated in ideation sessions were also those who designed the prototype. Thus, although the themes from affinity diagramming informed the design of the prototypes, the human factors study members were also able to draw from the contextual data that they experienced firsthand. Although the number of participants required for usability testing has been well-established [11], there is little consensus on the number of participants required for interviews and focus groups [21,50,51]. Given this limitation, the authors instead strived for saturation of the collected data and resulting themes [21] and triangulation using multiple sources of data (field observation, interviews, and focus groups) [51] to arrive at a comprehensive understanding of toxicity management from various stakeholders.

Conclusions

We have shown that using a human factors design approach for a Web-based application to support management of chemotherapy-related toxicities has the potential to address gaps in cancer care. As these gaps were identified within the local context of care, the design, iterated upon through prototyping and usability testing, seeks to address these needs directly. Our study highlighted that operationalizing a Web-based tool has significant system implications, including assigning responsibility for monitoring the tool to appropriate HCPs and the need to embed the tool into existing workflows and systems. Integration of new etools into self-care by patients and practice is a key issue that needs to be addressed for wide-scale adoption.

Acknowledgments

Funding for this study was provided by an Innovation Planning Fund grant from the Department of Medicine, University of Toronto, funds from the Princess Margaret Cancer Centre Foundation, and a Health Services Research Grant from the Ontario Institute for Cancer Research. Early versions of this work were presented at the American Society of Clinical Oncology Quality Symposium in February 2016, Phoenix, AZ; the Cancer Care Ontario Research Day in April 2016, Toronto, ON; and at the Applied Research in Cancer Control Conference in May 2016, Toronto, ON.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Prince RM, Atenafu EG, Krzyzanowska MK. Hospitalizations during systemic therapy for metastatic lung cancer: a systematic review of real world vs clinical trial outcomes. JAMA Oncol 2015 Dec;1(9):1333-1339. [doi: 10.1001/jamaoncol.2015.3440] [Medline: 26378774]
2. Hassett MJ, O'Malley AJ, Pakes JR, Newhouse JP, Earle CC. Frequency and cost of chemotherapy-related serious adverse effects in a population sample of women with breast cancer. J Natl Cancer Inst 2006 Aug 16;98(16):1108-1117. [doi: 10.1093/jnci/djj305] [Medline: 16912263]
3. Cancer Quality Council of Ontario. 2018. Unplanned Hospital Visits During Chemotherapy URL: http://www.csqi.on.ca/by_patient_journey/treatment/unplanned_hospital_visits_during_chemotherapy/ [WebCite Cache ID 76qhLFbn7]
4. Prince RM, Powis M, Zer A, Atenafu EG, Krzyzanowska MK. Hospitalisations and emergency department visits in cancer patients receiving systemic therapy: Systematic review and meta-analysis. Eur J Cancer Care (Engl) 2019 Jan;28(1):e12909. [doi: 10.1111/ecc.12909] [Medline: 30238542]
5. Given CW, Bradley C, You M, Sikorskii A, Given B. Costs of novel symptom management interventions and their impact on hospitalizations. J Pain Symptom Manage 2010 Apr;39(4):663-672 [FREE Full text] [doi: 10.1016/j.jpainsymman.2009.07.014] [Medline: 20413054]
6. Carlson LE, Speca M, Hagen N, Taenzer P. Computerized quality-of-life screening in a cancer pain clinic. J Palliat Care 2001;17(1):46-52. [Medline: 11324185]
7. Gaertner J, Elsner F, Pollmann-Dahmen K, Radbruch L, Sabatowski R. Electronic pain diary: a randomized crossover study. J Pain Symptom Manage 2004 Sep;28(3):259-267. [doi: 10.1016/j.jpainsymman.2003.12.017] [Medline: 15336338]
8. Mullen KH, Berry DL, Zierler BK. Computerized symptom and quality-of-life assessment for patients with cancer part II: acceptability and usability. Oncol Nurs Forum 2004 Sep;31(5):E84-E89. [doi: 10.1188/04.ONF.E84-E89] [Medline: 15378105]
9. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain 2001 Apr;91(3):277-285. [Medline: 11275385]
10. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. BMJ 2002 May 18;324(7347):1193-1194 [FREE Full text] [Medline: 12016186]
11. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76 [FREE Full text] [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
12. Kearney N, McCann L, Norrie J, Taylor L, Gray P, McGee-Lennon M, et al. Evaluation of a mobile phone-based, advanced symptom management system (ASyMS) in the management of chemotherapy-related toxicity. Support Care Cancer 2009 Apr;17(4):437-444. [doi: 10.1007/s00520-008-0515-0] [Medline: 18953579]
13. Apple. 2015. Cancer Emergency Response Tool- CERT URL: https://itunes.apple.com/gb/app/cancer-emergency-response-tool-cert/id711709486?mt=8 [accessed 2017-12-20] [WebCite Cache ID 6vrDn46iD]
14. Børøsund E, Cvancarova M, Moore SM, Ekstedt M, Ruland CM. Comparing effects in regular practice of e-communication and web-based self-management support among breast cancer patients: preliminary results from a randomized controlled trial. J Med Internet Res 2014 Dec 18;16(12):e295 [FREE Full text] [doi: 10.2196/jmir.3348] [Medline: 25525672]
15. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
16. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. J Am Med Assoc 2017 Dec 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
17. Blomberg J, Giacomi J, Mosher A, Swenton-Wall P. Ethnographic field methods and their relation to design. In: Participatory Design: Perspectives on Systems Design. Florida: CRC Press; 1993.
18. Carayon P, Kianfar S, Li Y, Xie A, Alyousef B, Wooldridge A. A systematic review of mixed methods research on human factors and ergonomics in health care. Appl Ergon 2015 Nov;51:291-321 [FREE Full text] [doi: 10.1016/j.apergo.2015.06.001] [Medline: 26154228]
19. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. Value Health 2009;12(8):1075-1083 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00603.x] [Medline: 19804437]
20. Holtzblatt K, Beyer H. Contextual Design. In: The Encyclopedia of Human-Computer Interaction, 2nd Ed. Soegaard M, Dam RF. editors. The Interaction Design Foundation. Denmark: Interaction Design Foundation; 2011.
21. Morse JM. The significance of saturation. Qual Health Res 1995 May 01;5(2):147-149. [doi: 10.1177/104973239500500201]
22. Yardley L, Morrison LG, Andreou P, Joseph J, Little P. Understanding reactions to an internet-delivered health-care intervention: accommodating user preferences for information provision. BMC Med Inform Decis Mak 2010 Sep 17;10:52 [FREE Full text] [doi: 10.1186/1472-6947-10-52] [Medline: 20849599]
23. Wickens CD, Hollands JG, Banbury S, Parasuraman R. Engineering Psychology And Human Performance. London, UK: Psychology Press; 2013.

XSL•FO
RenderX

24. Basch E, Artz D, Dulko D, Scher K, Sabbatini P, Hensley M, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. J Clin Oncol 2005 May 20;23(15):3552-3561 [FREE Full text] [doi: 10.1200/JCO.2005.04.275] [Medline: 15908666]
25. Basch E, Artz D, Iasonos A, Speakman J, Shannon K, Lin K, et al. Evaluation of an online platform for cancer patient self-reporting of chemotherapy toxicities. J Am Med Inform Assoc 2007;14(3):264-268 [FREE Full text] [doi: 10.1197/jamia.M2177] [Medline: 17329732]
26. Bock M, Moore D, Hwang J, Shumay D, Lawson L, Hamolsky D, et al. The impact of an electronic health questionnaire on symptom management and behavior reporting for breast cancer survivors. Breast Cancer Res Treat 2012 Aug;134(3):1327-1335. [doi: 10.1007/s10549-012-2150-1] [Medline: 22798157]
27. Judson TJ, Bennett AV, Rogak LJ, Sit L, Barz A, Kris MG, et al. Feasibility of long-term patient self-reporting of toxicities from home via the internet during routine chemotherapy. J Clin Oncol 2013 Jul 10;31(20):2580-2585 [FREE Full text] [doi: 10.1200/JCO.2012.47.6804] [Medline: 23733753]
28. Macpherson CF, Linder LA, Ameringer S, Erickson J, Stegenga K, Woods NF. Feasibility and acceptability of an iPad application to explore symptom clusters in adolescents and young adults with cancer. Pediatr Blood Cancer 2014 Nov;61(11):1996-2003. [doi: 10.1002/pbc.25152] [Medline: 25066927]
29. Snyder CF, Blackford AL, Aaronson NK, Detmar SB, Carducci MA, Brundage MD, et al. Can patient-reported outcome measures identify cancer patients' most bothersome issues? J Clin Oncol 2011 Mar 20;29(9):1216-1220. [doi: 10.1200/JCO.2010.33.2080] [Medline: 21343558]
30. Snyder CF, Blackford AL, Wolff AC, Carducci MA, Herman JM, Wu AW, PatientViewpoint Scientific Advisory Board. Feasibility and value of PatientViewpoint: a web system for patient-reported outcomes assessment in clinical practice. Psychooncology 2013 Apr;22(4):895-901 [FREE Full text] [doi: 10.1002/pon.3087] [Medline: 22544513]
31. Berry DL, Hong F, Halpenny B, Partridge A, Fox E, Fann JR, et al. The electronic self report assessment and intervention for cancer: promoting patient verbal reporting of symptom and quality of life issues in a randomized controlled trial. BMC Cancer 2014 Jul 12;14:513 [FREE Full text] [doi: 10.1186/1471-2407-14-513] [Medline: 25014995]
32. Berry DL, Hong F, Halpenny B, Partridge AH, Fann JR, Wolpin S, et al. Electronic self-report assessment for cancer and self-care support: results of a multicenter randomized trial. J Clin Oncol 2014 Jan 20;32(3):199-205 [FREE Full text] [doi: 10.1200/JCO.2013.48.6662] [Medline: 24344222]
33. Berry DL, Blonquist TM, Patel RA, Halpenny B, McReynolds J. Exposure to a patient-centered, web-based intervention for managing cancer symptom and quality of life issues: impact on symptom distress. J Med Internet Res 2015 Jun 03;17(6):e136 [FREE Full text] [doi: 10.2196/jmir.4190] [Medline: 26041682]
34. Head BA, Keeney C, Studts JL, Khayat M, Bumpous J, Pfeifer M. Feasibility and acceptance of a telehealth intervention to promote symptom management during treatment for head and neck cancer. J Support Oncol 2011 Jan 01;9(1):e1-e11 [FREE Full text] [doi: 10.1016/j.suponc.2010.12.006] [Medline: 21499540]
35. Klasnja P, Hartzler A, Powell C, Pratt W. Supporting cancer patients' unanchored health information management with mobile technology. AMIA Annu Symp Proc 2011;2011:732-741 [FREE Full text] [Medline: 22195130]
36. McGee M, Gray P. A handheld chemotherapy symptom management system: results from a preliminary outpatient field trial. Health Informatics J 2016 Jul 25;11(4):243-258. [doi: 10.1177/1460458205055686]
37. Chan MF, Ang E, Duong MC, Chow YL. An online symptom care and management system to monitor and support patients receiving chemotherapy: a pilot study. Int J Nurs Pract 2013 Feb;19(Suppl 1):14-18. [doi: 10.1111/ijn.12020] [Medline: 23425375]
38. Chan MF, Ang NK, Cho AA, Chow YL, Taylor B. Online chemotherapy symptom care and patient management system: an evaluative study. Comput Inform Nurs 2014 Feb;32(2):75-83. [doi: 10.1097/CIN.0000000000000036] [Medline: 24378370]
39. Maguire R, McCann L, Miller M, Kearney N. Nurse's perceptions and experiences of using of a mobile-phone-based Advanced Symptom Management System (ASyMS) to monitor and manage chemotherapy-related toxicity. Eur J Oncol Nurs 2008 Sep;12(4):380-386. [doi: 10.1016/j.ejon.2008.04.007] [Medline: 18539527]
40. Maguire R, Miller M, Sage M, Norrie J, McCann L, Taylor L, et al. Results of a UK based pilot study of a mobile phone based advanced symptom management system (ASyMS) in the remote monitoring of chemotherapy related toxicity. Clin Eff Nurs 2005 Sep;9(3-4):202-210. [doi: 10.1016/j.cein.2006.08.013]
41. Maguire R, Ream E, Richardson A, Connaghan J, Johnston B, Kotronoulas G, et al. Development of a novel remote patient monitoring system: the advanced symptom management system for radiotherapy to improve the symptom experience of patients with lung cancer receiving radiotherapy. Cancer Nurs 2015;38(2):E37-E47. [doi: 10.1097/NCC.0000000000000150] [Medline: 24836956]
42. McCann L, Maguire R, Miller M, Kearney N. Patients' perceptions and experiences of using a mobile phone-based advanced symptom management system (ASyMS) to monitor and manage chemotherapy related toxicity. Eur J Cancer Care (Engl) 2009 Mar;18(2):156-164. [doi: 10.1111/j.1365-2354.2008.00938.x] [Medline: 19267731]
43. Ruland CM, Andersen T, Jeneson A, Moore S, Grimsbø GH, Børøsund E, et al. Effects of an internet support system to assist cancer patients in reducing symptom distress: a randomized controlled trial. Cancer Nurs 2013;36(1):6-17. [doi: 10.1097/NCC.0b013e31824d90d4] [Medline: 22495503]

XSL•FO
RenderX

44. Weaver A, Love SB, Larsen M, Shanyinde M, Waters R, Grainger L, et al. A pilot study: dose adaptation of capecitabine using mobile phone toxicity monitoring-supporting patients in their homes. Support Care Cancer 2014 Oct;22(10):2677-2685. [doi: 10.1007/s00520-014-2224-1] [Medline: 24771299]
45. Mooney KH, Beck SL, Wong B, Dunson W, Wujcik D, Whisenant M, et al. Automated home monitoring and management of patient-reported symptoms during chemotherapy: results of the symptom care at home RCT. Cancer Med 2017 Dec;6(3):537-546 [FREE Full text] [doi: 10.1002/cam4.1002] [Medline: 28135050]
46. Denis F, Lethrosne C, Pourel N, Molinier O, Pointreau Y, Domont J, et al. Randomized trial comparing a web-mediated follow-up with routine surveillance in lung cancer patients. J Natl Cancer Inst 2017 Dec 01;109(9). [doi: 10.1093/jnci/djx029] [Medline: 28423407]
47. Holch P, Warrington L, Bamforth LC, Keding A, Ziegler LE, Absolom K, et al. Development of an integrated electronic platform for patient self-report and management of adverse events during cancer treatment. Ann Oncol 2017 Sep 01;28(9):2305-2311 [FREE Full text] [doi: 10.1093/annonc/mdx317] [Medline: 28911065]
48. Mooney KH, Beck SL, Friedman RH, Farzanfar R, Wong B. Automated monitoring of symptoms during ambulatory chemotherapy and oncology providers' use of the information: a randomized controlled clinical trial. Support Care Cancer 2014 Sep;22(9):2343-2350 [FREE Full text] [doi: 10.1007/s00520-014-2216-1] [Medline: 24687538]
49. Chumbler NR, Kobb R, Harris L, Richardson LC, Darkins A, Sberna M, et al. Healthcare utilization among veterans undergoing chemotherapy: the impact of a cancer care coordination/home-telehealth program. J Ambul Care Manage 2007;30(4):308-317. [doi: 10.1097/01.JAC.0000290399.43543.2e] [Medline: 17873662]
50. Kerr C, Nixon A, Wild D. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. Expert Rev Pharmacoecon Outcomes Res 2010 Jun;10(3):269-281. [doi: 10.1586/erp.10.30] [Medline: 20545592]
51. Fusch P, Ness L. Are We There Yet? Data saturation in qualitative research. The Qualitative Report 2005;20(9):1408-1416 [FREE Full text]

Abbreviations

ED: emergency department
HCP: health care provider
HIT: health information technology

Edited by G Eysenbach; submitted 01.02.18; peer-reviewed by R Collado-Borrell, D Howell; comments to author 28.07.18; revised version received 02.11.18; accepted 30.12.18; published 28.03.19.

Please cite as:
Prince RM, Soung Yee A, Parente L, Enright KA, Grunfeld E, Powis M, Husain A, Gandhi S, Krzyzanowska MK
User-Centered Design of a Web-Based Tool to Support Management of Chemotherapy-Related Toxicities in Cancer Patients
J Med Internet Res 2019;21(3):e9958
URL: http://www.jmir.org/2019/3/e9958/
doi:10.2196/jmir.9958
PMID:30920373

©Rebecca M Prince, Anthony Soung Yee, Laura Parente, Katherine A Enright, Eva Grunfeld, Melanie Powis, Amna Husain, Sonal Gandhi, Monika K Krzyzanowska. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

eHealth Apps Replacing or Complementing Health Care Contacts: Scoping Review on Adverse Effects

Wilhelmina Josepha Maria Stevens[1], MRes; Rob van der Sande[1], PhD; Lilian J Beijer[1], PhD; Maarten GM Gerritsen, MD; Willem JJ Assendelft[2], MD, PhD

[1]Faculty of Health, Hogeschool van Arnhem en Nijmegen University of Applied Sciences, Nijmegen, Netherlands

[2]Department for Primary and Community Care, Radboud University Medical Centre, Nijmegen, Netherlands

Corresponding Author:
Wilhelmina Josepha Maria Stevens, MRes
Faculty of Health
Hogeschool van Arnhem en Nijmegen University of Applied Sciences
Kapittelweg 33
Nijmegen, 6525 EN
Netherlands
Phone: 31 651450260
Email: marjo.stevens@han.nl

Abstract

Background: The use of eHealth has increased tremendously in recent years. eHealth is generally considered to have a positive effect on health care quality and to be a promising alternative to face-to-face health care contacts. Surprisingly little is known about possible adverse effects of eHealth apps.

Objective: We conducted a scoping review on empirical research into adverse effects of eHealth apps that aim to deliver health care at a distance. We investigated whether adverse effects were reported and the nature and quality of research into these possible adverse effects.

Methods: For this scoping review, we followed the five steps of Arksey and O'Malley's scoping review methodology. We searched specifically for studies into eHealth apps that replaced or complemented the face-to-face contact between a health professional and a patient in the context of treatment, health monitoring, or supporting self-management. Studies were included when eHealth and adverse effects were mentioned in the title or abstract and when empirical data on adverse effects were provided. All health conditions, with the exception of mental health conditions, all ages, and all sample sizes were included. We examined the literature published between December 2012 and August 2017 in the following databases: PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science, and the Cochrane Library. The methodological quality of the studies was assessed using the Critical Appraisal Skills Programme (CASP) checklists.

Results: Our search identified 79 papers that were potentially relevant; 11 studies met our inclusion criteria after screening. These studies differed in many ways and the majority were characterized by small research populations and low study quality. Adverse effects are rarely subject to systematic scientific research. So far, information on real adverse effects is mainly limited to incidental reporting or as a bycatch from qualitative pilot studies. Despite the shortage of solid research, we found some indications of possible negative impact on patient-centeredness and efficiency, such as less transparency in the relationship between health professionals and patients and time-consuming work routines.

Conclusions: There is a lack of high-quality empirical research on adverse effects of eHealth apps that replace or complement face-to-face care. While the development of eHealth apps is ongoing, the knowledge with regard to possible adverse effects is limited. The available research often focuses on efficacy, added value, implementation issues, use, and satisfaction, whereas adverse effects are underexplored. A better understanding of possible adverse effects could be a starting point in improving the positive impact of eHealth-based health care delivery.

(J Med Internet Res 2019;21(3):e10736) doi:10.2196/10736

KEYWORDS

eHealth; adverse effects; scoping review

XSL•FO
RenderX

Introduction

The use of eHealth has increased considerably in recent years. eHealth comprises all kinds of information and communication technologies, such as websites and apps for screening, assessment and self-monitoring, health promotion, physical training, and social support (eg, video-chat sessions with a therapist, moderated bulletin boards, chat rooms, and social media) [1,2]. eHealth is generally considered to have a positive effect on health care quality and to be a promising alternative to face-to-face health care contacts [3]. Moreover, the use of eHealth apps is expected to reduce health care consumption and health care costs [4]. Also, eHealth is supposed to contribute to the fast availability of updated medical information, as well as to the provision of tailored care, independent of place and time [5]. In addition, although research in the field is not conclusive, eHealth may improve self-management, health literacy, and healthy behavior [6,7].

In line with these high expectations, large companies and health care organizations invest many millions of dollars in the development of eHealth apps [8,9]. Given the large investments into eHealth, it is remarkable that there is a lack of information with regard to possible unfavorable outcomes of eHealth interventions. It is not known whether there are any adverse effects or whether they are not included in scientific research. The limited number of studies in this field have reported, for example, adverse events such as patients' anxiety caused by monitoring vital signs [10]. In another publication, it was mentioned that professionals get overwhelmed by the amount of data, workload, and workarounds [11].

We conducted a scoping review on empirical studies into adverse effects of eHealth apps that aim to deliver health care at a distance. We welcome the advantages that eHealth interventions will bring; a better understanding of what is known about possible adverse effects will help to improve the use of eHealth.

Methods

Overview of Scoping Review Methodology

For this scoping review, we followed the five steps according to Arksey and O'Malley's scoping review methodology, supplemented with recommendations by Levac et al and Daudt et al [12-14]. We used this method because the aim of a scoping review is to assess the available research literature in order to chart the nature, range, and extent of the research evidence and to identify gaps in the existing literature [15].

Step 1: Identifying the Research Question

In our study, we defined eHealth as the use of information and communication technologies to support or improve health care and health care delivery [1,16]. We searched specifically for studies into eHealth apps that replaced or complemented the face-to-face contact between a health professional and a patient in the context of treatment, health monitoring, supporting self-management, or their communication [17]. Examples are online tools for patients with a chronic condition that replace some of the outpatient checks with online care or online rehabilitation programs that integrate outpatient treatment with exercises performed at home. An adverse effect was defined as any unfavorable outcome on the quality of care that occurred as a result of the use of an eHealth intervention [18].

The following research questions were formulated:

1. Which adverse effects of eHealth apps are reported in empirical studies?
2. What is the nature and quality of the research into the adverse effects of eHealth apps?

Step 2: Identifying Relevant Studies

We examined the literature published between December 2012 and August 2017 in the following databases: PubMed, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science, and the Cochrane Library. We used RefWorks 2.0 (ProQuest), a Web-based bibliographic manager, to import all citations. After an initial broad search and consultations with a librarian, the final search query was composed; we used the set of comprehensive Medical Subject Headings (MeSH) and the free-text search term *eHealth* or its synonyms, as well as the term *adverse effects* or its synonyms. To find as many relevant articles as possible, we decided to add *quality of care* and *risks* as title words because both appeared to be related to articles on adverse effects (see Multimedia Appendix 1 for search query).

Step 3: Selecting Studies

The primary search resulted in 6010 records. After removing duplicates, 5523 titles and abstracts were screened for relevance and for the inclusion and exclusion criteria by two researchers, independently (WJMS and MGMG). Our aim was to include articles that described apps that replaced or complemented the face-to-face contact between a health professional and a patient in the context of treatment, health monitoring, supporting self-management, or their communication. Studies were also included when eHealth and adverse effects were mentioned in the title or abstract and when empirical data—qualitative or quantitative—on adverse effects were provided. Adverse effects had to be related to patients or quality of care.

We wanted a broad search and, at the same time, homogeneity in apps. Mental health conditions were excluded from our scoping study, as the eHealth studies found were already very diverse and this would make the diversity too great. We have tried to achieve as much homogeneity as possible. Except for mental health conditions, all other health conditions, ages, and sample sizes were included. The papers had to be written in English and published between December 2012 and August 2017. We did not use studies in the field of public health or studies on electronic health records, electronic medical records, education, or surgical technology. Exchange of patient information between stakeholders generates different problems and challenges, such as technical matters and privacy issues. We also excluded adverse effects related to security and privacy of data storage and transmission. Prior to inclusion, two of the authors (WJMS and MGMG) verified their agreement in applying the inclusion criteria. Disagreement was solved by discussion. In case no consensus was reached, a third expert

XSL•FO
RenderX

(RvdS) was consulted. This resulted in 79 studies meeting the inclusion criteria.

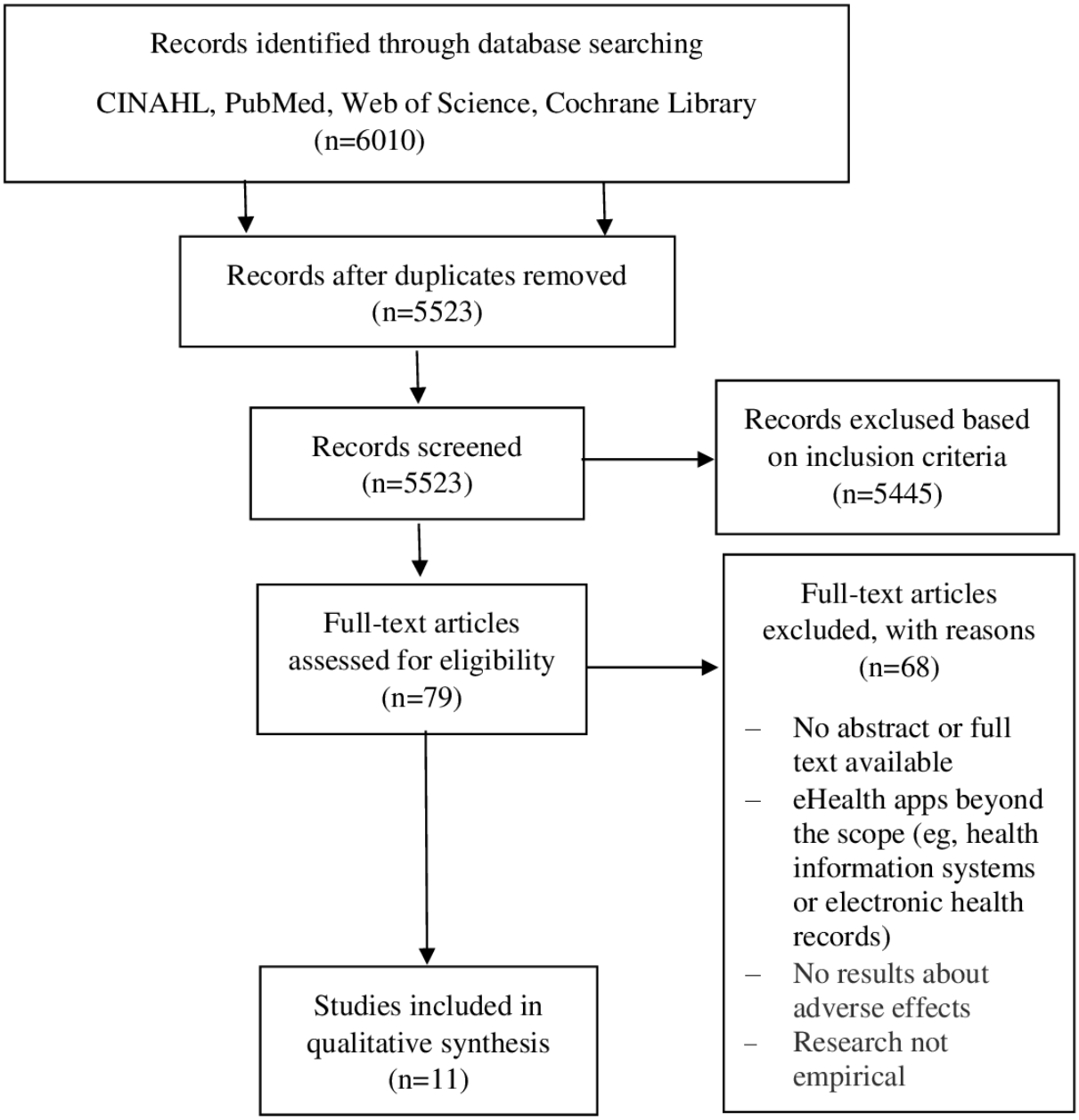
Two of the authors (WJMS and MGMG) subsequently screened the full text of the selected articles, independently, for information on adverse events. The screening results were compared and any discrepancies were resolved by discussion. If the outcome was unclear, two other authors (LJB and RvdS) from the research team arbitrated. For the final synthesis, we excluded 68 studies that did not meet the inclusion criteria, leaving 11 studies for final synthesis (see Figure 1). Study quality was independently evaluated by two researchers (WJMS and MGMG) and disagreements were resolved through discussion in order to reach final study-quality ratings.

Step 4: Charting the Data

We extracted and summarized information for author, year, geographic area, title, name and function of the intervention, study population, study design, outcome and measurements, results, and conclusions of adverse effects. All articles were assessed and data were extracted independently by WJMS and MGMG.

We used the Critical Appraisal Skills Programme (CASP) Qualitative Research Checklist and the Randomised Controlled Trial Checklist for appraisal [19,20]. WJMS and MGMG assessed the studies independently and any disagreement was resolved by discussion. Scores were displayed as the proportion of the number items filled in with *yes* in relation to the total number of items (see Multimedia Appendix 2 for overview).

Figure 1. Study selection process. CINAHL: Cumulative Index to Nursing and Allied Health Literature.



XSL•FO
RenderX

Step 5: Collating, Summarizing, and Reporting the Results

In this phase, results related to adverse effects were discussed by the authors WJMS and MGMG and by the experts RvdS, LJB, and WJJA. Results were categorized following the six domains of quality as formulated by the Agency for Healthcare Research and Quality (AHRQ): safety , effectiveness, patient-centeredness, timeliness, efficiency, and equitability [21,22].

Results

Overview

We found 11 studies that met our inclusion criteria. The characteristics of included studies are summarized in Multimedia Appendix 3. A total of 9 studies out of 11 (82%) used qualitative data. Observation, semistructured interviews, questionnaires, and documentary study are the forms of qualitative research that we have come across in the selected studies.

General Characteristics

eHealth Terminology and Functions

Different terms were used for the eHealth app within articles: telerehabilitation [23], telecare [24,25], telemonitoring [26], telemedicine consultation [27], video telehealth [28], video teleconsultation [29], Internet intervention [30], eVisit [31], mobile health [32], and digital communication [33].

These various apps served different functions: supporting exercise [23,24,32], (video)consultation [25,27,29,33], supporting self-management [25,26,30], triage [28], and primary care [31].

Study Participants

Study participants were patients [23,27,29-32], health care professionals [25], or a mix of patients and health care professionals [24,26,28,33].

A total of 8 out of 11 studies (73%) involved an intervention targeting patients with a chronic health problem [23-26,29,30,32,33].

Study Design and Quality

A total of 8 out of 11 studies (73%) had a quantitative design, of which 2 studies (25%) were randomized controlled trials (RCTs). Multimedia Appendix 2 shows an overview of the quality appraisal.

Study quality, in general, was very low. The sample size in the studies varied from 2 [28] to 564 [31] participants. There were 2 studies out of 11 (18%) that explicitly used adverse effects, a priori, as a primary outcome measure.

Adverse Effects

Overview

Table 1 shows study results from studies about adverse effects categorized following the AHRQ six domains of quality.

Safety

Griffiths et al performed a study into the use of text messages, email, and social media in the communication between young people and a clinical team. They found an increased risk of communication failures, failure to record the content of the communication, and failure to consult the patient's notes prior to engaging in communication [33].

Furthermore, in the studies of Benvenutti et al and Buvik et al, safety was part of the outcome. In both studies, the eHealth intervention was judged to be safe because no study-related adverse events were observed among, or reported by, study participants [23,27].

Effectiveness

In an RCT by Petrella et al, an online exercise program for people with metabolic risk was used. After 12 weeks, this program resulted in a lower systolic blood pressure (SBP) in the active control group compared to the eHealth group. By 52 weeks, the reduction in SBP was similar in both groups [32].

Patient-Centeredness

The study by Bodker et al showed that when eHealth interventions aimed to support patients in doing their physical exercise at home, responsibilities between the health professional and patient became less transparent [24].

A total of 3 out of 11 articles (27%) reported a lack of human face-to-face contact between the health professional and patient or reported on the impossibility for physical examination, leading to "...a new perceptual distance between patients and therapists" [24,29,30].

eHealth apps that support self-management cause patients to express a sense of loss of privacy and stigmatization [25], loss of trust by patients [30], or poor cooperation or lack of willingness on the part of patients [25,30]. Fairbrother et al explored patient and professional views on self-management in the context of telemonitoring. They reported that professionals expressed concerns about promoting the sick role and creating dependence on telemonitoring and professionals [26].

Efficiency

In the study by Bodker et al, health care professionals mentioned new time-consuming work routines (ie, a significant amount of coordination tasks) [24]. Cady et al investigated triage nurse workflow before and after the implementation of video telehealth. They found an increased triage time [28]. Buvik et al, however, found no differences in consultation duration in their RCT investigation into video-assisted remote orthopedic consultations in an orthopedic outpatient clinic [27].

In a study by Mehrotra et al on a comparison of care between eVisits and physician office visits, physicians using eHealth were less likely to order relevant tests or order preventive care and researchers established the occurrence of overprescribing of antibiotics [31].

XSL•FO
RenderX

Table 1. Reported results from studies about adverse effects.

AHRQ[a] domains of quality	Adverse effects
Safety	No study-related adverse events [23]
	No serious adverse advents were found to be related to the mode of the consultation [27]
	Communication failures [33]
	Failure to record the content of the communication [33]
	Failure to consult the patient's notes prior to engaging in communication [33]
Effectiveness	No difference in re-referrals [27]
	Reduction in SBP[b] greater in control group at 12 weeks and similar at 52 weeks [32]
Patient-centeredness	A new perceptual distance between patients and therapists [24]
	Sense of losing privacy [25]
	Stigmatization [25]
	Poor cooperation of the patient [25]
	Loss of trust [30]
	Lack of human face-to-face contact [30]
	Lack of willingness [30]
	Concerns about promoting the sick role [26]
	Creates dependence on telemonitoring and professionals [26]
	Lack of physical contact causes concerns about long-term complications of diabetes [29]
Timeliness	N/A[c]
Efficiency	New time-consuming work routines [24]
	Responsibilities less transparent [24]
	Mean consultation duration not different [27]
	Increased triage time [28]
	Less likely to order urinary tract infection-relevant tests [31]
	No difference for follow-up visits [31]
	Overprescribing of antibiotics [31]
	Less likely to order preventive care [31]
Equitability	N/A[c]

[a]AHRQ: Agency for Healthcare Research and Quality.

[b]SBP: systolic blood pressure.

[c]N/A: not applicable, as no adverse effects related to timeliness and equitability were reported in the selected studies.

Timeliness and Equitability

Adverse effects related to timeliness and equitability were not reported in the selected studies.

Discussion

Principal Findings

Our scoping review shows that there is a clear lack of empirical research on adverse effects of eHealth apps that replace or complement face-to-face care. After a broad search for empirical studies, we were only able to include 11 studies. These studies differed not only in the function (eg, monitoring or assessment) of the eHealth intervention, but also in study population, methodology, and outcome; the majority of the studies entailed small research populations and low study quality.

Adverse effects are rarely subject to systematic scientific research. So far, information on real adverse effects is mainly limited to incidental reporting or as a bycatch from qualitative pilot studies. The diversity and the low quality among studies in our scoping review do not provide a good understanding of the nature and size of these possible adverse effects and do not offer a good enough understanding to make a meaningful benefit-harm analysis.

Despite this shortage of solid research, we suggest that eHealth may have a negative impact on the transparency of the relationship between health professionals and patients regarding their responsibilities [24]. Furthermore, because in some apps there is no nonverbal communication and no ability to perform a physical examination, health professionals worry about the effect of eHealth on the quality of the communication that will, in turn, affect the quality of care. Confidentiality issues and potential negative feelings can arise as a result of this changing relationship [28]. Patients may overemphasize the impact of their condition by getting *fixated* on readings and the monitoring of data. This can lead to dependence on telemonitoring and professionals [26,33].

Thereby, complex programs of therapeutic exercises delivered by technology had limited success in engaging people in chronic

XSL•FO
RenderX

pain. Patients showed a lack of willingness and engagement because they missed some help and face-to-face acknowledgement or the content did not seem very relevant to them [30].

Furthermore, our findings show that eHealth may have an adverse effect on efficiency because of new time-consuming work routines [24,28]. Bodker et al reported subtle transformations of work activities, such as recruiting patients, conducting home visits to give personalized advice on home training, and invisible work necessary to uphold the telerehabilitation infrastructure.

In the study by Cady et al, time spent on video triage activities was significantly longer than the time spent on equivalent telephone triage. Their workflow analysis revealed that new activities were added, such as preparing for video telehealth sessions, troubleshooting, and the possibility to arrange an appointment with the physician to participate in the session. In addition, the possibility to interact not only with the parent(s), but also with the child during the video telehealth assessment, caused an increased workflow [28].

Limitations

Various limitations of this scoping review need to be considered. Some articles may have been missed when the search was undertaken. Due to the research question, we only searched for empirical studies that primarily focused on adverse effects. Studies may not be explicit about their findings related to adverse effects in the title or the abstract, as it may not be the main goal of the study. In addition, alternative terminology is used, such as *unintended consequences*, *negative effects*, or *quality of care* to report relevant findings , which means this review may not be inclusive of all papers that have reported relevant results.

Furthermore, we searched specifically for eHealth apps that delivered health care at a distance. It is possible that we missed studies that did not meet the inclusion criteria, but that do offer value in understanding the phenomenon of adverse effects of eHealth. Our findings address different eHealth apps, goals, and implementation contexts; different users, communities, and countries; and different chronic conditions. Due to the small number of often poorly qualitative studies and the diversity of the apps examined, we believe that our findings cannot apply to eHealth in general.

Comparison With Prior Literature

The concept *eHealth* is a relatively new way of providing care and is used for different applications, technologies, and care processes. Although the number of articles reporting on eHealth interventions has increased in the past 10 years, it is still a relatively new field of research. Most eHealth interventions are now at a pilot phase and as their implementation is often halted by organizational, cultural, or financial barriers, most studies focus on implementation and organizational issues.

In an overview of systematic reviews of studies into the impact of telehealth care on the quality and safety of care in 2013, McLean et al report, “It was not clear whether adverse events did not occur or whether there was a lack of reporting.” They did not come across any studies that explicitly examined impacts of telehealth care on patient safety [34]. Our findings confirm that we still do not know if there are adverse effects or if the issue of adverse effects is simply not addressed.

Research on the risks of eHealth have mostly focused on factors such as infrastructure, technological issues, implementation issues, and lower adherence. Unfavorable patient outcomes are rarely mentioned [11,34,35].

Recommendations for Future Design and Research

While this scoping review highlights few adverse effects of eHealth interventions, there remains a gap in empirical research that should be addressed in the future. Researchers need to consider and anticipate these adverse effects of eHealth interventions. The changing relationship between, and responsibilities of, health professionals and patients, greater dependence of patients on health care, potential negative feelings, and new time-consuming work routines are important subjects for research in the future. Furthermore, for future research, it is worthwhile to discriminate apps used as replacements or as complements to regular care. A proper insight into the size and nature of adverse effects will only arise if these effects are systematically investigated, preferably by RCTs. We therefore recommend that adverse effects be included as a standard in studies into the effects of eHealth apps.

Conclusions

eHealth may contribute to more accessible and more efficient health care. So far, possible negative effects have not been thoroughly investigated. The little research that has been done suggests that they do exist. Given the rapid expansion of eHealth, there is an urgent need for further research on this issue.

Acknowledgments

Funding for this study was provided by Hogeschool van Arnhem en Nijmegen University of Applied Sciences, the Netherlands. The funders had no role in the study design, data collection, and analysis; decision to publish; or preparation of the manuscript. The authors gratefully thank Thomas Pelgrim and Alice Tillema, the medical information specialists, for their help and advice on search strategies in scientific databases. MGMG was previously affiliated with the Diafaan Foundation, Zevenaar, Netherlands, but has been retired since January 2015.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Search strategies for the different online databases.

[PDF File (Adobe PDF File), 31KB - jmir_v21i3e10736_app1.pdf]

Multimedia Appendix 2

Appraisal of studies by study design using Critical Appraisal Skills Programme (CASP) tools.

[PDF File (Adobe PDF File), 33KB - jmir_v21i3e10736_app2.pdf]

Multimedia Appendix 3

Description of the 11 included articles.

[DOCX File, 27KB - jmir_v21i3e10736_app3.docx]

References

1. Boogerd EA, Arts T, Engelen LJ, van de Belt TH. "What is eHealth": Time for an update? JMIR Res Protoc 2015 Mar 12;4(1):e29 [FREE Full text] [doi: 10.2196/resprot.4065] [Medline: 25768939]
2. Oh H, Rizo C, Enkin M, Jadad A. What is eHealth?: A systematic review of published definitions. World Hosp Health Serv 2005;41(1):32-40. [Medline: 15881824]
3. 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]
4. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: A systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014 Apr 16;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
5. Taylor K. Connected Health: How Digital Technology is Transforming Health and Social Care. London, UK: Deloitte Centre for Health Solutions; 2015. URL: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/life-sciences-health-care/deloitte-uk-connected-health.pdf [accessed 2019-01-22] [WebCite Cache ID 75d3iUyW0]
6. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: A scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
7. Notenboom A, Blankers I, Goudriaan R, Groot W. E-health en Zelfmanagement: Een Panacee voor Arbeidstekorten en Kostenoverschrijdingen in de Zorg?. The Hague, Netherlands: Aarts De Jong Wilms Goudriaan Public Economics bv (APE); 2012 Feb. URL: http://www.ape.nl/include/downloadFile.asp?id=286 [accessed 2019-01-22] [WebCite Cache ID 75d4OgocN]
8. Wouters M, Swinkels I, Sinnige J, de Jong J, Brabers A, van Lettow B, et al. Kies Bewust Voor eHealth. The Hague, Netherlands: Nictiz en het NIVEL; 2017 Nov. URL: https://www.nictiz.nl/wp-content/uploads/2018/03/1_eHealth-monitor-2017.pdf [accessed 2019-01-22] [WebCite Cache ID 75d57wo2J]
9. Wicks P, Stamford J, Grootenhuis MA, Haverman L, Ahmed S. Innovations in eHealth. Qual Life Res 2014 Feb;23(1):195-203 [FREE Full text] [doi: 10.1007/s11136-013-0458-x] [Medline: 23852096]
10. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: A systematic overview. PLoS Med 2011 Jan 18;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
11. Ossebaard H, Bruijn AD, Geertsma R. Magnitude of eHealth technology risks largely unknown. Int J Adv Syst Meas 2013;6(1):57-71.
12. 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]
13. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. Res Synth Methods 2014 Dec;5(4):371-385 [FREE Full text] [doi: 10.1002/jrsm.1123] [Medline: 26052958]
14. Daudt HM, van Mossel C, Scott SJ. Enhancing the scoping study methodology: A large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013 Mar 23;13:48 [FREE Full text] [doi: 10.1186/1471-2288-13-48] [Medline: 23522333]
15. Peters MD, 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]
16. Oh H, Rizo C, Enkin M, Jadad A. What is eHealth (3): A systematic review of published definitions. J Med Internet Res 2005 Feb 24;7(1):e1 [FREE Full text] [doi: 10.2196/jmir.7.1.e1] [Medline: 15829471]

XSL•FO
RenderX

17. Krijgsman J, Wolterink GK. Ordening in de Wereld van eHealth. The Hague, Netherlands: Nictiz; 2012 Aug 24. URL: https://www.nictiz.nl/wp-content/uploads/2012/08/Whitepaper-Ordening-in-de-wereld-van-eHealth.pdf [accessed 2019-01-22] [WebCite Cache ID 75d6NYAKQ]
18. Zorzela L, Loke YK, Ioannidis JP, Golder S, Santaguida P, Altman DG, PRISMAHarms Group. PRISMA harms checklist: Improving harms reporting in systematic reviews. BMJ 2016 Feb 01;352:i157. [doi: 10.1136/bmj.i157] [Medline: 26830668]
19. CASP Qualitative Checklist. Oxford, UK: Critical Appraisal Skills Programme (CASP) URL: https://casp-uk.net/wp-content/uploads/2018/01/CASP-Qualitative-Checklist-2018.pdf [accessed 2019-01-22] [WebCite Cache ID 75d6nBd4J]
20. CASP Randomised Controlled Trial Checklist. Oxford, UK: Critical Appraisal Skills Programme (CASP) URL: https://casp-uk.net/wp-content/uploads/2018/03/CASP-Randomised-Controlled-Trial-Checklist-2018_fillable_form.pdf [accessed 2019-01-22] [WebCite Cache ID 75cY6AlNz]
21. Berwick D. A user's manual for the IOM's 'Quality Chasm' report. Health Aff (Millwood) 2002;21(3):80-90. [doi: 10.1377/hlthaff.21.3.80] [Medline: 12026006]
22. Agency for Healthcare Research and Quality. Rockville, MD: Agency for Healthcare Research and Quality; 2018 Nov. Six domains of health care quality URL: https://www.ahrq.gov/talkingquality/measures/six-domains.html [accessed 2019-01-22] [WebCite Cache ID 75cYMP3w1]
23. Benvenuti F, Stuart M, Cappena V, Gabella S, Corsi S, Taviani A, et al. Community-based exercise for upper limb paresis: A controlled trial with telerehabilitation. Neurorehabil Neural Repair 2014 Sep;28(7):611-620. [doi: 10.1177/1545968314521003] [Medline: 24515928]
24. Bødker M, Juul Nielsen A. Providing rehabilitation online: Invisible work and diagnostic agents. J Health Organ Manag 2015;29(7):948-964. [doi: 10.1108/JHOM-06-2014-0091] [Medline: 26556161]
25. Chang C, Lee T, Chou C, Mills ME. Telecare for diabetes mellitus: Case managers' experiences. Comput Inform Nurs 2013 Oct;31(10):505-511. [doi: 10.1097/01.NCN.0000432128.07268.cc] [Medline: 23958966]
26. Fairbrother P, Pinnock H, Hanley J, McCloughan L, Sheikh A, Pagliari C, et al. Exploring telemonitoring and self-management by patients with chronic obstructive pulmonary disease: A qualitative study embedded in a randomized controlled trial. Patient Educ Couns 2013 Dec;93(3):403-410. [doi: 10.1016/j.pec.2013.04.003] [Medline: 23647981]
27. Buvik A, Bugge E, Knutsen G, Småbrekke A, Wilsgaard T. Quality of care for remote orthopaedic consultations using telemedicine: A randomised controlled trial. BMC Health Serv Res 2016 Dec 08;16:483 [FREE Full text] [doi: 10.1186/s12913-016-1717-7] [Medline: 27608768]
28. Cady RG, Finkelstein SM. Mixed-methods approach for measuring the impact of video telehealth on outpatient clinic triage nurse workflow. Comput Inform Nurs 2013 Sep;31(9):439-449 [FREE Full text] [doi: 10.1097/01.NCN.0000432126.99644.6c] [Medline: 24080753]
29. Fatehi F, Martin-Khan M, Smith AC, Russell AW, Gray LC. Patient satisfaction with video teleconsultation in a virtual diabetes outreach clinic. Diabetes Technol Ther 2015 Jan;17(1):43-48. [doi: 10.1089/dia.2014.0159] [Medline: 25296189]
30. Duggan GB, Keogh E, Mountain GA, McCullagh P, Leake J, Eccleston C. Qualitative evaluation of the SMART2 self-management system for people in chronic pain. Disabil Rehabil Assist Technol 2015 Jan;10(1):53-60. [doi: 10.3109/17483107.2013.845696] [Medline: 24112276]
31. Mehrotra A, Paone S, Martich G, Albert S, Shevchik G. A comparison of care at e-visits and physician office visits for sinusitis and urinary tract infection. JAMA Intern Med 2013 Jan 14;173(1):72-74 [FREE Full text] [doi: 10.1001/2013.jamainternmed.305] [Medline: 23403816]
32. Petrella RJ, Stuckey MI, Shapiro S, Gill DP. Mobile health, exercise and metabolic risk: A randomized controlled trial. BMC Public Health 2014 Oct 18;14:1082 [FREE Full text] [doi: 10.1186/1471-2458-14-1082] [Medline: 25326074]
33. Griffiths F, Bryce C, Cave J, Dritsaki M, Fraser J, Hamilton K, et al. Timely digital patient-clinician communication in specialist clinical services for young people: A mixed-methods study (the LYNC study). J Med Internet Res 2017 Dec 10;19(4):e102 [FREE Full text] [doi: 10.2196/jmir.7154] [Medline: 28396301]
34. McLean S, Sheikh A, Cresswell K, Nurmatov U, Mukherjee M, Hemmi A, et al. The impact of telehealthcare on the quality and safety of care: A systematic overview. PLoS One 2013;8(8):e71238 [FREE Full text] [doi: 10.1371/journal.pone.0071238] [Medline: 23977001]
35. Guise V, Anderson J, Wiig S. Patient safety risks associated with telecare: A systematic review and narrative synthesis of the literature. BMC Health Serv Res 2014 Nov 25;14:588 [FREE Full text] [doi: 10.1186/s12913-014-0588-z] [Medline: 25421823]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
CASP: Critical Appraisal Skills Programme
CINAHL: Cumulative Index to Nursing and Allied Health Literature
MeSH: Medical Subject Headings
N/A: not applicable
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by G Eysenbach; submitted 14.04.18; peer-reviewed by R van der Vaart, A Zeleke, H Al-Samarraie, D Nault; comments to author 25.08.18; revised version received 31.10.18; accepted 22.11.18; published 01.03.19.

Please cite as:
Stevens WJM, van der Sande R, Beijer LJ, Gerritsen MGM, Assendelft WJJ
eHealth Apps Replacing or Complementing Health Care Contacts: Scoping Review on Adverse Effects
J Med Internet Res 2019;21(3):e10736
URL: https://www.jmir.org/2019/3/e10736/
doi:10.2196/10736
PMID:30821690

©Wilhelmina Josepha Maria Stevens, Rob van der Sande, Lilian J Beijer, Maarten GM Gerritsen, Willem JJ Assendelft. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User Experience of an App-Based Treatment for Stress Urinary Incontinence: Qualitative Interview Study

Ina Asklund[1], MD; Eva Samuelsson[1], MD, PhD; Katarina Hamberg[1], MD, PhD; Göran Umefjord[1], MD, PhD; Malin Sjöström[1], MD, PhD

Department of Public Health and Clinical Medicine, Family Medicine, Umeå University, Umeå, Sweden

Corresponding Author:
Ina Asklund, MD
Department of Public Health and Clinical Medicine, Family Medicine
Umeå University
The Research Unit, Region Jämtland Härjedalen
Box 654
Umeå, SE-831 27
Sweden
Phone: 46 704788148
Fax: 46 63177580
Email: ina.asklund@umu.se

Abstract

Background: Stress urinary incontinence (SUI) affects 10%-39% of women. Its first-line treatment consists of lifestyle interventions and pelvic floor muscle training (PFMT), which can be performed supervised or unsupervised. Health apps are increasing in number and can be used to improve adherence to treatments. We developed the Tät app, which provides a 3-month treatment program with a focus on PFMT for women with SUI. The app treatment was evaluated in a randomized controlled trial, which demonstrated efficacy for improving incontinence symptoms and quality of life. In this qualitative interview study, we investigated participant experiences of the app-based treatment.

Objective: This study aimed to explore women's experiences of using an app-based treatment program for SUI.

Methods: This qualitative study is based on telephone interviews with 15 selected women, with a mean age of 47 years, who had used the app in the previous randomized controlled trial. A semistructured interview guide with open-ended questions was used, and the interviews were transcribed verbatim. Data were analyzed according to the grounded theory.

Results: The results were grouped into three categories: "Something new!" "Keeping motivation up!" and "Good enough?" A core category, "Enabling my independence," was identified. The participants appreciated having a new and modern way to access a treatment program for SUI. The use of new technology seemed to make incontinence treatment feel more prioritized and less embarrassing for the subjects. The closeness to their mobile phone and app features like reminders and visual graphs helped support and motivate the women to carry through the PFMT. The participants felt confident that they could perform the treatment program on their own, even though they expressed some uncertainty about whether they were doing the pelvic floor muscle contractions correctly. They felt that the app-based treatment increased their self-confidence and enabled them to take responsibility for their treatment.

Conclusions: Use of the app-based treatment program for SUI empowered the women in this study and helped them self-manage their incontinence treatment. They appreciated the app as a new tool for supporting their motivation to carry through a slightly challenging PFMT program.

Trial Registration: ClinicalTrials.gov NCT01848938; https://clinicaltrials.gov/ct2/show/NCT01848938 (Archived by WebCite at https://clinicaltrials.gov/ct2/show/NCT01848938)

(J Med Internet Res 2019;21(3):e11296) doi:10.2196/11296

KEYWORDS

mobile applications; urinary incontinence; stress; pelvic floor muscle training; self-management; qualitative research; grounded theory; women's health

XSL•FO
RenderX

Introduction

Urinary incontinence (UI) is defined as complaints of any involuntary leakage of urine [1]. There are different types of UI, the most common of which is stress urinary incontinence (SUI). SUI is defined as leakage of urine with exercise, coughing, or sneezing and affects 10%-39% of women [1-3]. Although common, many women do not seek care because they regard their incontinence as "normal" or find it too embarrassing to talk about [4]. The first-line treatment of SUI consists of lifestyle interventions and pelvic floor muscle training (PFMT) and can be initiated among most women without extensive preliminary evaluation [5]. A recent Cochrane review confirms that PFMT is effective and can cure or improve symptoms of SUI, reduce the number of leakage episodes, and improve UI-specific quality of life [6]. There is no consensus on the optimal PFMT program, but it usually involves exercises designed to increase muscle strength, endurance, rapidity, and coordination [7]. Adherence to the exercise program is important for effectiveness, and strategies to increase adherence need more attention among clinicians and researchers [8].

Smartphone use is increasing worldwide, and in 2017, more than 32% of the world's population used a smartphone [9]. In Sweden, 85% of the population owned a smartphone in 2017, and the corresponding number in the United States was 77% in 2018 [10-11]. The number of health apps is increasing rapidly, with 325,000 mobile health apps available in 2017 [12]. Health apps have the potential to increase treatment adherence and improve clinical outcomes in chronic disease management [13]; however, data about user experience with health apps are limited, because many qualitative studies focused on the developmental stage of the apps, not experiences with long-term use [14].

Within the eContinence project [15], the Tät app was developed by ES, MS, and GU in cooperation with engineers at ITS (ICT Services and System Development), Umeå University, Sweden. The app, which provides a treatment program for SUI with a focus on PFMT, is referred to here as the "app-based treatment." We demonstrated efficacy regarding symptoms, quality of life, and urinary leakage in a previous randomized controlled trial (RCT) [16]. The treatment was cost effective [17] and had persistent long-term effects [18].

As part of the evaluation of the app-based treatment, we aimed to learn more about the participants' experiences of using the app and engaging in the PFMT program. We wanted to determine how and why the app-based treatment could be effective. The aim of this study was therefore to explore women's experiences of using an app-based treatment program for SUI.

Methods

Study Setting and Design

This qualitative study was based on telephone interviews in Sweden between January and November 2014. All participants had taken part in our previous RCT (Trial Registration: ClinicalTrials.gov NCT01848938) [16] that evaluated the effect of the mobile app Tät. We received ethical approval from the Regional Ethical Review Board of Umeå University (UMU DNR: 2012-325-31M).

For the RCT, we recruited adult women who experienced SUI at least once a week through our website. Exclusion criteria were pregnancy, previous urinary incontinence surgery, malignancy in the lower abdomen, neurologic disease affecting the lower abdomen or legs, irregular bleeding, visible blood in urine, difficulty passing urine, and severe psychiatric disorders. Participants completed an informed consent form, a Web-based questionnaire, and a 2-day leakage diary before randomization. We randomized 123 women to treatment with the app for 3 months (n=62) or a postponed treatment group that did not receive the app until the follow-up was completed (control group, n=61). During the whole RCT study, from online recruitment to follow-up, we had no face-to-face contact with any of the women in any of the study groups. We did not provide any reimbursements. At 3 months of follow-up, we evaluated the effect of the app-based treatment with Web-based questionnaires. The methods and results from the RCT are presented in detail elsewhere [16].

The app included a treatment program focused on PFMT exercises and contained information about SUI; the pelvic floor; and factors related to incontinence such as overweight, physical activity, smoking, and fluid intake. There were six basic and six advanced PFMT exercises with increasing difficulty, and the participants were advised to proceed according to their own improvement. The exercises included different types of pelvic floor muscle contractions such as quick, strength, and endurance contractions. To illustrate each contraction, a moving graph showed how long and with what intensity the user should contract and relax (Figure 1). Within the app, the participants could set three reminders a day and log the number of exercises performed in a statistics function.

Participants

Participants eligible for this interview study were women from the RCT app group (who received the app directly after randomization) from our previous RCT, who had completed the 3-month follow-up. The first author IA performed the selection and approached the participants by email invitation. Our aim was to conduct the interviews within 3 months after they had completed the follow-up in the RCT in order to minimize the risk of recall bias. In Figure 2, we summarize the flow of the previous RCT study and the present interview study.

We purposefully selected participants with a goal of ensuring diversity of age, residence, and type of mobile phone operating system (Apple or Android). For the last interviews, we also selected two participants based on their treatment outcome in the RCT, because we wanted experiences from women who did not show improvement in incontinence symptoms.

XSL•FO
RenderX

Figure 1. Screenshots of the mobile app Tät version 1.9.

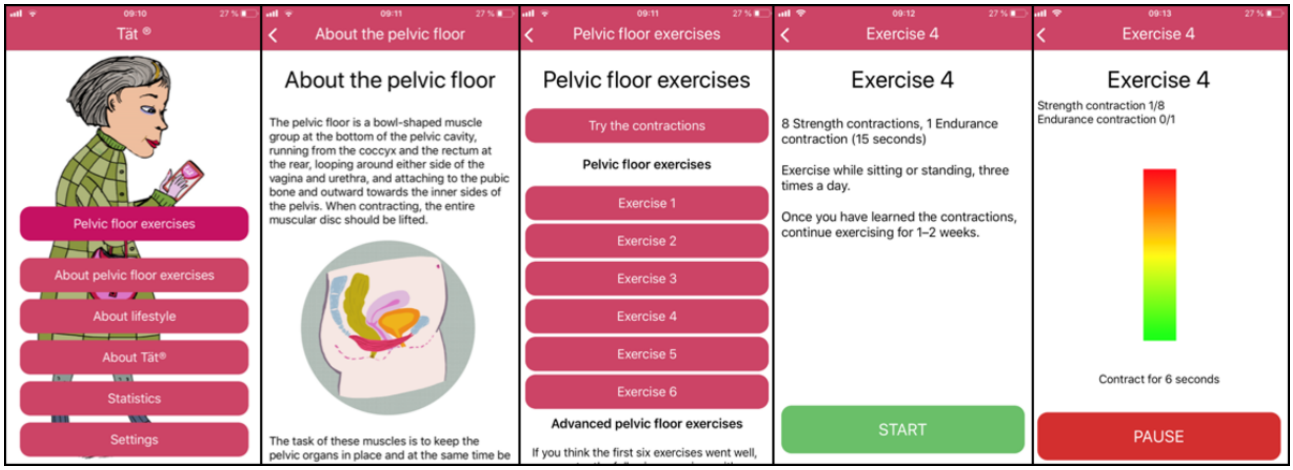
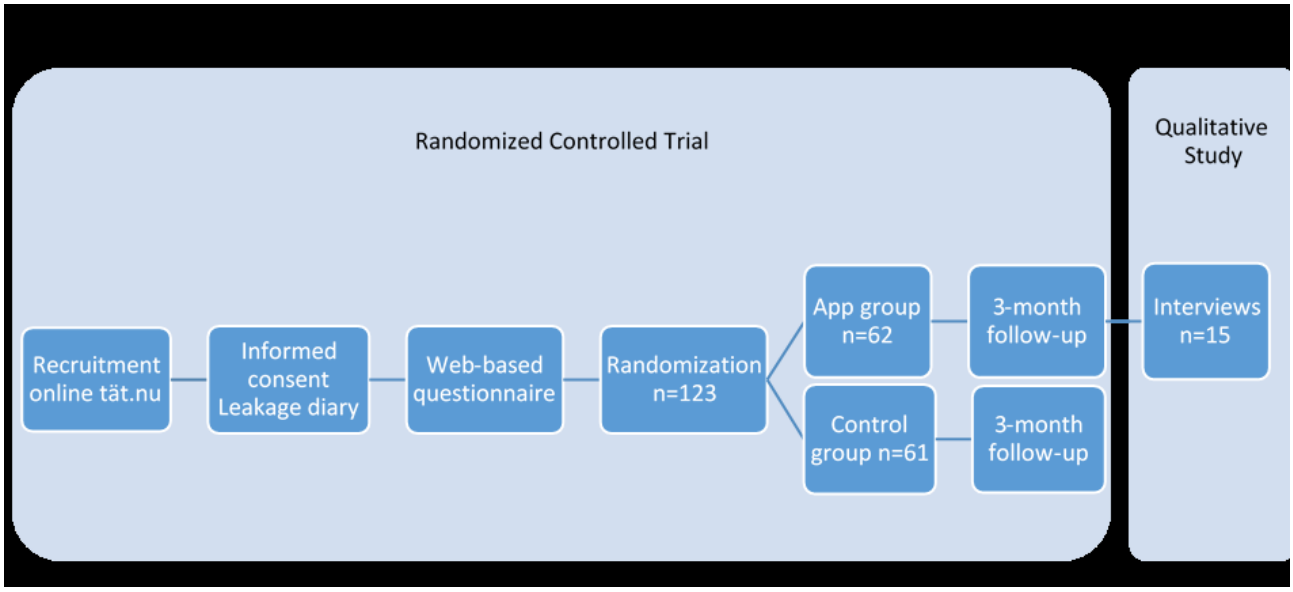


Figure 2. Flow of the randomized controlled study and the present interview study.



Data Collection

Individual telephone interviews were conducted using a semistructured interview guide with open-ended questions. The first author developed the interview guide, which was discussed in the research group to refine the questions. To determine which topics could be worth investigating further, we used information from the participants' answers to open-ended questions about their experiences of the app-based treatment in the follow-up questionnaire from the RCT. Some of the research team also had previous experience in internet-based treatment of incontinence, which influenced our preunderstanding.

The topics covered in the interviews were as follows: the participants' expectations about using an app for treatment, their experiences of using the app, their relationship and interaction with the app during the study, and their experiences performing PFMT independently.

The first author IA, who is a general practitioner, conducted all interviews. As suggested for qualitative analyses, we performed data collection, transcription, and preliminary analysis in parallel [19,20]. This approach made it possible to adapt the interview guide when new, interesting aspects emerged. For example, we added questions to explore in more detail how the participants interacted with the app. We also added questions about the participants' expectations for the app and their views on PFMT, in general.

Interviews were recorded and transcribed verbatim. The first author transcribed 8 of the 15 interviews, and an experienced medical secretary transcribed the remaining 7 interviews. The interviewer wrote short memos after the interviews.

Data Analysis

We performed the analysis according to the constant comparison techniques in the grounded theory [20]. Our goal was to provide an informative, rich description rather than develop a theory. The analysis started directly after the first interview. All authors independently read and coded the second and third interviews according to meaning and content. The codes were then discussed and sorted into different categories in a joint session. Most of the categories were developed at this stage. Interviews 4 to 10 were coded individually by the first and last authors (IA and MS). The last five interviews were coded only by the first author but the last author read them thoroughly, following which, no new categories emerged.

XSL•FO
RenderX

Table 1. Examples of codes, subcategories, and categories derived from a quotation.

Quotation	Code	Subcategory	Category
That was what was so good about this, I can do this at home myself, no need to book an appointment, find the time and suit others.	• Do it at home • Do it myself • Having to book appointments • Finding the time	• Easily accessible • Low threshold	• Something new!
Well, to get it sorted, get a regular reminder, get the support day-in, day-out.	• To get it sorted • Reminders • Daily support	• Motivational support	• Keeping motivation up!
But I feel like I have found my muscles, but that I'm still a little bit uncertain about it, kind of.	• Feel I have found my muscles • Still a bit uncertain	• Insecurity about training correctly	• Good enough?

After these steps, all authors read all interviews and met to discuss the categories in relation to the original research questions and see if the data were sufficient to answer our questions. After two additional meetings, we reached conclusions about the categories. The first author reread all interviews to check for similarities and differences in the experiences described. This step was also a way to see that the categories were well grounded in the data and that there were quotations that supported the findings (Table 1). We assumed that we had reached saturation and included a sufficient number of participants when we did not find any new categories in the analysis and the categories we had found were well developed.

Results

Principal Findings

Fifteen women between the ages of 27 and 72 years were interviewed (mean age, 47 years). They lived in different parts of Sweden—rural areas, smaller towns, and larger cities. Ten of the participants used iPhones and five used Android operating systems on their mobile phones. The mean interview time was 34 minutes (range, 22-55 minutes). Most of the interviews took place within 3 months after participants completed the follow-up questionnaire in the RCT. Two interviews were delayed until 4 months after follow-up.

At the beginning of the interviews, the participants spoke about their incontinence problems as they were when the women entered the study. Many of them described difficulties when exercising as the biggest problem, followed by problems when playing with their children or coughing or sneezing. There was an ambivalence regarding how they viewed the severity of their incontinence: Although they described it as annoying and affecting their everyday lives, they did not call it a big problem or something that restricted their lives.

Most participants had used apps for other purposes such as bank transactions, social media, calendars, and training and had good experiences. They found this technology uncomplicated and useful in their daily lives.

The analysis resulted in a core category called "Enabling my independence." The underlying three categories were called "Something new!" "Keeping motivation up!" and "Good enough?" and were each supported by several subcategories (Textbox 1).

Enabling My Independence

Participants in our study seemed to be health conscious and self-confident, because many of them were actively looking for a self-management program to treat their incontinence. The core category "Enabling my independence" refers to their wish to be independent while needing some external support (like an app) to make it possible. Although disappointed with previous advice and lack of support from health care, the women realized that the only ones responsible for actually performing the PFMT were themselves. They had a strong sense of their own responsibility and felt confident that they could perform the training themselves. The participants were satisfied with trying an app as a tool to support their training and felt that the app helped them realize their wish to follow a PFMT program. It made them feel confident and empowered to know that they had "done what they could" by themselves to treat their incontinence. Although some women expressed a wish for more personalized support, they prioritized the fact that they could handle the treatment independently.

Something New!

New technology, Something New and Modern

Many of the women had searched the internet for new treatment programs and apps addressing PFMT. They had felt that an app could be a suitable tool to use for treatment of incontinence:

> *And then I thought, maybe there's an app, there must be an app that helps me to do my pelvic floor exercises, to remind me to do it, because it was mostly the reminder I needed...*

The participants perceived the app as a new and modern way of helping them with their incontinence and described their excitement and curiosity about trying it once they had found it. They appreciated that there was an evolution of treatment methods for this condition and that someone was doing research on it. It made them feel that incontinence was an important and prioritized problem:

> *Yes, great, solve it in a good, I mean use technology in a good way to solve a typically feminine problem, which otherwise is just a scrappy piece of paper from the healthcare centre, you know.*

XSL•FO
RenderX

Textbox 1. Categories and subcategories under the core category "Enabling my Independence."

Something new!

- New technology, something new and modern
- Low threshold, easily accessible

Keeping motivation up!

- Motivational support
- Visualizing treatment

Good enough?

- Unsure expectations of the result
- Insecurity about training correctly

This feeling was in contrast with their previous experiences with health care in which they felt that their incontinence had been dismissed as a banal problem. They described treatment programs that they had received as old, black-and-white photocopies that were uninspiring and easily forgotten. One woman felt that the advice she received was ignorant and old fashioned:

> *And then to get a poor photocopy, you know, one of those kind of cut-away views of the genital area. And then just, yes well, it's also important for you to exercise. And then, well I've tried to exercise and it didn't go well, and then the next person I meet says, like, yes, well, you just have to get used to this being the way it is. Kind of despondent.*

It seemed like a treatment program delivered via an app was appealing, as much for its form—a new, modern, and "fun" tool—as for anything else. Some participants also felt that it was easier to talk to friends about a new app for PFMT than to talk about incontinence.

Low Threshold, Easily Accessible

One main advantage with following a treatment program in an app seemed to be the proximity to the app and the mobile phone, making the treatment program easily available and convenient to use:

> *As I use my phone every day and it feels so accessible, well I thought, of course, I always have it with me, so it's easy to find a moment to do the exercises.*

The participants appreciated that this treatment was easily accessible from home and was less time consuming than booking appointments in conventional health care:

> *That was what was so good about this, I can do this at home myself, no need to book an appointment, find the time and suit others, and you know, that process of booking a time.*

The participants saw the app treatment as a "first-line" treatment, something to try first, before eventually looking for other help, if necessary. However, the threshold for seeking help in ordinary health care seemed high, and for most women, it was not an alternative at the time.

Keeping Motivation Up!

Motivational Support

Our participants described many barriers to achieving regular PFMT, the most common of which was problems prioritizing training, not being motivated enough, or forgetting to practice:

> *Well, to get it sorted, get a regular reminder, get the support day-in, day-out. Because there's so much other stuff that, well you know, life just goes on as it does, and then you end up never doing anything about it because it feels a bit uncomfortable.*

The possibility of receiving reminders was the main reason participants had chosen to try an app, and they were convinced that reminders were what they needed to incorporate the routine of training into their daily lives. Most women used the reminder function and described how they tried to perform the exercises when planned, although they sometimes had to postpone the reminders because they were busy at work:

> *Of course, I've read all the brochures and I've searched on the internet for pelvic-floor exercises, but you just forget all about it, so what I wanted help with was exactly this, to remember it more.*

Participants not using the reminders stated that just seeing the app on the mobile screen reminded them to do the exercises, and one woman even described it as having a person next to her, reminding her to perform the exercises:

> *It's works pretty much like when somebody tells you, you know, that you have to exercise.*

Some of the participants did, however, describe difficulties in maintaining the same enthusiasm and motivation in the long run. They reported that when the novelty of the app wore off, the training program felt boring. In addition, they reported that in the beginning, they were eager to follow each reminder directly, but at the end of the 3 months, they sometimes ignored the reminders or became annoyed with them. One woman described it somewhat dramatically:

> *So, during the first month I thought it was kind of fun, because it was new. In the second month I thought like this; oh no...that phone is ringing again, time for me to do my exercises, God how boring. And then*

XSL•FO
RenderX

> *during the last month I really had to, pretty much, force myself to do it. But I thought it was so boring.*

Another woman said that when her incontinence improved, she lost some of her motivation to do the exercises. This loss led to an increase in her symptoms, and she planned to make a new effort with the training:

> *My biggest problem was that I got pretty good results, so I stopped doing them...So now the problems have come back again. But I just need to start over.*

Visualizing Treatment

The participants described the app as easy to use and having few technical problems. The training program gave them a fixed structure to follow, and they could individualize how fast they wanted to proceed in the program. They particularly appreciated the way the app visualized the training by offering a moving graph for each contraction. This visual helped them see for how long and at what intensity they should contract and relax the muscles:

> *Being able to see, partly to read the short intro and also see the picture properly, well yes, it just gives a completely different dimension in solving the problem.*

Although the participants could memorize some of the exercises, after a while, they mostly chose to look at the app when performing them and reported that it helped them stay focused on the exercise.

Good Enough?

Unsure Expectations of the Result

Despite high expectations from an app that supports training, the participants were unsure of their expectations about the actual treatment effect. They were "giving it a chance" or at least getting a little bit better with the app:

> *I probably expected that after three months I would be able to feel like I had done everything I could. I did not expect to feel like, oh, now I'm going to be cured for the rest of my life, but to know that I had at least given it an honest chance.*

Some argued that it was difficult to know whether the result was good enough or if they could get even better if they exercised more:

> *It has become much better. And that makes me wonder about something, how good can it get actually?" Can I actually become, like, really, really good?*

Some raised questions about how to evaluate or measure muscle strength:

> *And how good are my muscles, how well have I trained them? You never know that, I don't think you ever get the answer sheet on that. Compared with others, or compared with people that don't have leakage. Are their muscles stronger than mine, and things like that?*

Some of the participants who did not get better after 3 months expressed that they would have liked to talk to someone about their results or lack thereof. They raised questions about what results could be expected, if they could get better by training for another 3 months, and where they should go if they wanted further help. At the same time, they felt satisfied that they had done what they could and felt that it helped them in their future decisions about whether to seek further help.

Insecurity About Training Correctly

After the treatment period, most participants felt confident that they could perform pelvic floor muscle contractions, had better control over their muscles, and had improved their muscle strength. At the same time, they were unsure about whether they could have been even stronger if they had contracted their muscles "better" or in a more "correct" way:

> *But I feel like I have found my muscles, but that I'm still a little bit uncertain about it, kind of...But you never know of course, would I have been even better if I had found my muscle groups better.*

Although some participants would have liked someone to confirm that they were contracting correctly, most felt that the instructions within the app and their own experience performing PFMT were sufficient for them to feel confident that they did it right.

> *Well, I seem to remember that there was some text that said that in the beginning, you could feel for yourself to see if you were contracting in the right way, and I think that was good enough.*

Participants who felt that their incontinence symptoms decreased considered this decrease a confirmation of the fact that they had performed the training correctly.

> *No, it's more that I've tested it out and feel that I'm becoming successively stronger and stronger, so I have gotten that confirmation, that this was something important. Of course, the support in the instructions, what I would experience, that I would feel it dropping, feel it relaxing, and that's exactly what I felt.*

Discussion

Principal Findings

This qualitative study explored women's experiences using an app-based treatment for SUI. We found that their experiences could be grouped into three categories. When deciding to try PFMT to treat their incontinence, the participants wanted some kind of support, and they searched for something new and something better than what they had previously tried, which is described in the category "Something new!" They also needed some help with motivation and coaching to be able to carry through the treatment they had engaged in, which is described in the category "Keeping motivation up!" In addition, they had to deal with the fact that performing a treatment independently involves a certain amount of uncertainty, which is described in the category "Good enough?" The core category "Enabling my independence" was related to all other categories and illustrates the women's wish to take responsibility for their health and treatment.

XSL•FO
RenderX

Core Category: Enabling My Independence

Our participants wanted to improve their health by self-treatment and thought that an app could be helpful. They had positive attitudes toward apps and were motivated to engage with them. This finding is in line with conclusions from a review by O'Connor et al, which showed that people who want to be healthier or have more control over their well-being are more likely to engage with digital health interventions [21].

Our core category summarizes the participants' wish to manage their incontinence and PFMT independently and the ways the app could support them in achieving this. Similar findings have been described in a qualitative study on a mobile app to increase physical activity, where the participants appreciated the opportunity to use the app independently and described a feeling of being in control [22]. In addition, in a 2017 review of perceptions of health-tracking apps for chronic illnesses, the authors found that apps were an important support for managing the chronic illness, and users felt empowered to independently manage their illness [14].

Providing Treatment Using an App

Our participants compared the app-based treatment to other self-management programs for SUI that they had tried before. They perceived the app as a more modern and effective method. It also seemed that the chance to use new technology made them feel like incontinence was prioritized. This finding is interesting and might be specific to an app dealing with a health condition considered to be "embarrassing" or "taboo." As in our previous study on internet-based treatment, some women reported that participating in the study made it easier to bring up the otherwise sensitive topic of incontinence with friends and family [23]. Moreover, in this case, it is possible that the use of a new, modern, and "fun" tool like an internet program or a training app facilitated a willingness to "open up" the subject.

Our participants emphasized the convenience of using apps, which has been well described in previous studies [24,25]. Anderson et al found that convenience was the main reason people engaged with health apps [24]. The easy availability makes interaction, recording, and reminders easy. Our participants also found it less time consuming and embarrassing to use an app for advice about incontinence compared to booking appointments in conventional health care.

The participants described many barriers to performing regular PFMT, and by using an app, they hoped to overcome some of these barriers. The app was seen as a tool to put good intentions into practice. The app supported them in several ways—by its mere presence on their mobile phone, by providing a fixed structure for their training, and by reminding them to adhere to that structure.

Forgetting to practice is a common problem in the self-management of SUI, and reminders could be important for adherence [26]. A study of apps for nutrition behavior change for weight management found mixed perspectives on reminders [27]. Some participants found them useful, but others found them annoying, for example, if they arrived at inappropriate times or were not discreet enough. Our participants expressed little concern about this aspect, and one reason could be that they could individualize the reminders within the app to times that were convenient. The ability to personalize or customize app features is known to motivate app use [12]. Additionally, the reminders in our study could be described as "cues to action," whereas in the study on apps for weight management, the reminders appeared after a person had forgotten to record something in the app.

The app helped visualize or "to see" the training, with the moving graphs showing each contraction. The participants continued to use this visual help even when they had memorized the exercise, because they found it helpful for staying focused. In a study on mobile apps for young people with type 1 diabetes [28], an important finding was that the app provided a new visual understanding of diabetes self-management. All the adolescent participants in that study [28] used the verb "to see," and interestingly, our participants also commonly used this exact expression. Thus, it seems that visual support/visualization could be helpful for different populations and different diseases and is therefore likely to be of general importance.

Some participants reported that their motivation declined during the 3-month study period. One reason was getting bored when the novelty of the app had faded. We also found that some participants lost their motivation to continue when their incontinence improved, whereas some lost motivation for the opposite reason—lack of improvement. Both scenarios have been previously described with apps for other health conditions such as pain management and nutrition behavior change for weight management [24,27].

Performing Pelvic Floor Muscle Training

An interesting finding is that most participants found PFMT difficult and challenging, but at the same time, they felt confident that they could master it well on their own. Some did not feel completely sure that they were doing the pelvic floor muscle contractions correctly, even though they could feel that their pelvic floor muscle strength improved and their incontinence symptoms reduced.

Difficulties in understanding how to do the exercises and knowledge of whether they are done correctly are known barriers to adherence to PFMT. Enablers related to these difficulties are feedback, affirmation of progress, and professional involvement [29]. In the context of self-management with the help of an app, there are, of course, limitations to what kind of feedback is possible. Nevertheless, most women can correctly perform pelvic floor muscle contractions after a simple verbal cue [30]. There was information within the app explaining how to find the "right" pelvic floor muscles and how to evaluate progress. Most participants found that the information within the app and their own experience made them confident that their pelvic floor muscle contractions were "good enough." The additional feeling of uncertainty was possibly unrelated to a lack of knowledge or skills for performing PFMT but rather related to a lack of reassurance, which might be an inevitable aspect of self-management treatment. Participants who did not experience improvement in symptoms where naturally more likely to wish for personalized feedback.

XSL•FO
RenderX

One explanation for why some participants had difficulties evaluating their results could be that they had unclear expectations about what results to expect with their incontinence symptoms when entering the study. Goal setting plays a role in promoting adherence to PFMT, as described in a report on adherence strategies [31]. Similarly, in a study of experiences with behavioral interventions for urinary incontinence, such as PFMT, French et al found that factors that enabled clients to adhere to PFMT were having realistic goals and expectations [29]. Our app did not include an individualized goal-setting feature. However, we encouraged participants to follow the treatment program and practice PFMT three times daily for 3 months, which could be considered a training goal.

Further Research and Implications for Practice

Further research on app features that can increase adherence to app-based treatments would be valuable, such as ways to encourage goal setting for people using apps for self-management. In addition, ways to increase motivation to use apps in the long-term would be of interest. The impact of more advanced features within apps, like more individualized information and feedback, competition, or social sharing, could be evaluated, bearing in mind that more complex apps might lose some of their "ease of use."

The app Tät was designed for sustained use among women with SUI. The idea of using the app for more periodic treatment when the symptoms (and the motivation to reduce symptoms) increase is interesting. Periodic use is probably common with self-management tools like apps. Although evaluated as a stand-alone treatment program, in clinical practice, the app could be used as a complement to other treatments such as supervised PFMT within conventional health care.

The Tät app was designed as a first-line, conservative treatment program for SUI with a focus on PFMT. We have evaluated the app with its different components as a whole, but we could not determine which parts of the program have been effective. This qualitative study has provided more insight into which parts of the app treatment women appreciated and which parts they found difficult. This information has already led us to make improvements in the app, such as adding audio sounds to the exercises to simplify the training and including a calendar function to facilitate self-monitoring of the number of exercises performed. After the RCT was completed, the app was released for free in the App Store and Google Play in Swedish, English, German, Spanish, Arabic, and Finnish. To date, 54,000 people have downloaded the app and answered our short questionnaire. The app is CE (Conformité Européenne) marked as a Class 1 medical device in accordance with Swedish regulation LVFS 2003:11 [32]. We are following the use of the app to learn more about who uses it and how well it works outside a study setting.

Strengths and Limitations

This is the first qualitative study of an app-based treatment for SUI. We considered this interview study an important complement to our RCT, because we wanted to understand how and why an app treatment could be effective for SUI. Grounded theory was a suitable method of analysis, because little previous knowledge is available about this subject and we could not know in advance what findings to expect.

One strength was that our participants used the app for a long time (3 months). This duration made it possible to explore attitudes towards app interventions, factors that affected motivation to use the app, and challenges with long-term adherence.

The same interviewer performed all the interviews and did a majority of the transcription, thereby obtaining a thorough understanding of the material. Telephone interviews have some obvious limitations such as the lack of body language and the challenge of achieving personal contact without seeing the other person. This difficulty might be one reason for the relatively short interviews. On the other hand, the method made it possible to interview participants from all over Sweden.

The interviewer and all coauthors are family physicians, which could be an advantage, since we are used to talking about urinary incontinence with our patients. On the other hand, we might have preconceptions about what constitutes the best treatment method for women with incontinence.

We purposefully selected participants to represent a broad range of experiences, with variations in age and residence. The educational level was overall high in our RCT, limiting transferability to other groups of women with other educational backgrounds. However, educational level has not been found to affect the ability to learn or perform correct PFMT contractions [30]. The high educational level might also reflect the fact that use of health apps is still more common among individuals with higher education [33].

Conclusions

Women using an app-based treatment program for SUI reported that the app treatment enabled them to self-manage an up-to-date treatment for their condition. They perceived the app as a new and modern tool that is useful for maintaining motivation when performing a PFMT program. Many women had doubts about whether their training results were good enough but accepted that some level of uncertainty is inevitable with this kind of independent self-treatment. The app is an appreciated way of providing first-line treatment to women with SUI.

Acknowledgments

We thank all the participants for making this study possible. We also gratefully acknowledge and thank our study coordinator Susanne Johansson. The study was supported by grants from the Region Jämtland Härjedalen (DNR: JLL-843181); the Swedish Research Council for Health, Working Life and Welfare (Ref no 2014-5140); the Kamprad Family Foundation for Entrepreneurship, Research & Charity (Ref no 20170202); and Visare Norr, Northern County Councils, Sweden (DNR : VISARENORR656551).

XSL•FO
RenderX

The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

The Tät app was developed by ES, MS, and GU in cooperation with engineers at ITS (ICT Services and System Development), Umeå University, Sweden. The name Tät (mobile app) and the logo Tät.nu are registered as trademarks by The Swedish Patent and Registration office for ES at Umeå University. There is a copyright for Tät.nu (eContinence.se) at Umeå University. None of the researchers have any financial interest in the product.

Conflicts of Interest

None declared.

References

1. Haylen BT, de Ridder D, Freeman RM, Swift SE, Berghmans B, Lee J, International Urogynecological Association, International Continence Society. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Neurourol Urodyn 2010;29(1):4-20. [doi: 10.1002/nau.20798] [Medline: 19941278]
2. Abrams P, Cardozo L, Khoury S, Wein A, editors. Epidemiology of Urinary Incontinence (UI) other Lower Urinary Tract Symptoms (LUTS), Pelvic Organ Prolapse (POP), Anal Incontinence (AI). In: Incontinence: Fifth Edition. Arnhem, The Netherlands: EAU; 2013.
3. Ebbesen MH, Hunskaar S, Rortveit G, Hannestad YS. Prevalence, incidence and remission of urinary incontinence in women: longitudinal data from the Norwegian HUNT study (EPINCONT). BMC Urol 2013 May 30;13:27 [FREE Full text] [doi: 10.1186/1471-2490-13-27] [Medline: 23721491]
4. Siddiqui NY, Levin PJ, Phadtare A, Pietrobon R, Ammarell N. Perceptions about female urinary incontinence: a systematic review. Int Urogynecol J 2014 Jul;25(7):863-871. [doi: 10.1007/s00192-013-2276-7] [Medline: 24310988]
5. Lukacz ES, Santiago-Lastra Y, Albo ME, Brubaker L. Urinary Incontinence in Women: A Review. JAMA 2017 Oct 24;318(16):1592-1604. [doi: 10.1001/jama.2017.12137] [Medline: 29067433]
6. Dumoulin C, Cacciari LP, Hay-Smith EJC. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. Cochrane Database Syst Rev 2018 Dec 04;10:CD005654. [doi: 10.1002/14651858.CD005654.pub4] [Medline: 30288727]
7. Dumoulin C, Hunter KF, Moore K, Bradley CS, Burgio KL, Hagen S, et al. Conservative management for female urinary incontinence and pelvic organ prolapse review 2013: Summary of the 5th International Consultation on Incontinence. Neurourol Urodyn 2016 Jan;35(1):15-20. [doi: 10.1002/nau.22677] [Medline: 25400065]
8. Dumoulin C, Hay-Smith J, Frawley H, McClurg D, Alewijnse D, Bo K, International Continence Society. 2014 consensus statement on improving pelvic floor muscle training adherence: International Continence Society 2011 State-of-the-Science Seminar. Neurourol Urodyn 2015 Sep;34(7):600-605. [doi: 10.1002/nau.22796] [Medline: 25998603]
9. Statista. Smartphone user penetration as percentage of total global population from 2014 to 2021 URL: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/ [accessed 2018-06-11] [WebCite Cache ID 70641wggx]
10. Svenskarna och internet 2017. URL: https://www.iis.se/docs/Svenskarna_och_internet_2017.pdf [accessed 2018-06-12] [WebCite Cache ID 7079eWo0C]
11. Pew Research Center. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-06-12] [WebCite Cache ID 707B6Ysy8]
12. Research 2 Guidance. 84,000 health app publishers in 2017 – Newcomers differ in their go-to-market approach URL: https://research2guidance.com/84000-health-app-publishers-in-2017/ [accessed 2018-06-12] [WebCite Cache ID 707C66AR2]
13. 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;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
14. Birkhoff SD, Smeltzer SC. Perceptions of Smartphone User-Centered Mobile Health Tracking Apps Across Various Chronic Illness Populations: An Integrative Review. J Nurs Scholarsh 2017 Jul;49(4):371-378. [doi: 10.1111/jnu.12298] [Medline: 28605151]
15. Tät.nu. Tät.nu – research into the treatment of urinary incontinence URL: http://xn--tt-via.nu/?page_id=84&lang=en [accessed 2019-01-25]
16. Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Mobile app for treatment of stress urinary incontinence: A randomized controlled trial. Neurourol Urodyn 2017 Jun;36(5):1369-1376. [doi: 10.1002/nau.23116] [Medline: 27611958]
17. Sjöström M, Lindholm L, Samuelsson E. Mobile App for Treatment of Stress Urinary Incontinence: A Cost-Effectiveness Analysis. J Med Internet Res 2017 May 08;19(5):e154 [FREE Full text] [doi: 10.2196/jmir.7383] [Medline: 28483745]

XSL•FO
RenderX

18. Hoffman V, Söderström L, Samuelsson E. Self-management of stress urinary incontinence via a mobile app: two-year follow-up of a randomized controlled trial. Acta Obstet Gynecol Scand 2017 Oct;96(10):1180-1187 [FREE Full text] [doi: 10.1111/aogs.13192] [Medline: 28718223]
19. Hamberg K, Johansson E, Lindgren G, Westman G. Scientific rigour in qualitative research--examples from a study of women's health in family practice. Fam Pract 1994 Jun;11(2):176-181. [Medline: 7958583]
20. Corbin J, Strauss A. Basics of qualitative research. 3rd edition. Techniques and procedures for developing grounded theory. Thousand Oaks: Sage Publications, Inc; 2008.
21. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
22. Casey M, Hayes PS, Glynn F, OLaighin G, Heaney D, Murphy AW, et al. Patients' experiences of using a smartphone application to increase physical activity: the SMART MOVE qualitative study in primary care. Br J Gen Pract 2014 Aug;64(625):e500-e508 [FREE Full text] [doi: 10.3399/bjgp14X680989] [Medline: 25071063]
23. Björk A, Sjöström M, Johansson EE, Samuelsson E, Umefjord G. Women's experiences of internet-based or postal treatment for stress urinary incontinence. Qual Health Res 2014 Apr;24(4):484-493. [doi: 10.1177/1049732314524486] [Medline: 24598777]
24. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
25. Tatara N, Arsand E, Skrøvseth SO, Hartvigsen G. Long-term engagement with a mobile self-management system for people with type 2 diabetes. JMIR Mhealth Uhealth 2013;1(1):e1 [FREE Full text] [doi: 10.2196/mhealth.2432] [Medline: 25100649]
26. Borello-France D, Burgio KL, Goode PS, Ye W, Weidner AC, Lukacz ES, Pelvic Floor Disorders Network. Adherence to behavioral interventions for stress incontinence: rates, barriers, and predictors. Phys Ther 2013 Jun;93(6):757-773 [FREE Full text] [doi: 10.2522/ptj.20120072] [Medline: 23431210]
27. Lieffers JRL, Arocha JF, Grindrod K, Hanning RM. Experiences and Perceptions of Adults Accessing Publicly Available Nutrition Behavior-Change Mobile Apps for Weight Management. J Acad Nutr Diet 2018 Feb;118(2):229-239.e3. [doi: 10.1016/j.jand.2017.04.015] [Medline: 28625662]
28. Frøisland DH, Arsand E, Skårderud F. Improving diabetes care for young people with type 1 diabetes through visual learning on mobile phones: mixed-methods study. J Med Internet Res 2012;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
29. French B, Thomas L, Harrison J, Coupe J, Roe B, Booth J, ICONS Project Team and the ICONS Patient, Public and Carer Groups. Client and clinical staff perceptions of barriers to and enablers of the uptake and delivery of behavioural interventions for urinary incontinence: qualitative evidence synthesis. J Adv Nurs 2017 Jan;73(1):21-38. [doi: 10.1111/jan.13083] [Medline: 27459911]
30. Henderson JW, Wang S, Egger MJ, Masters M, Nygaard I. Can women correctly contract their pelvic floor muscles without formal instruction? Female Pelvic Med Reconstr Surg 2013;19(1):8-12 [FREE Full text] [doi: 10.1097/SPV.0b013e31827ab9d0] [Medline: 23321652]
31. Dumoulin C, Alewijnse D, Bo K, Hagen S, Stark D, Van Kampen M, et al. Pelvic-Floor-Muscle Training Adherence: Tools, Measurements and Strategies-2011 ICS State-of-the-Science Seminar Research Paper II of IV. Neurourol Urodyn 2015 Sep;34(7):615-621. [doi: 10.1002/nau.22794] [Medline: 25998493]
32. Läkemedelsverket. Legislation - Medical Devices URL: https://lakemedelsverket.se/english/product/Medical-devices/An-Introduction-to-Rules-and-Regulations/Legislation/ [accessed 2019-01-25]
33. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]

Abbreviations

CE: Conformité Européenne
PFMT: pelvic floor muscle training
SUI: stress urinary incontinence
UI: urinary incontinence

XSL•FO
RenderX

Edited by G Eysenbach; submitted 14.06.18; peer-reviewed by K Grindrod, S Bahanshal; comments to author 04.10.18; revised version received 23.11.18; accepted 12.12.18; published 14.03.19.

Please cite as:
Asklund I, Samuelsson E, Hamberg K, Umefjord G, Sjöström M
User Experience of an App-Based Treatment for Stress Urinary Incontinence: Qualitative Interview Study
J Med Internet Res 2019;21(3):e11296
URL: http://www.jmir.org/2019/3/e11296/
doi:10.2196/11296
PMID:30869644

©Ina Asklund, Eva Samuelsson, Katarina Hamberg, Göran Umefjord, Malin Sjöström. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Association Between Medication Adherence for Chronic Conditions and Digital Health Activity Tracking: Retrospective Analysis

Tom Quisel[1], BS; Luca Foschini[1], PhD; Susan M Zbikowski[2,3], PhD; Jessie L Juusola[1], PhD

[1]Evidation Health, San Mateo, CA, United States
[2]inZights Consulting, LLC, Seattle, WA, United States
[3]Humana, Inc, Louisville, KY, United States

Corresponding Author:
Jessie L Juusola, PhD
Evidation Health
167 2nd Avenue
San Mateo, CA,
United States
Phone: 1 650 279 8855
Email: jjuusola@evidation.com

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2019/12/e17375/

Abstract

Background: Chronic diseases have a widespread impact on health outcomes and costs in the United States. Heart disease and diabetes are among the biggest cost burdens on the health care system. Adherence to medication is associated with better health outcomes and lower total health care costs for individuals with these conditions, but the relationship between medication adherence and health activity behavior has not been explored extensively.

Objective: The aim of this study was to examine the relationship between medication adherence and health behaviors among a large population of insured individuals with hypertension, diabetes, and dyslipidemia.

Methods: We conducted a retrospective analysis of health status, behaviors, and medication adherence from medical and pharmacy claims and health behavior data. Adherence was measured in terms of proportion of days covered (PDC), calculated from pharmacy claims using both a fixed and variable denominator methodology. Individuals were considered adherent if their PDC was at least 0.80. We used step counts, sleep, weight, and food log data that were transmitted through devices that individuals linked. We computed metrics on the frequency of tracking and the extent to which individuals engaged in each tracking activity. Finally, we used logistic regression to model the relationship between adherent status and the activity-tracking metrics, including age and sex as fixed effects.

Results: We identified 117,765 cases with diabetes, 317,340 with dyslipidemia, and 673,428 with hypertension between January 1, 2015 and June 1, 2016 in available data sources. Average fixed and variable PDC for all individuals ranged from 0.673 to 0.917 for diabetes, 0.756 to 0.921 for dyslipidemia, and 0.756 to 0.929 for hypertension. A subgroup of 8553 cases also had health behavior data (eg, activity-tracker data). On the basis of these data, individuals who tracked steps, sleep, weight, or diet were significantly more likely to be adherent to medication than those who did not track any activities in both the fixed methodology (odds ratio, OR 1.33, 95% CI 1.29-1.36) and variable methodology (OR 1.37, 95% CI 1.32-1.43), with age and sex as fixed effects. Furthermore, there was a positive association between frequency of activity tracking and medication adherence. In the logistic regression model, increasing the adjusted tracking ratio by 0.5 increased the fixed adherent status OR by a factor of 1.11 (95% CI 1.06-1.16). Finally, we found a positive association between number of steps and adherent status when controlling for age and sex.

Conclusions: Adopters of digital health activity trackers tend to be more adherent to hypertension, diabetes, and dyslipidemia medications, and adherence increases with tracking frequency. This suggests that there may be value in examining new ways to

XSL•FO
RenderX

further promote medication adherence through programs that incentivize health tracking and leveraging insights derived from connected devices to improve health outcomes.

(J Med Internet Res 2019;21(3):e11486) doi:10.2196/11486

KEYWORDS

activity tracker; health behavior; eHealth; mHealth; medication adherence

Introduction

Background

Chronic diseases affect approximately half of all adults in the United States, and they are the leading cause of death and disability [1]. They also create a substantial cost burden; patients with chronic diseases accounted for 86% of all US health care spending in 2010 [2]. Of the biggest contributors, 2 are heart disease and stroke, estimated to cost US $315 billion in 2010, and diabetes, estimated to cost US $245 billion in 2012 [3,4]. Prescription medication is a key component of treatment for these diseases and their underlying risk factors, but adherence to medication has historically been low for patients with chronic diseases [5]. This is problematic as poor medication adherence can lead to poor health outcomes, which then increase health care utilization and costs [6].

The benefits of high medication adherence have been well established in diabetes, hypertension, and dyslipidemia, all of which are major risk factors for heart disease and stroke. One meta-analysis of studies in various disease areas, including diabetes and heart disease prevention, found that good medication adherence was associated with lower mortality, as compared with poor adherence [7]. Medical and pharmacy claims analyses have examined the relationship between medication adherence and health care utilization and costs. In 1 analysis across diabetes, dyslipidemia, and hypertension, the authors found that the annual total health care spending was significantly lower for adherent patients than nonadherent patients despite higher pharmacy costs in adherent patients [8]. The overall decrease in costs was driven by fewer hospitalizations and emergency department visits in the adherent population. Analyses of other claims data sources have led to similar conclusions [9-11].

Objectives

Given the relationship between medication adherence and health care utilization and costs, efforts to increase medication adherence have been well studied. Patient behavior is often a key factor in individual medication adherence patterns, but research on the relationship between adherence and health activity behavior is limited [12]. Better understanding the link between behavior and medication adherence could facilitate the development of programs, tools, and approaches that improve adherence and thus lead to lower disease burden. With the recent proliferation of digital health trackers for activity, sleep, and diet, new data are available on these types of behaviors, which provide new opportunities to examine how health behaviors and lifestyles are linked to medication adherence. In 2013, an estimated 2% of the US population had used a wearable device. The use is growing quickly, with some estimates suggesting that over 20% of the population owned a wearable device in 2016 and annual sales projecting to increase to more than US $50 billion by 2018 [13,14]. We leveraged medical and pharmacy claims and other health behavior data from insured individuals to examine the relationship between health behavior and medication adherence. We sought to understand the connection between digital activity-tracking behavior and adherence for people with diabetes, dyslipidemia, and hypertension, and we sought to understand whether engaging with digital health trackers ties to changes in medication adherence. This can provide insight on the value of using data from connected devices to understand health behaviors such as medication adherence and improve health outcomes through health engagement strategies.

Methods

Study Sample

Claims Data

The analytic sample was derived from Humana medical and pharmacy claims and other health behavior data for an insured population. All data were deidentified and complied with requirements set forth by Humana's Protected Health Information and Vendor Ethics committee. The study received institutional review board (IRB) exemption from Solutions IRB.

We identified cases with continuous health insurance coverage from January 1, 2015 to June 1, 2016 with at least 1 of 3 medical conditions: diabetes, dyslipidemia, or hypertension. Health conditions were established based on the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) and ICD, Tenth Revision, CM (ICD-10-CM) codes, which are official systems of assigning codes to diagnoses and procedures associated with hospital utilization in the United States. Individuals were included if they had at least 1 outpatient visit or hospitalization with a specified ICD-9-CM or ICD-10-CM code, and they were included if there was a relevant diagnosis between January 1, 2005 to September 1, 2015 (Table 1). People with multiple conditions were included in each disease cohort.

The disease cohorts were further limited on the basis of pharmacy claims. We defined 2 periods for examining medication utilization: (1) the supply period and 2) the analysis period. The supply period (January 1, 2015-May 31, 2015) was the baseline time frame during which we utilized pharmacy claims to identify individuals who should be included in the analysis on the basis of prescription fills and to estimate the supply of medication in each patient's possession at the start of the analysis period. Understanding the medication supply at the start of the analysis period is necessary for correctly computing

XSL•FO
RenderX

medication adherence during that period. The analysis period (June 1, 2015-June 1, 2016) was the time frame in which we tracked prescription refills and adherence rates as well as activity-tracker use. To be included in the analysis, individuals must have had at least 1 pharmacy claim for an oral medication relevant to the disease cohort during the supply or analysis period (Table 1).

Table 1. Medical diagnosis codes used in condition cohort creation and classes of medication included in medication adherence analysis.

Condition	ICD-9-CM[a] Codes	ICD-10-CM[b] Codes	Pharmaceutical treatment classes
Diabetes	250.x	E08.x-E11.x	Alpha-glucosidase inhibitors; dipeptidyl peptidase - 4 inhibitors; glucagon-like peptide-1 receptor agonists; meglitinides; metformin/metformin combinations; sodium glucose cotransporter - 2 inhibitors; sulfonylureas; thiazolidinediones
Dyslipidemia	272.0, 272.2. 272.4	E78.0, E78.2, E78.4, and E78.5	Bile acid sequestrants; cholesterol absorption inhibitors; fibrates; lipid-regulating agents; nicotinic acid derivatives; statins
Hypertension	401.x-405.x	I10.x-I15.x	angiotensin-converting–enzyme inhibitors; Angiotensin II receptor blockers; beta blockers; calcium channel blockers; clonidine; diuretics; hydralazine; renin inhibitors

[a]ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

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

Activity-Tracking Data

Additional health behavior data were available for a subsample of participants who chose to use and link their activity-tracking device to track and earn points and status awards (eg, gold and silver) for participating in health and wellness activities. These points could be redeemed via a Web-based store for activity trackers, other fitness-related items (apparel, gear), and gift cards. We extracted activity-tracking data—including step counts, sleep duration, weight, and food logs. To focus on individuals who make a daily active choice of whether to track, we limited step counts in this analysis to those from tracker devices rather than step counts collected passively via smartphone.

Study Variables and Analysis

The primary analysis examined the association between medication adherence, as measured by proportion of days covered (PDC), and activity-tracking metrics.

Medication Adherence

PDC is a common metric in retrospective medication adherence research, using pharmacy claims data to calculate the proportion of time that an individual had a prescribed medication on hand for a given condition [15]. We computed PDC using both a fixed and a variable denominator, as both methodologies are commonly used [8-11,16,17]. Fixed PDC generally serves as a lower bound for actual adherence, whereas variable PDC generally serves as an upper bound [18].

The fixed PDC methodology assumes that an individual should be taking medication during the entire analysis period [8,11]. Thus, we calculated the fixed PDC over a denominator of 1 year. The numerator is the number of days on which the individual had the medication on hand, including medication remaining from the supply period. In contrast, the variable PDC methodology assumes that the individual should be taking the medication only during the variable analysis period, defined as the time period between the first prescription fill and the end of the supply of the last refill [9]. Its numerator is the number of days on which the individual had the medication on hand during this period. Individuals must have had at least 2 prescription fills during the analysis period to prevent trivial variable PDC values of 1.0 from individuals who filled just 1 prescription during the analysis period [9]. For both the fixed and variable PDC methodology, PDC values ranged from 0.0 to 1.0.

We calculated a drug-class-level PDC over the analysis period for each therapeutic class of drugs used to treat the condition for which the individual filled prescriptions. We then calculated the condition-level PDC for each individual as an average of the per-individual drug-class-level PDCs for the fixed PDC methodology and as a weighted average of the per-individual drug-class-level PDCs for the variable PDC methodology. We used the length of the variable analysis period as the weight for each drug class' variable PDC.

Finally, we used a threshold of condition-level PDC≥0.80 to define each individual as adherent or nonadherent for a particular condition. This is the threshold most commonly used in medication adherence research [5,6,8,9,19]. In the sensitivity analysis, we also considered a scenario where any individual with at least 1 drug-class-level PDC≥0.80 was classified as adherent.

Activity Tracking

We computed several activity metrics for each participant. First, we created a binary variable indicating if the individual had ever used an activity tracker during the analysis period (June 1, 2015-June 1, 2016). To understand how consistently individuals tracked their activities, we computed a tracking ratio for each individual: the ratio of days on which the individual logged at least 1 activity to the total number of days between the first activity logged and the most recent activity logged in the analysis period.

To correct for highly variable tracking ratios for individuals with little data, we created an adjusted tracking ratio. For each individual m we modelled the observed tracking ratio $r = k / n$ (where k is the number of days tracked and n is the number of

XSL•FO
RenderX

total days between *m*'s first and last day tracked) as a sample from a Binomial(*n*, *p*). To create a shrinkage estimator of *m*'s true tracking ratio using the Bayesian framework, we chose a Beta(α,β) distribution as our prior for *m*'s tracking ratio. We fixed α / (α+β) to *rmean*, the sample mean of tracking ratios across all individuals and fixed α+β-2 to 10. After performing a Bayesian update using *m*'s observed tracking ratio, Maximum A Posteriori estimation of *m*'s posterior tracking ratio was used to generate a point estimate of *m*'s true tracking ratio. With our choices of α and β, the point estimate reduced to (*nr* +10 *rmean*) / (*n* +10). We call this final result the adjusted tracking ratio.

Finally, to assess activity level, we computed steps taken per week during which a tracker was used. We divided the total number of steps taken by each individual in the measurement period by the number of weeks during which they tracked steps. We limited this metric to step count, which directly corresponds to the desired healthy behavior (ie, exercise). Sleep, weight, and food logging data were not used to assess activity level, as they do not have a direct linear tie to the desired healthy behavior (eg, sleep duration can be too short or too long and healthy weight is dependent on many other personal characteristics).

Statistical Analysis

We performed all analyses using both the fixed and variable condition-level PDC adherent or nonadherent status, in turn, as the outcome variable. We considered a result significant if it achieved a *P* value of .05 or less using a 2-sided *t* test, or in the case of logistic regression, a 2-sided Wald test.

Association Between Adherence and Activity-Tracker User (Trackers vs Nontrackers)

We used logistic regression to assess the relationship between tracker versus nontracker status and fixed and variable adherent versus nonadherent status, including an interaction term to control for age and sex. We first built an overall model to measure the effect across all conditions and activities. Individuals with multiple conditions were accounted for under each condition. We then built 3 separate models, 1 for each condition, and then 4 additional models, 1 for each activity (steps, sleep, weight, and food logs), for a total of 7 models to investigate any differences in the tracker-adherence relationship across conditions or activity types.

Association Between Adherence and Activity-Tracking Metrics

We then used logistic regression to assess the relationship between fixed and variable individual adherent versus nonadherent status, frequency of tracking activities (as measured by adjusted tracking ratio), and activity level (as measured by steps per week tracked), including an interaction term to control for age and sex. We performed this analysis across individuals from all conditions, limited to those with at least 2 tracking events at least 10 days apart in the analysis period. Tracker data are significantly autocorrelated over the span of a few days, so the 10-day separation ensures at least 2 independent measurements per individual. We built a model for each of the 2 activity metrics independently as well as a combined model including both.

We performed all analysis using Python version 2.7.10 (Python Software Foundation), Spark version 1.3.1 (The Apache Software Foundation), Pandas version 0.15.2 (Python Software Foundation), and numpy 1.9.2 (NumPy Developers). We used significance tests from the stats module in scipy version 0.15.1 (Scipy Developers). We performed our logistic regressions using R version 3.2.0 (The R Foundation).

Results

Sample Characteristics

We identified 117,765 individuals with diabetes, 317,340 individuals with dyslipidemia, and 673,428 individuals with hypertension who were included in the fixed PDC analysis. Slightly fewer individuals qualified for the variable PDC analysis, as it required at least 2 pharmacy claims for at least 1 relevant medication—102,322 for diabetes, 286,640 for dyslipidemia, and 642,818 for hypertension. There was an overlap among the disease cohorts, ranging from 11.89% (37,719/317,340) of dyslipidemia individuals also having a diabetes diagnosis to 86.66% (102,050/117,765) of individuals with diabetes also having a hypertension diagnosis (Table 2). Coronary artery disease was also a common comorbidity, seen in approximately one-thirds of each cohort. On average, the population screened for diabetes was aged 70.5 years and 51.49% (60,641/117,765) females, for dyslipidemia was aged 70.8 years and 53.66% (170,289/317,340) females, and for hypertension was aged 70.8 years and 56.01% (377,183/673,428) females (Table 2). Individuals with tracker activity were younger with an average age of 52.3 years for diabetes, 55.1 years for dyslipidemia, and 52.5 years for hypertension (Table 3).

Medication Adherence

Average PDC was similar across the conditions and was higher using the variable methodology than using the fixed methodology. Fixed methodology PDCs ranged from 0.673 for diabetes to 0.756 for dyslipidemia and hypertension, and variable methodology PDCs ranged from 0.917 for diabetes to 0.929 for hypertension (Table 4). The percent of individuals classified as adherent (PDC≥0.80) showed similar trends, ranging from 48.09% (56,630/117,765) for diabetes to 61.64% (195,606/317,340) for dyslipidemia for the fixed methodology and 85.38% (87,365/102,322) for diabetes to 89.07% (572,553/642,818) for hypertension for the variable methodology.

XSL•FO
RenderX

Table 2. Participant characteristics summarized by condition for all individuals.

Characteristic	Diabetes	Dyslipidemia	Hypertension
Individual population, n			
Fixed PDC[a] methodology	117,765	317,340	673,428
Variable PDC methodology	102,322	286,640	642,818
Age (years), mean (SD)	70.5 (10.7)	70.8 (10.8)	70.8 (11.7)
Sex, n (%)			
Female	60,641 (51.49)	170,289 (53.66)	377,183 (56.01)
Male	57,124 (48.51)	147,051 (46.34)	296,245 (43.99)
Comorbid conditions, n (%)			
Diabetes	117,765 (100)	37,719 (11.89)	102,050 (15.15)
Dyslipidemia	37,719 (32.03)	317,340 (100)	235,863 (35.02)
Hypertension	102,050 (86.66)	235,863 (74.33)	673,428 (100)
Coronary artery disease	40,931 (34.76)	96,169 (30.30)	254,212 (37.75)
Heart failure	16,745 (14.22)	35,936 (11.32)	81,362 (12.08)
End stage renal disease	2534 (2.15)	5055 (1.59)	10,041 (1.49)
Depression	12,085 (10.26)	29,655 (9.35)	78,035 (11.59)
Digital tracking activity, n (%)			
None	116,974 (99.33)	313,741 (98.87)	666,956 (99.04)
Steps	758 (0.64)	3498 (1.10)	6243 (0.93)
Sleep	530 (0.45)	2460 (0.78)	4414 (0.66)
Weight	179 (0.15)	736 (0.23)	1513 (0.22)
Food logs	112 (0.10)	463 (0.15)	1050 (0.16)

[a]PDC: proportion of days covered.

Table 3. Participant characteristics summarized by condition for individuals with tracker data.

Characteristic	Diabetes	Dyslipidemia	Hypertension
Individual population, n			
Fixed PDC[a] methodology	791	3599	6472
Variable PDC methodology	653	3196	6064
Age (years), mean (SD)	52.3 (10.2)	55.1 (9.6)	52.5 (10.5)
Sex, n (%)			
Female	444 (56.1)	1813 (50.48)	3784 (58.47)
Male	347 (43.8)	1786 (49.62)	2688 (41.53)

[a]PDC: proportion of days covered.

XSL•FO
RenderX

Table 4. Medication adherence by condition as measured by mean proportion of days covered and percent of individuals classified as adherent.

Metric	Diabetes	Dyslipidemia	Hypertension
PDC[a], mean (SD)			
Fixed PDC[a] methodology	0.673 (0.320)	0.756 (0.291)	0.756 (0.256)
Variable PDC methodology	0.917 (0.130)	0.921 (0.127)	0.929 (0.110)
Adherent (PDC≥0.80), n (%)			
Fixed PDC methodology	56,630 (48.09)	195,606 (61.64)	373,515 (55.46)
Variable PDC methodology	87,365 (85.38)	248,709 (86.77)	572,553 (89.07)

[a]PDC: proportion of days covered.

Table 5. Medication adherence metrics by age and sex. Results are significantly different between age groups and female versus male at $P<.001$ using a 2-sided *t* test.

Metric	Age		Sex	
	<50 years (n=43,545)	≥50 years (n=723,469)	Female (n=427,546)	Male (n=339,936)
PDC[a], mean (SD)				
Fixed PDC methodology	0.629 (0.321)	0.753 (0.271)	0.745 (0.275)	0.750 (0.275)
Variable PDC methodology	0.882 (0.154)	0.928 (0.115)	0.925 (0.118)	0.927 (0.116)
Adherent (PDC≥0.80), n (%)				
Fixed PDC methodology	17,785 (40.84)	414,009 (57.23)	238,638 (55.82)	194,503 (57.22)
Variable PDC methodology	34,010 (78.10)	640,439 (88.52)	375,430 (87.81)	300,410 (88.37)

[a]PDC: proportion of days covered.

Medication adherence was significantly associated with age and sex. On the basis of age alone, there was a clear separation at 50 years between those who were more adherent and those who were less adherent. Adherence increased with age, with 40.84% (17,785/43,545) of individuals younger than 50 years being adherent across conditions as compared with 57.23% (414,009/723,469) of those aged 50 years and older, using fixed methodology (Table 5). Variable methodology showed a similar pattern, with 78.10% (34,010/43,545) of those younger than 50 years being adherent across conditions as compared with 88.52% (640,439/723,469) of those aged 50 years and older. Males were slightly but still significantly more adherent than females across conditions, with 57.22% (194,503/339,936) versus 55.82% (238,638/427,546) and 88.37% (300,410/339,936) versus 87.81% (375,430/427,546) being adherent under fixed and variable methodologies, respectively.

Activity Tracking

We identified 8553 individuals who chose to link and share data from a digital health tracker for at least 1 activity across step counts, sleep, weight, and food logs, with step trackers being the most common. More than 75% of the individuals were using a Fitbit device to log steps and/or sleep, approximately 10% were logging steps via Garmin or Jawbone, and the remaining 15% were logging weight and food via Apple products (Watch and Health), MyFitnessPal, and RunKeeper, in decreasing order of prevalence. Tracker usage was more common in people younger than 50 years than in older populations, with 7.21% (3139/43,545) tracking at least 1 of the 4 activities compared with only 0.75% (5414/723,469) for individuals older than 50 years. Details on tracker usage by condition and type of activity can be seen in Table 2.

Association Between Adherence and Activity-Tracker Use (Trackers vs Nontrackers)

Across conditions, simply engaging in activity tracking was positively associated with medication adherence. When controlling for age and sex, people who tracked at least 1 of the 4 activities were significantly more adherent to medication than those who did not use any trackers, for both fixed adherent status (OR 1.33, 95% CI 1.29-1.36) and variable adherent status (OR 1.37, 95% CI 1.32-1.43; Figure 1). The results were similar when broken down by condition and specific activity tracked, with individuals who tracked a given activity more likely to be adherent than those who did not track the activity, for both fixed and variable methodologies and controlling for age and sex (Tables 6 and 7). The only exception was for diabetes in the fixed PDC model; the tracker-adherent relationship was not statistically significant (OR 1.12, 95% CI 0.96-1.30).

XSL•FO
RenderX

Figure 1. Odds ratios of percent of individuals who are adherent (proportion of days covered, PDC≥0.80) for individuals who use activity trackers versus nontrackers illustrate the association between activity-tracker use and medication adherence. Lines within the bars represent 95% CIs. Results are shown across conditions and activities tracked, while controlling for age and sex, and are significant at *P*<.001. PDC: proportion of days covered.

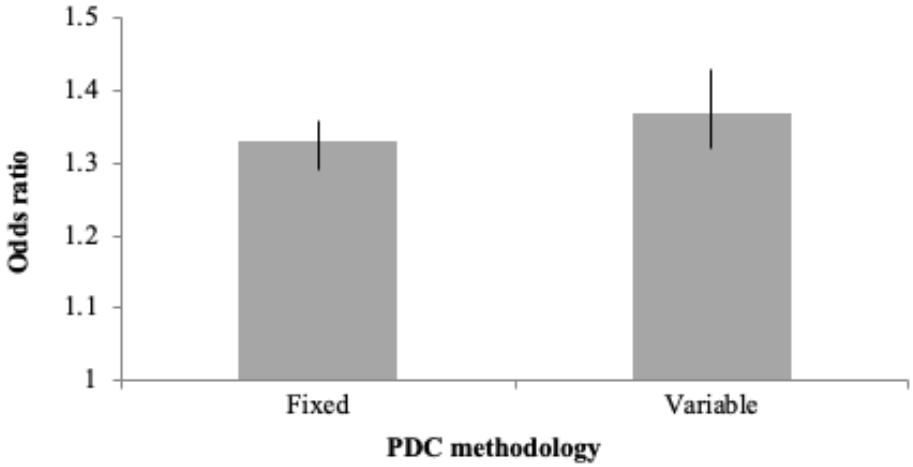


Table 6. The association between tracker use and medication adherence by condition in terms of the odds ratio of percent of individuals who are adherent (proportion of days covered≥0.80) for individuals who use activity trackers versus nontrackers. Odds ratios are significant at *P*<.001 unless otherwise noted.

Metric	Diabetes (n=791), OR[b] (95% CI)	Dyslipidemia (n=3599), OR (95% CI)	Hypertension (n=6472), OR (95% CI)
Fixed PDC[a] methodology	1.12 (0.96-1.30)[c]	1.23 (1.15-1.31)	1.38 (1.31-1.45)
Variable PDC methodology	1.31 (1.07-1.61)[d]	1.24 (1.13-1.37)	1.47 (1.36-1.59)

[a]PDC: proportion of days covered.

[b]OR: odds ratio.

[c]*P*=.14.

[d]*P*=.009.

Table 7. The association between tracker use and medication adherence by activity tracked in terms of the odds ratio of percent of individuals who are adherent (proportion of days covered≥0.80) for individuals who use activity trackers versus nontrackers. Odds ratios are significant at *P*<.001 unless otherwise noted.

Metric	Steps (n=10,499), OR[b] (95% CI)	Sleep (n=7404), OR (95% CI)	Weight (n=2428), OR (95% CI)	Food logs (n=1625), OR (95% CI)
Fixed PDC[a] methodology	1.34 (1.29-1.40)	1.35 (1.29-1.42)	1.21 (1.12-1.32)	1.20 (1.09-1.33)
Variable PDC methodology	1.37 (1.29-1.46)	1.38 (1.28-1.48)	1.37 (1.21-1.54)	1.26 (1.09-1.45)[c]

[a]PDC: proportion of days covered.

[b]OR: odds ratio.

[c]*P*=.001.

XSL•FO
RenderX

Association Between Adherence and Activity-Tracking Metrics

Frequency of activity tracking and activity level were also tied to medication adherence. When controlling for age and sex, individuals who tracked activities more frequently or who were more active on the basis of step counts were significantly more likely to be adherent for both fixed adherent status and variable adherent status. In the logistic regression model, increasing the adjusted tracking ratio by 0.5 increased the fixed adherent status OR by a factor of 1.11 (95% CI 1.06-1.16) and the variable adherent status OR by a factor of 1.14 (95% CI 1.07-1.22). Increasing the steps-per-day-tracked by 2000 increased the fixed adherent status OR by a factor of 1.07 (95% CI 1.04-1.09) and the variable adherent status OR by a factor of 1.05 (95% CI 1.01-1.09). The combined logistic regression model, which included both adjusted tracking ratio and steps per week tracked to assess whether each predictor had an additive effect, gave similar results for the fixed methodology adherent status. However, in the combined variable adherent status model, activity level did not have a significant association with adherent status.

Sensitivity Analysis

If any individual with at least 1 drug-class-level PDC≥0.80 was classified as adherent rather than basing the classification of adherent or nonadherent on whether the condition-level PDC was ≥0.80, engaging in activity tracking remained significantly and positively associated with medication adherence (fixed adherent status OR 1.21, 95% CI 1.18-1.25; variable adherent status OR 1.29, 95% CI 1.23-1.35). Just as in the base case, results were statistically significant when broken down by condition and specific activity tracked, except for diabetes in the fixed PDC model (OR 1.14, 95% CI 0.98-1.31). Frequency of tracking and activity level were also significantly tied to adherent status in the logistic regression model, except for the case of frequency of tracking for the variable adherent status.

Discussion

Principal Findings

Our analysis demonstrates that there is a significant relationship between medication adherence and activity tracking in individuals with chronic diseases, after controlling for age and sex. In particular, we found that people with diabetes, dyslipidemia, or hypertension who use activity trackers are more adherent to their medication than those who do not use activity trackers. In addition, medication adherence improves as consistency of tracker use and activity level increase. This analysis is an important first step in using additional available digital behavioral data to understand how health behavior ties to medication adherence.

These initial findings can be leveraged in a variety of clinically meaningful settings, such as targeting medication adherence initiatives. For example, considering that activity-tracking indices like the ones described require significantly shorter observation periods to be computed accurately for an individual as compared with PDC (days vs months), health plans or provider systems may use this information as a predictive score to selectively target new members for various medication adherence programs. Another use case is in enrollment for clinical trials of new chronic disease therapies. Trial recruitment can be targeted at patients who use activity trackers on a consistent basis, as these patients can be expected to have better medication adherence than nontracking patients of similar age and gender. This can minimize the sample size necessary for the trial. Targeting patients in this manner may be especially useful for enrolling younger patients, who tend to be less adherent yet significantly more likely to adopt activity trackers than older patients.

Our analysis also exposes many areas for further exploration into the relationship between patient behavior and medication adherence. We demonstrate a relationship between engaging in and consistency of activity tracking and medication adherence. One could suggest that individuals with tracker activity are being adherent to that intervention as they have a trait of behavior by which they are generally adherent to healthy activities and interventions. This is similar to the concept of the "healthy adherer"—the idea that healthier people are generally more adherent to medication than unhealthy people [6]. An interesting avenue of exploration is whether activity tracking is an independent indicator of adherence, in the sense that it has additional explanatory power in predicting medication adherence as compared with other healthy activities and interventions that show an association with medication adherence or adherence to other clinical care (eg, annual physical exams and glucose monitoring). Another future research avenue is to understand whether indices based on adherence to activity tracking are responsive to longitudinal changes in medication adherence at the individual level: does a sudden drop in activity tracking predict a time period of lower medication adherence? If this were to be the case, activity-tracking patterns could be used to responsively monitor medication adherence and deploy timelier interventions.

Limitations

This study has several limitations. First, the findings are associative and do not demonstrate a causal link between adherence to activity tracking and medication adherence. Further research should be conducted to determine experimentally whether active manipulation of the activity-tracking behavior (eg, through incentives) leads to improved medication adherence. Second, only approximately 1% of the population in this study had data available from a linked digital health tracker, which limits the generalizability of study results. The study did not include traditional pedometers, and it did not include data from individuals who used digital devices but who did not link and share their data. Digital health tracker use is more common in younger populations than older adults; however, we expect applicability of these study results to grow over time as digital health tracker use continues to grow. This also may be enhanced as tracking capabilities, especially accelerometers for tracking steps, have become the standard for smartphones and thus are more accessible to wide audiences as adoption of new smartphones grows. However, it remains to be seen if the link uncovered between tracking and adherence will still be present for those who use built-in smartphone technology for tracking, as less effort is required to track activity in this

XSL•FO
RenderX

scenario than purchasing or downloading a dedicated wearable or app.

This study is also subject to limitations common across claims-based analyses. Only insured individuals with pharmacy claims were included in the analysis; data from uninsured individuals or individuals who paid out of pocket for their medications were not captured. We used PDC to measure medication adherence, but it is an indirect measure of adherence that assumes filling a prescription equates to taking the medication. Individuals could fill but not consume prescriptions. This analysis also does not capture any medication discontinuations advised by a prescriber. Finally, we were only able to include interaction terms for age and sex in our analysis as other demographic factors such as income and education level were not available in the dataset.

Conclusions

This study demonstrates that individuals who engage in activity tracking have significantly higher medication adherence than those who do not track their activities when controlling for age and sex across thousands of people with diabetes, hypertension, and dyslipidemia. The results were typically not dependent on a specific condition or activity tracked. The positive association with medication adherence extended to frequency of activity tracking as well as to physical activity level, as measured by step count. Given the well-established link between poor medication adherence and increased health care costs and utilization, as well as mortality, improving medication adherence in chronic conditions continues to be a high-value objective. This study is the first step in developing a better understanding of how to use digital health tools to understand and drive medication adherence and subsequently lower the cost of managing chronic diseases.

Acknowledgments

The authors would like to thank Jennifer L Tran, PharmD; Laura E Happe, PharmD, MPH; and Courtney Brown, PharmD, for their assistance in the preparation and review of this manuscript.

SZ was affiliated with Humana, Inc, in Louisville, Kentucky, at the time of the research and is currently affiliated with inZights Consulting, LLC, in Seattle, Washington.

Conflicts of Interest

None declared.

References

1. Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. Prev Chronic Dis 2014 Apr 17;11:E62 [FREE Full text] [doi: 10.5888/pcd11.130389] [Medline: 24742395]
2. Gerteis J, Izrael D, Deitz D, LeRoy L, Ricciardi R, Miller T, et al. Agency for Healthcare Research and Quality. Rockville, MD; 2014 Apr. Multiple Chronic Conditions Chartbook URL: https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/prevention-chronic-care/decision/mcc/mccchartbook.pdf [accessed 2019-02-11] [WebCite Cache ID 766H5Ukcn]
3. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2014 update: a report from the American Heart Association. Circulation 2014 Jan 21;129(3):e28-292 [FREE Full text] [doi: 10.1161/01.cir.0000441139.02102.80] [Medline: 24352519]
4. American Diabetes Association. Economic costs of diabetes in the US in 2012. Diabetes Care 2013 Apr;36(4):1033-1046 [FREE Full text] [doi: 10.2337/dc12-2625] [Medline: 23468086]
5. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 4;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]
6. Iuga AO, McGuire MJ. Adherence and health care costs. Risk Manag Healthc Policy 2014;7:35-44 [FREE Full text] [doi: 10.2147/RMHP.S19801] [Medline: 24591853]
7. Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, et al. A meta-analysis of the association between adherence to drug therapy and mortality. Br Med J 2006 Jul 1;333(7557):15 [FREE Full text] [doi: 10.1136/bmj.38875.675486.55] [Medline: 16790458]
8. Roebuck MC, Liberman JN, Gemmill-Toyama M, Brennan TA. Medication adherence leads to lower health care use and costs despite increased drug spending. Health Aff (Millwood) 2011 Jan;30(1):91-99 [FREE Full text] [doi: 10.1377/hlthaff.2009.1087] [Medline: 21209444]
9. Pittman DG, Tao Z, Chen W, Stettin GD. Antihypertensive medication adherence and subsequent healthcare utilization and costs. Am J Manag Care 2010 Aug;16(8):568-576 [FREE Full text] [Medline: 20712390]
10. Zhao Y, Zabriski S, Bertram C. Associations between statin adherence level, health care costs, and utilization. J Manag Care Spec Pharm 2014 Jul;20(7):703-713 [FREE Full text] [doi: 10.18553/jmcp.2014.20.7.703] [Medline: 24967523]
11. Boye KS, Curtis SE, Lage MJ, Garcia-Perez L. Associations between adherence and outcomes among older, type 2 diabetes patients: evidence from a Medicare Supplemental database. Patient Prefer Adherence 2016;10:1573-1581 [FREE Full text] [doi: 10.2147/PPA.S107543] [Medline: 27574406]

XSL•FO
RenderX

12. Brookhart MA, Patrick AR, Dormuth C, Avorn J, Shrank W, Cadarette SM, et al. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. Am J Epidemiol 2007 Aug 1;166(3):348-354. [doi: 10.1093/aje/kwm070] [Medline: 17504779]
13. Brady CJ, Villanti AC, Pearson JL, Kirchner TR, Gupta OP, Shah CP. Rapid grading of fundus photographs for diabetic retinopathy using crowdsourcing. J Med Internet Res 2014;16(10):e233 [FREE Full text] [doi: 10.2196/jmir.3807] [Medline: 25356929]
14. [PwC Sverige | Audit, Business Advice, Tax].: PwC; 2016. The Wearable Life 2.0: Connected living in a wearable world URL: https://www.pwc.se/sv/pdf-reports/the-wearable-life-2-0.pdf [accessed 2019-02-08] [WebCite Cache ID 762lksnW7]
15. Peterson AM, Nau DP, Cramer JA, Benner J, Gwadry-Sridhar F, Nichol M. A checklist for medication compliance and persistence studies using retrospective databases. Value Health 2007;10(1):3-12 [FREE Full text] [doi: 10.1111/j.1524-4733.2006.00139.x] [Medline: 17261111]
16. Abughosh SM, Wang X, Serna O, Henges C, Masilamani S, Essien EJ, et al. A pharmacist telephone intervention to identify adherence barriers and improve adherence among nonadherent patients with comorbid hypertension and diabetes in a medicare advantage plan. J Manag Care Spec Pharm 2016 Jan;22(1):63-73 [FREE Full text] [doi: 10.18553/jmcp.2016.22.1.63] [Medline: 27015053]
17. Slejko JF, Ho M, Anderson HD, Nair KV, Sullivan PW, Campbell JD. Adherence to statins in primary prevention: yearly adherence changes and outcomes. J Manag Care Pharm 2014 Jan;20(1):51-57 [FREE Full text] [doi: 10.18553/jmcp.2014.20.1.51] [Medline: 24372460]
18. Kozma CM, Dickson M, Phillips AL, Meletiche DM. Medication possession ratio: implications of using fixed and variable observation periods in assessing adherence with disease-modifying drugs in patients with multiple sclerosis. Patient Prefer Adherence 2013;7:509-516 [FREE Full text] [doi: 10.2147/PPA.S40736] [Medline: 23807840]
19. Halpern MT, Khan ZM, Schmier JK, Burnier M, Caro JJ, Cramer J, et al. Recommendations for evaluating compliance and persistence with hypertension therapy using retrospective data. Hypertension 2006 Jun;47(6):1039-1048 [FREE Full text] [doi: 10.1161/01.HYP.0000222373.59104.3d] [Medline: 16651464]

Abbreviations

ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.
ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification.
IRB: institutional review board
OR: odds ratio
PDC: proportion of days covered

Edited by G Eysenbach; submitted 04.07.18; peer-reviewed by M Boukhechba, K Stawarz; comments to author 18.10.18; revised version received 13.12.18; accepted 31.12.18; published 20.03.19.

Please cite as:
Quisel T, Foschini L, Zbikowski SM, Juusola JL
The Association Between Medication Adherence for Chronic Conditions and Digital Health Activity Tracking: Retrospective Analysis
J Med Internet Res 2019;21(3):e11486
URL: http://www.jmir.org/2019/3/e11486/
doi:10.2196/11486
PMID:30892271

©Tom Quisel, Luca Foschini, Susan M Zbikowski, Jessie L Juusola. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention

Kevin Wright[1*], PhD; Carla Fisher[2*], PhD; Camella Rising[1*], PhD; Amelia Burke-Garcia[3*], PhD; Dasha Afanaseva[3*], MPH; Xiaomei Cai[1*], PhD

[1]George Mason University, Fairfax, VA, United States

[2]University of Florida, UF Health Cancer Center, Gainesville, FL, United States

[3]Westat, Rockville, MD, United States

[*]all authors contributed equally

Corresponding Author:
Kevin Wright, PhD
George Mason University
Northeast Module, Room 102
Fairfax, VA,
United States
Phone: 1 413 362 5611
Email: kwrigh16@gmu.edu

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2019/6/e14158/

Abstract

Background: Women are concerned about reducing their breast cancer risk, particularly if they have daughters. Social media platforms, such as blogs written by mothers, are increasingly being recognized as a channel that women use to make personal and family health–related decisions. Government initiatives (eg, Interagency Breast Cancer and Environmental Research Coordinating Committee) and researchers have called for scientists and the community to partner and disseminate scientifically and community-informed environmental risk information.

Objective: We developed and evaluated a blog intervention to disseminate breast cancer and environmental risk information to mothers. We teamed with *mommy bloggers* to disseminate a message that we developed and tailored for mothers and daughters based on scientific evidence from the Breast Cancer and the Environment Research Program (BCERP). We posited that the intervention would influence women's exposure to, acceptance of, and beliefs about environmental risks while promoting their intention to adopt risk-reducing behaviors.

Methods: Using a quasi-experimental design, we recruited 75 mommy bloggers to disseminate the breast cancer risk message on their respective blogs and examined the impact of the intervention on (1) readers exposed to the intervention (n=445) and (2) readers not exposed to the intervention (comparison group; n=353).

Results: Following the intervention, blog reader scores indicating exposure to the breast cancer risk and prevention information were greater than scores of blog readers who were not exposed (or did not recall seeing the message; mean 3.92, SD 0.85 and mean 3.45, SD 0.92, respectively; P<.001). Readers who recalled the intervention messages also had higher breast cancer risk and prevention information satisfaction scores compared with readers who did not see (or recall) the messages (mean 3.97, SD 0.75 and mean 3.57, SD 0.94, respectively; P<.001). Blog readers who recalled seeing the intervention messages were significantly more likely to share the breast cancer risk and prevention information they read, with their daughters specifically, than readers who did not recall seeing them (χ^2_1=8.1; P=.004). Those who recalled seeing the intervention messages reported significantly higher breast cancer risk and prevention information influence scores, indicative of behavioral intentions, than participants who did not recall seeing them (mean 11.22, SD 2.93 and mean 10.14, SD 3.24, respectively; P=.003). Most women ranked Facebook as their first choice for receiving breast cancer risk information.

Conclusions: Results indicated that blog readers who were exposed to (and specifically recalled) the BCERP-adapted intervention messages from mommy bloggers had higher breast cancer risk and prevention information exposure scores and higher breast

XSL•FO
RenderX

cancer risk and prevention information satisfaction and influence scores than those who did not see (or recall) them. Mommy bloggers may be important opinion leaders for some women and key to enhancing the messaging, delivery, and impact of environmental breast cancer risk information on mothers.

(J Med Internet Res 2019;21(3):e12441) doi:10.2196/12441

KEYWORDS

breast cancer; environment and public health; risk reduction behavior; blogging; social networking; social media; health communication; information dissemination

Introduction

Overview

Breast cancer is the second most common cancer diagnosed in women [1]. Women, particularly those with a personal or family history of the disease and mothers with daughters, are concerned about what they can do to reduce the risk of breast cancer [2-7]. In addition, when mothers advise their daughters about how to reduce the risk, daughters tend to follow their advice into adulthood [7]. Although clinicians, scientists, and advocacy groups prioritize cancer screening to reduce the risk (eg, mammograms and genetic testing), scientists have discovered that environmental exposures (eg, having contact with certain chemicals or particles through food, air, water, or touch) impact women's risk [8].

In response to increasing research on breast cancer risk and environmental exposures, Congress passed the *Breast Cancer and Environmental Research Act* in 2008. This resulted in the initiation of the Interagency Breast Cancer and Environmental Research Coordinating Committee (IBCERCC), which was tasked with examining the state of this research. Their report identified 7 recommendations, 2 of which centered on the need for scientists to become engaged in dissemination, specifically prioritizing community stakeholder engagement and translating research for the public [9].

The authors are members of 1 group actively involved in the creation and dissemination of such research—the Breast Cancer and the Environment Research Program (BCERP). Founded in 2003 and funded by the National Institute of Environmental Health Sciences (NIEHS) and the National Cancer Institute (NCI), BCERP researchers have conducted decades of studies linking environmental factors and breast cancer risk. This includes identifying endocrine-disrupting chemicals (eg, phthalates and bisphenol A [BPA]) that could increase women's risk and are found in some personal care and household products (eg, shampoo, detergent, and plastic bottles) [8]. BCERP provides actionable steps that women and their family members can take to reduce their exposure to these environmental risks. In recent years, BCERP has funded projects focused specifically on dissemination. The program currently disseminates this evidence-based information free to the public via their website [10].

Although women seek Web-based information about how to reduce personal and familial risk of breast cancer [11], they may not be obtaining reputable or scientifically based information that organizations such as BCERP provide. Readability and accessibility are also issues. A recent review of breast cancer and environmental risk information online showed that most content is not disseminated at a readable level with health literacy in mind [12]. There is also a plethora of information on the internet from divergent sources, including medical websites that are not regulated [13]. Websites are discovered based on users' search terms with no guarantee that reputable sources such as BCERP are even reached. Thus, as the IBCERCC report also identified, researchers need to do a better job of disseminating breast cancer information in ways that reach women. Dissemination approaches should also strive to reach more women or have the potential for rapid diffusion of information, the process by which information is spread across channels rapidly [14].

Social media offers an ideal dissemination channel with the potential to reach an expansive group of women. Current evidence [15] has shown that:

> *Social media are becoming preferred methods of health promotion as evidence builds showing their effectiveness in reaching public audiences...Evidence about social media's impact on health knowledge, behavior, and outcomes shows these tools can be effective in meeting individual and population health needs.*

Social media platforms, such as blogs, offer a way to disseminate health information rapidly and expansively, educate the public, and promote healthy behavior. Still, *careful application* and evaluation is needed to obtain the desired outcomes [15] as blog-based interventions can result in limited exposure, messages not being tailored to the targeted audience, or messages that do not account for health literacy or culture [12,13,16]. Also, when it comes to online dissemination about environmental breast cancer risk, accuracy is of concern [12]. This is particularly true with blogs [17], which have been shown to be the least accurate in comparison with other online sources [12].

A recent review also showed that online information about environmental exposures is particularly complicated and, at times, misunderstood [12], with more commercial-focused online sources reporting known risk associations within exposures like deodorant, which opposes research showing no associated risk presented on more credible sites (eg, American Cancer Society). Additionally, this review found that while research may be presented accurately, the conclusions are not always accurately drawn and appear to be driven by personal motivations [12].

To ensure access and accuracy in the dissemination of breast cancer risk information, scholars have recently called for a

XSL•FO
RenderX

community-engaged approach, linking scientists with community stakeholders to ensure that the public has access to the information and that it is disseminated in a manner that is readable and relatable [18]. In line with the IBCERCC's recommendation [9], mothers who blog, or *mommy bloggers*, are an ideal community partner to disseminate information via blogs to reach women. Their blog posts could integrate scientifically based information in a manner that blog readers can identify. Also, blogs written by mothers are increasingly recognized as a channel that women utilize to make personal and family health–related decisions [19-21], and as such, mommy bloggers are often viewed as relatable and trustworthy by other mothers.

In addition, evidence-based strategies are integral to successfully disseminating health information via blogs [15]. These strategies include tailoring messages with the audience or readers in mind. Tailoring may involve user-generated (ie, blogger-generated) content to promote acceptance (by using a trusted source that content readers identify with), encouraging multipronged approaches (multiple social media platforms), integrating theory during intervention development, and using tools to evaluate the impact of the intervention on readers [15,22].

With these evidence-based strategies in mind, we aimed to develop and evaluate a targeted blog intervention, teaming with mommy bloggers to disseminate scientifically informed breast cancer and environmental risk information (adapted from BCERP) with the potential for rapid diffusion. The goal was to influence readers' (ie, mothers') exposure to, acceptance of, and beliefs about environmental risks while promoting intentions to adopt risk-reducing behaviors. In addition to testing the efficacy of this approach, we sought to shed light on how such an intervention may extend farther than targeted readers or how readers may share the information across other platforms.

Getting the Message Out There: Blogging About Reducing Breast Cancer Risk

For many, the internet is a primary source of health information, including information about reducing breast cancer risk [11,23,24]. People increasingly seek health information through social media, such as Facebook, Twitter, and blogs, where knowledge is shared within a relational network facilitated through an online community [25]. The interactive and relational nature of social media allows individuals to connect with and reach broader audiences to address a range of health issues [26,27]. Therefore, blogs may represent an optimal social media channel for disseminating breast cancer risk information [19].

Blogs can be rich in information and promote interaction among bloggers and readers. For instance, bloggers provide unique content for their readers compared with other online communities [28]. They display a keen understanding of readers' health needs and beliefs [19,28] and typically deliver information in varied forms (eg, text, multimedia, and links to Web resources) that are shared interactively via bloggers' posts and exchanges of readers' comments [19,29]. This online interaction generates a larger community network, extending beyond the blog [30-32]. Bloggers and readers may share blog content online (eg, via Facebook, Twitter, and Pinterest) and offline (eg, with family, friends, coworkers, and community groups). Thus, blogging can reach a broad audience.

In addition to facilitating interaction and reach, blogs represent a unique channel where specific audiences (such as mothers) can be reached and influenced. Bloggers can customize health information in ways that readers relate to, as bloggers are likely perceived by readers as being similar in terms of beliefs, experiences, and language [33,34]. In other words, blog readers identify with the bloggers (eg, as a woman, mother, or survivor). This increases the likelihood of readers' acceptance, sharing, and adoption of the information posted by the blogger [35,36].

Currently, there are about 3.9 million mothers in the United States who identify as a blogger, and many mothers with children at home turn to blogs for advice about health issues [19,37]. Blogs written by mothers or those that focus on motherhood topics—sometimes called *mommy blogs* —offer the possibility of reaching an important target audience for breast cancer risk information: women and their family members [20,38,39].

Teaming With Mommy Bloggers for Rapid Diffusion of Information: Trusted Sources Women Can Identify With

Readers relate to their blogs in part because mommy bloggers routinely tailor messages to their target audience (eg, using strategies such as feminine rhetoric, humor, or personal stories with photographs that women identify with) [40-43]. Bloggers also direct their readers to relevant and accurate health information on the internet [44]. Given their trusted status among mothers, mommy bloggers could be used to disseminate risk-reducing information and persuade readers to adopt or respond to it [45]. In line with Rogers' widely used diffusion of innovation theory [14], mommy bloggers may be key to disseminating such information.

Influencing Mothers' Risk-Reducing Perceptions and Behavior: Diffusion of Innovation Approach

As diffusion of innovation theory purports, early adopters or innovators (readers) and opinion leaders (bloggers) are central to the dissemination of new ideas to the public. Mommy bloggers could promote rapid diffusion of the information through larger blogging networks when bloggers adopt the message (ie, post the content and endorse it), which can influence readers to do the same. Thus, bloggers can be viewed as opinion leaders who may drive diffusion and uptake of environmental breast cancer risk information by their readers [14,19,45]. The risk-related content may heighten the urgency of the message, and the message's effects may be further amplified through bloggers' and readers' comments, shares, and likes [46].

This interactive and dynamic diffusion of information also has the potential to influence women's behavior, which is essential to reducing cancer risk. Blog intervention studies show that blogger-reader interactions can create a sense of immediacy about health topics that encourages readers' adoption of healthy behaviors [47]. Moreover, the tailored context of the blog, in conjunction with readers' perceived source similarity [35], can

XSL•FO
RenderX

influence women's adoption of recommended health behaviors [47].

Despite the potential for mommy bloggers to serve as trusted or relatable sources to disseminate breast cancer risk information to mothers, it is critical that bloggers disseminate accurate, scientifically informed content [12]. Moreover, content should in part be *user-generated* with a target audience in mind, meaning that it should be produced in part with the bloggers' input [15]. To achieve this, collaborative message development (between scientists and bloggers) is warranted to ensure women receive accurate information in a format they trust and relate to.

Targeting Mommy Blog Readers: Collaborative Message Development

As described, federally funded BCERP scientists and community partners have developed and disseminated research to bring awareness to the public about linkages between environmental exposures and breast cancer risk. BCERP provides several educational materials online, including a toolkit for mothers with daughters. These mother-daughter–focused materials specifically address how mothers can reduce their own and their daughter's risk by adopting healthier lifestyle choices together (eg, eliminating products with BPA and phthalates).

Many mothers with daughters are concerned and feel uncertain about breast cancer prevention and their daughters' risk [2,3,48]. However, when mothers advise daughters about how to reduce risk, daughters adhere to their mother's advice into adulthood [7]. Talking about risk can be challenging. Younger daughters often avoid or withdraw from such conversations [4,49], which can trigger a physiological stress response[50]. Mothers have reported using third-party approaches (eg, a magazine article) to prompt interaction and ease their daughter's comfort during discussions [3,4]. Relatedly, recent research shows that third-party Web-based approaches (eg, videos about BPA/perfluorooctanoic acid and radiation risk) favorably influence mothers' and daughters' prevention behavior [16,51].

Mommy bloggers' posts could change mothers' knowledge and beliefs about environmental breast cancer risk factors and simultaneously function as a third-party approach to facilitate mother-daughter communication. Even though the information on BCERP's website is scientifically informed and expansive, the format and manner in which these materials (a brochure, flyer, and public service announcement) are delivered assume women will obtain them, first by finding their website and then by reading them on their own. Passive dissemination approaches such as these do not ensure the information reaches the target audience and, when used alone, are not likely to be translated to practice or result in behavioral change [52]. Moreover, these materials are often lengthy and not in a format ideal for social media dissemination.

A more active diffusion of information or dissemination approach on an interactive Web platform in which an influencer (eg, blogger) communicates with women (their readers) is more likely to reach the targeted audience. Additionally, this approach may prompt interaction within and outside the blog network. By teaming with mommy bloggers, the evidence-based information could be integrated into a user-generated format ideal for social media and delivered in a manner that their readers (mothers) can relate to.

Research Foci

As the aforementioned research demonstrates, mothers look to mommy bloggers for health information. By partnering with mommy bloggers to disseminate evidence-based environmental risk information about how mothers (and daughters) can reduce breast cancer risk, we sought to increase women's exposure to, satisfaction with, and acceptance of environmentally focused risk-reducing information. Thus, the first hypothesis (H1) was posited:

> *H1: The use of a targeted online blog intervention will increase blog reader's exposure to and satisfaction with the breast cancer risk and prevention information compared with blog readers who are not exposed to (or who did not recall seeing) the intervention messages.*

Given the interactive nature of blogging, we also wanted to encourage the diffusion of this information within women's larger social networks. Previous research has not examined how intervention messages stemming from mommy blogs might influence interaction and information sharing among other social media platforms. Therefore, the following research question (RQ) was posed:

> *RQ1: How does the use of a targeted online blog intervention encourage interaction and information sharing about breast cancer risk and prevention messages across other online social media networks?*

Behavior change is the ultimate goal of a health promotion intervention. According to the integrative model of behavioral prediction (a reasoned action theory approach to health promotion), several antecedents of behavior change should be evaluated to understand and predict whether women will take action [53,54]. These variables include acceptance of a health-related message, beliefs about health and risk, and intention to change. On the basis of this framework, we posited the second hypothesis (H2):

> *H2: The use of a targeted online blog intervention will increase breast cancer risk and prevention message acceptance, beliefs, and intentions to adopt the guidance.*

Finally, to reach a larger audience, it is critical that this information is disseminated (or diffused) through channels perceived as optimal by women [15]. Therefore, we aimed to understand mothers' preferred communication channels for breast cancer and environmental risk and prevention information. More than half of the US population use 2 or more social media platforms [55], and some individuals prefer traditional, interpersonal (eg, face-to-face) communication channels because of cultural norms about discussing health and risk topics [56]. To ensure women perceive mommy bloggers as an optimal channel but also provide opportunities to refine our intervention as necessary, we sought to learn about perceived optimal channels more broadly. Therefore, the following inquiry was posed:

XSL•FO
RenderX

RQ2: What are blog readers' preferred media/communication channels for receiving information about breast cancer risk and prevention?

Methods

This targeted social media intervention study involved a quasi-experimental design to assess women's exposure to, acceptance of, and beliefs about environmental risks while promoting intentions to adopt risk-reducing behaviors.

Participants and Recruitment

After the institutional review board approval, 3 groups of participants were recruited: (1) bloggers involved in the intervention, (2) readers exposed to the intervention (intervention group), and (3) readers not exposed to the intervention (comparison group). A convenience sample of mommy bloggers, all women, was recruited through The Motherhood [57], a network of more than 3000 diverse mothers who blog about various topics including health. During the recruitment process, efforts were made to find participants of different racial or ethnic groups and geographic locations, as well as prior experiences with breast cancer (either personally and/or by family members). A total of 75 mommy bloggers agreed to participate in the intervention (see Table 1 for blogger characteristics).

Upon consent, all participants completed an online survey with items about sociodemographics and breast cancer history. About 35 mommy bloggers had previously written about breast cancer in their blogs in the past year. All of the bloggers in the sample had children and 52 bloggers had daughters, specifically. Bloggers posted messages on their blogs to recruit readers for the intervention group.

The reader intervention group was comprised 445 blog readers (435 women and 10 men) who follow one or more of the 75 bloggers involved in the intervention. The participating mommy bloggers were asked to recruit their readers and direct them to an online postintervention survey link. A total of 353 blog readers (341 women and 12 men) made up the reader comparison group. To minimize contamination between the 2 groups, readers in the comparison group were recruited through a separate set of mommy bloggers who were affiliated with The Motherhood network, but who did not blog about breast cancer and were not involved in the intervention (see Table 2 for blog readers' characteristics). The comparison group bloggers wrote about a wide range of issues, including parenting concerns and other health issues besides breast cancer.

Table 1. Sociodemographics and breast cancer or risk history of bloggers (n=75).

Characteristics	Bloggers
Age (years), mean (SD); range	37.88 (7.02); 25-61
Income (US $) based on reported zip code, median	65,611
Race or ethnicity, n (%)	
White	43 (58)
African American or black	15 (19)
Asian or Pacific Islander	2 (3)
Hispanic	14 (18)
Other	1 (1)
Education level, n (%)	
High school diploma	6 (8)
Some college	13 (14)
2-year college degree	6 (8)
4-year college degree	29 (37)
Graduate degree	21 (27)
Breast cancer or risk history, n (%)	
Diagnosed with breast cancer	3 (4)
Breast cancer 1 or 2 mutation positive	1 (1)
Family history of breast cancer	50 (64)
Diagnosed first-degree relative	13 (17)

XSL•FO
RenderX

Table 2. Sociodemographics and breast cancer or risk history of readers, intervention group (n=445) and comparison group (n=353).

Characteristics	Intervention group	Comparison group
Age (years), mean (SD); range	39.33 (10.82)[a]; 19-83	37.74 (10.18); 19-81
Average income (US $) by reported zip code	65,709	67,435
Race or ethnicity, n (%)		
White	389 (87.4)	304 (86.1)
African American or black	24 (5.4)	18 (5.1)
Asian	13 (2.9)	5 (1.4)
Hispanic	28 (6.3)	27 (7.6)
Native American or Alaska native	5 (1.1)	4 (1.1)
Other	11 (2.5)	3 (0.8)
Education level, n (%)		
Less than high school	5 (1.1)	2 (0.5)
High school graduate	37 (8.3)	26 (7.4)
Some college	86 (19.3)	69 (19.5)
2-year college degree	65 (14.6)	32 (9.1)
4-year college degree	152 (34.2)	145 (41.1)
Graduate degree	100 (22.4)	79 (22.4)
Breast cancer or risk history, n (%)		
Diagnosed with breast cancer	16 (3.5)	10 (2.8)
Breast cancer 1 or 2 mutation positive	7 (1.5)	7 (1.9)
Family history of breast cancer	119 (26.7)	150 (42.4)
Diagnosed first-degree relative	72 (16.2)	59 (16.7)

[a]Significant difference for age between intervention and comparison groups, t_{794}=2.110; P=.03.

The reader intervention group completed an online postintervention survey after the message exposure period. In terms of frequency of use, 367 intervention group respondents reported reading mommy blogs one or more days per week (78 indicated less than once per week). Many (n=280) reported having daughters.

Blog readers recruited for the reader comparison group completed the same postintervention online survey as the intervention group. Respondents from this group (n=307) mentioned reading mommy blogs one or more days per week (46 reported less than once per week). Reader comparison group respondents (n=227) also reported having daughters. Participants in the intervention and comparison groups only differed significantly by age (t_{794}=2.110; P=.03).

Intervention bloggers were provided incentives depending on their reach. Bloggers with an overall reach of up to 100,000 were given US $125; bloggers with an overall reach of up to 200,000 were given US $200; and bloggers with an overall reach above 250,000 were given US $300. Bloggers who disseminated the reader comparison group survey were offered US $50. Readers had the option to enter a raffle to win 1 out of 5 US $100 gift cards.

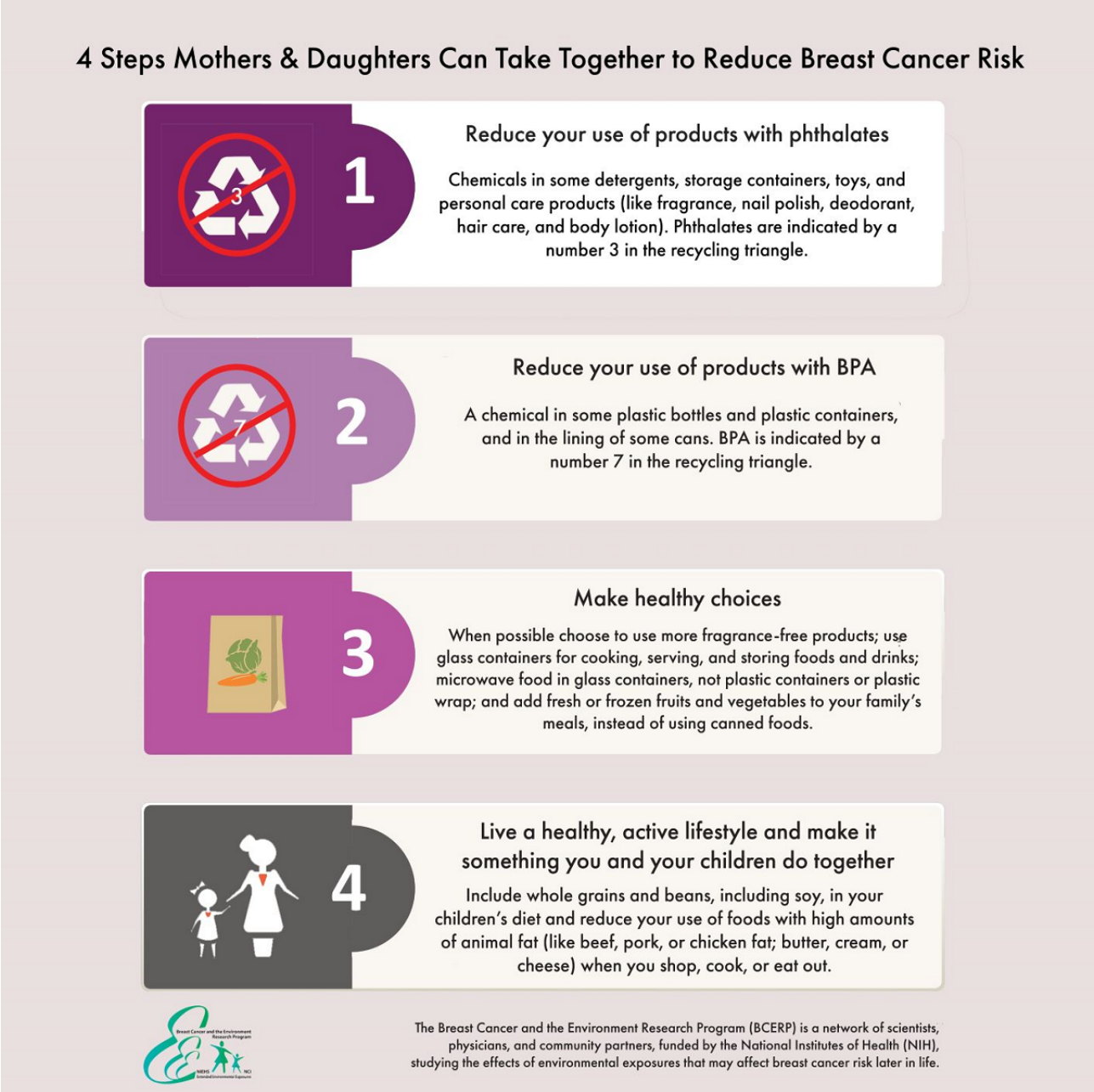
Intervention Development and Dissemination

Message Development

Our goal was to develop a message that was based on scientific evidence as well as informed by the community we sought to reach. First, to develop a scientifically informed message suitable for social media dissemination, 2 authors analyzed BCERP's online educational materials targeting mothers and daughters [58]. The authors identified salient themes that would provide mothers with actionable steps they could take to reduce their own and their daughter's environmental risk for breast cancer development. The remaining authors reviewed the analysis and confirmed the identification of 4 salient messages (or 4 steps mothers and daughters could take) communicated across the materials. To facilitate social media dissemination, a JPEG image with the 4 steps was created using language extracted from BCERP's materials to ensure that information was accurate and uneditable (see Figure 1). Second, we sought to ensure that the message disseminated would also be community-informed (or user-generated) to better ensure women identified with the information. This aspect involved the bloggers adding their own content (which is further described below in procedures).

XSL•FO
RenderX

Figure 1. Intervention message adapted from the Breast Cancer and the Environment Research Program materials.



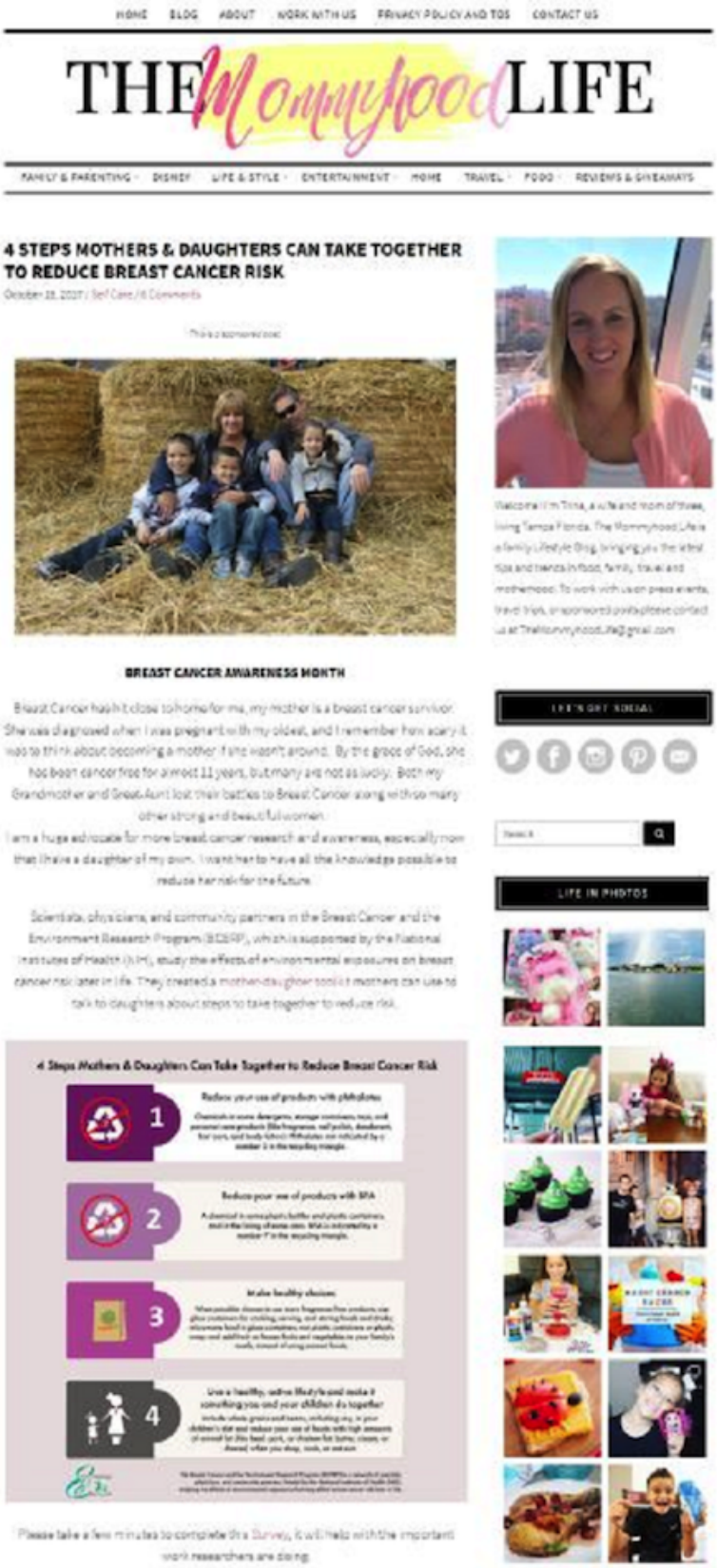
Intervention Procedures

Bloggers were provided with a document that included the study protocol and instructions about when and how to integrate the image or intervention message into their blog post. They also received procedures for completing the online survey and for recruiting their readers to do the same. Bloggers were asked to write a blog post during the National Breast Cancer Awareness Month (October-November 2017) that includes the image and targets mothers with daughters. In addition to the image, bloggers were asked to include 2 sentences of text to identify the source of the information (ie, BCERP). They were also given a link to additional online BCERP resources to include in their post.

To ensure that information disseminated on the blog was in part user-generated [15], bloggers were asked to use their experience with their blog audience to integrate the image and required text in a manner that they felt would be appealing to them (see Figure 2). Bloggers were not required but encouraged to promote their post via other social media channels (eg, Facebook, Twitter, and Instagram) and to consider using the hashtags #BCERP, #MotherDaughter, #BreastCancerRisk, and/or #BreastCancerAwareness in their posts. They were told that they could provide a link to their original blog post or link to the BCERP toolkit directly. Finally, bloggers were asked to post a message with links to the surveys to remind readers to complete the online reader survey.

XSL•FO
RenderX

Figure 2. Example mommy blog post with adapted intervention message.



XSL•FO
RenderX

Analysis of Breast Cancer Messages Associated With Seeded Mommy Blogs

To assess the degree to which the intervention encouraged interaction about breast cancer risk messages among blog users (research question [RQ] 1), the researchers used an online tracking program, *Tracx* [59], to capture blog and related online social media content associated with the targeted online blog intervention. *Tracx* is a social media monitoring and analytics program that uses Boolean search querying to track posts, mentions, and conversations across different online channels (eg, social media platforms, news outlets, blogs, and forums) through the program algorithms. Additionally, it populates the number of author's followers and total engagements per post.

The sample was collected between October 1 and December 31, 2017, by one of the authors, corresponding with the start of the intervention and 2 months following it. To ensure exhaustiveness, manual identification of posts and ensuing reactions across all platforms was performed in addition to using *Tracx*. Out of 7897 total posts reported, 293 additional blog posts, social media posts, and engagements were added manually.

The frequency of the following key variables was reported by *Tracx*: A post was the initial piece of breast cancer content submitted to a blog post and/or other social media used by the participating bloggers to reach their reader audience. Engagement referred to any interaction with a post on a social network including Google+, Facebook, Twitter, and Instagram. An interaction included likes, retweets, replies, shares, favorites, and comments in relation to the original post. Comments were defined as any text or emoji response to a post on either the original blog post or a post on social media.

Reach was operationalized as the number of unique people who were potentially exposed to the original blog posts. The overall reach of bloggers was calculated by the number of readers who followed their blog. Multiple reactions from the same person across channels were counted as one reach. *Tracx* was able to match up the engagement, comment, and reach with each mommy blogger original post. Figure 3 demonstrates the flow and timeline of the intervention process.

Blog Reader Survey Measures

The blog reader survey included a variety of measures that are consistent with the diffusion of innovation theory [14], including measures of message exposure, message satisfaction, message acceptance or influence, and likelihood of sharing the BCERP messages with the members of other social networks.

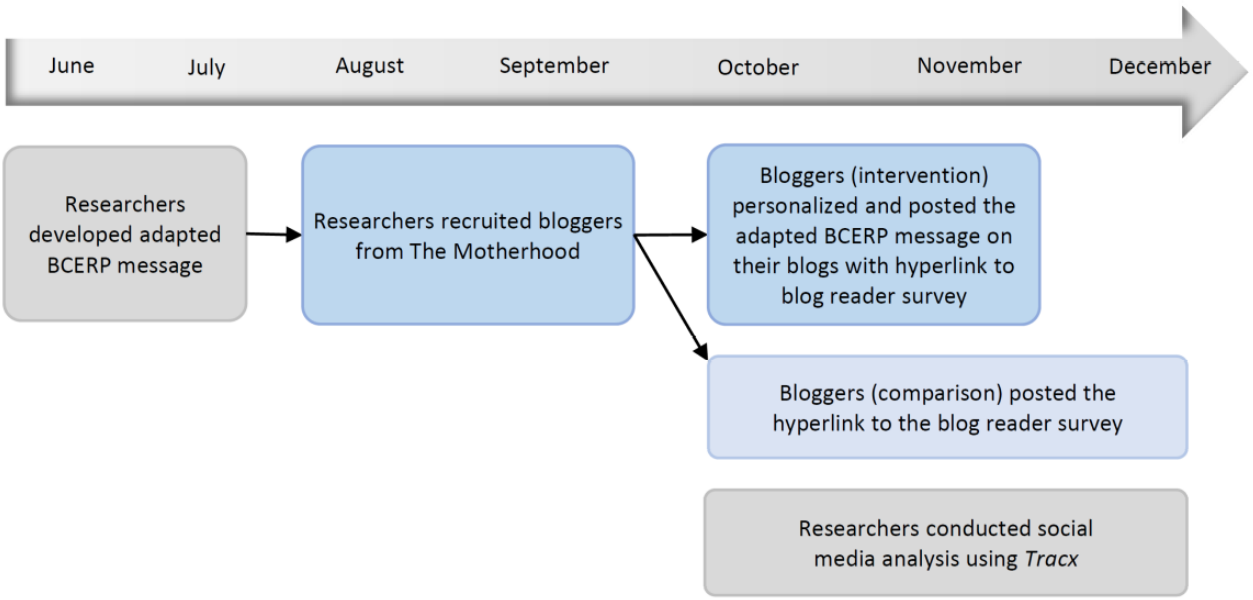
Intervention Message Recall

Recall of the intervention messages was assessed by the following question: *Did you notice breast cancer risk and prevention information from the BCERP (Breast Cancer and the Environment Research Program of the National Institutes of Health) organization in any of the blog posts you read recently?* Respondents were also asked to list up to 5 different blogs they followed in the past year to help the researchers detect exposure to the intervention message on other blogs the readers follow. In this way, the researchers could also determine whether reader comparison group participants had been exposed to the intervention message.

Exposure to Breast Cancer Risk and Prevention Information

Exposure to information about breast cancer risk and prevention was measured with a 5-point Likert-type scale item: *The blog(s) I follow has/have increased my exposure to information about breast cancer risk and prevention*, ranging from *Strongly Disagree* to *Strongly Agree*.

Figure 3. Intervention process and timeline (2017). BCERP: Breast Cancer and the Environment Research Program.



XSL•FO
RenderX

Satisfaction With Breast Cancer Risk and Prevention Information From Blogs

Satisfaction with the breast cancer risk and prevention information that blog readers were exposed to was measured using a 5-point Likert-type scale item: *Please indicate how satisfied you are with the way the blog(s) you follow present breast cancer risk and prevention information*, ranging from *Not Satisfied at All* to *Very Satisfied*.

Breast Cancer Risk and Prevention Message Acceptance or Influence

Breast cancer risk and prevention message acceptance or influence was measured with 3 5-point Likert-type scale items, including: *How much did reading the recent breast cancer risk and prevention blog posts influence your acceptance of the message regarding breast cancer risk?*; *How much did reading the recent breast cancer risk and prevention blog posts influence your beliefs about breast cancer risk?*; and *How much did reading the recent breast cancer risk and prevention blog posts influence your behaviors to reduce your risk of breast cancer?* Each response ranged from *Not at All* to *Very Much*. These measures were added together to create a composite score (Cronbach alpha=.87).

Likelihood of Sharing Breast Cancer Risk and Prevention Information From Blogs

Participant likelihood of sharing breast cancer risk and prevention information from blogs was measured with a 5-point Likert-type scale item: *In the future, how likely are you to share with others any information about breast cancer risk and prevention from the breast cancer risk blog post(s) you read in the past year?*, ranging from *Not Likely at All* to *Very Likely*.

Sharing of Breast Cancer Risk and Prevention Information

All participants were asked, *With whom did you share the information from the breast cancer risk and prevention blog post(s) you read?* Participants could make their selection from a drop-down menu that included various family members, friends, coworkers, and others within their social networks.

Perceived Importance of Breast Cancer Risk and Prevention as a Topic

Participant perceptions of the importance of breast cancer risk and prevention as a blog topic was measured with a 5-point Likert-type scale item: *How important to you is breast cancer risk and prevention as a topic?*, ranging from *Unimportant* to *Very Important*.

Statistical Analyses

Significance Tests

Statistical analyses were conducted using IBM SPSS. Independent sample *t* tests were performed in an initial assessment of differences between blog reader intervention and comparison groups. Analysis of covariance (ANCOVA) was conducted to test H1 and H2. Finally, a chi-square analysis was performed to analyze between-group differences in sharing breast cancer risk and prevention information with others, specifically daughters. Values were considered statistically significant at $P<.05$.

Initial Comparison of Blog Reader Intervention and Comparison Groups

In terms of the perceived importance of breast cancer risk and prevention as a topic, there was no significant difference between members of the intervention and comparison groups, $t_{794}=-1.004$; $P=.31$. As expected, participants in the reader intervention group were significantly more likely to recall seeing one or more of the intervention messages (n=165), $\chi^2_1=42.9$; $P<.001$. However, 32 individuals in the comparison group also recalled seeing the intervention messages on 1 of 5 other blogs they follow. These individuals were treated as members of the intervention or exposed condition in analyses, bringing the total number of participants who recalled seeing one or more of the intervention messages to 197 (n=197) versus the comparison group (n=321).

Results

H1 posited that the use of a targeted online blog intervention would increase blog reader's exposure to and satisfaction with the breast cancer risk and prevention information they received compared with blog readers who were not exposed to (or who did not recall seeing) the intervention messages. An ANCOVA test comparing breast cancer risk and prevention information exposure scores between participants who recalled seeing the BCERP intervention messages versus those who did not, while controlling for level of education, revealed that participants who recalled seeing the intervention messages had higher breast cancer risk and prevention information exposure scores (mean 3.92, SD 0.85) than those who did not (mean 3.45, SD 0.92; $F_{1,409}=25.78$; $P<.001$; partial eta$^2=0.06$). In addition, those readers who recalled the intervention messages had higher breast cancer risk and prevention information satisfaction scores (mean 3.97, SD 0.75) than participants who did not recall seeing the intervention messages (mean 3.57, SD 0.94; $F_{1,409}=19.86$; $P<.001$; partial eta$^2=0.047$) supporting H1.

RQ1 asked how a targeted online blog intervention might encourage interaction or information sharing about breast cancer risk and prevention messages across other social media networks. The *Tracx*-assisted content analysis of the breast cancer messages associated with the participating bloggers' adapted intervention messages captured message reach, message engagement, and number of comments as detailed in Table 3.

XSL•FO
RenderX

Table 3. Frequencies of message posts, reach, engagement, and comments.

Key metrics for social networking sites, n	Blog	Google+	Facebook	Twitter	Instagram	Total
Posts	75	12	25	98	12	222
Reach[a]	310	21	247	648	1048	2153
Engagement[b]	1778	11	432	1742	3712	7675
Comments	283	0	60	0	181	524

[a]Defined as the number of unique people who were potentially exposed to the original blog posts.

[b]Defined as any interaction with a post on a social network.

Moreover, a chi-square test revealed that blog readers who recalled seeing the intervention messages were significantly more likely to share the breast cancer risk and prevention information they read, with their daughters specifically, when compared with individuals who did not recall seeing them, χ^2_1=8.1; *P*=.004.

H2 predicted that a targeted online blog intervention will influence breast cancer risk and prevention message acceptance, beliefs, and intentions to act on the information provided to reduce breast cancer risk. An ANCOVA test comparing breast cancer risk information influence scores of intervention and comparison groups, while controlling for level of education ($F_{1,407}$=13.89; *P*<.001) and age ($F_{1,407}$=8.93; *P*=.003), showed significantly higher breast cancer risk and prevention information influence scores (mean 11.22, SD 2.93) for those who recalled seeing the intervention message when compared with those who did not (mean 10.14, SD 3.24; $F_{1,407}$=9.16; *P*=.003; partial eta^2=0.022, supporting H2).

The second research question (RQ2) asked about mommy blog readers' media or channel preferences for receiving future information about breast cancer risk and prevention. For the intervention group, 14.4% (64/445) of participants ranked Facebook as their first choice for receiving future information about breast cancer risk, followed by 0.07% (31/445) who ranked Twitter as their first choice, 0.05% (22/445) who ranked *other social media* first, 0.02% (8/445) who ranked blogs as their first choice, and 0.01% (4/445) who ranked email as their first choice.

In the comparison group, 0.09% (35/353) of participants ranked Facebook as their first choice for receiving future information about breast cancer risk and prevention, 0.06% (21/353) ranked Twitter as their first choice, 0.01% (6/353) ranked *other social media* as their first choice, 0.01% (5/353) ranked blogs as their first choice, 0.008% (3/353) ranked email as their first choice, and 0.003% (1/353) ranked mail (eg, postal system) as their first choice. In terms of other preferred social media, the most frequently mentioned channels were Instagram (n=27) and Pinterest (n=9). Other suggested channels for disseminating breast cancer risk and prevention messages were YouTube (n=3), text messages (n=3), medical websites, such as WebMD or PubMed (n=2), Google (n=2), and Tumblr (n=1).

Discussion

Principal Findings

The aim of this study was to examine whether a targeted intervention in which mommy bloggers disseminate evidence-based information could increase mothers' exposure to and dissemination of breast cancer environmental risk and prevention information and, ultimately, be a potential means of persuading mothers to engage in behavioral changes that could reduce their disease risk. To better determine the potential for women to engage in risk-reducing lifestyle behaviors advocated for in the messages, we investigated several antecedents to behavior change (message satisfaction, acceptance and beliefs about the information, and behavioral intentions to adopt the guidance). Additionally, to ascertain the potential reach of this intervention approach, we explored women's dissemination of and preferences for information via other social media platforms, allowing us to explore the potential for rapid diffusion via this intervention approach.

Results indicated that mommy blog readers who were exposed to (and specifically recalled) the seeded intervention message adapted from BCERP guidance had higher breast cancer risk information exposure scores and higher breast cancer risk and prevention information satisfaction and influence scores than those who did not see (or recall) them. These findings are consistent with previous theory and research on the influence of social networks and opinion leaders within them on health perceptions [14,19,32,45]. Thus, mommy bloggers may serve as important opinion leaders for some women. Moreover, mommy bloggers may be key to enhancing the messaging, delivery, and impact of environmental breast cancer risk information on mothers.

Our intervention involved both evidence-based and user-generated message development. As such, our targeted intervention demonstrated that future designers of breast cancer communication interventions can provide online opinion leaders such as mommy bloggers with health information adapted from evidence-based research from authoritative groups (eg, National Institutes of Health [NIH]). This may be especially important in better ensuring that information posted on blogs is accurate and credible given a recent study that showed that blogs are often less accurate than other online sources [12]. Our findings also show that it is important to include content generated by the user (blogger). Online opinion leaders can tailor more stylistic aspects of intervention messages and communicate them according to their readers' needs, experiences, and

XSL•FO
RenderX

preferences. Using established relationships between bloggers and their readers may provide a more organic and collaborative means of tailoring health-related messages compared with interventions where the message tailoring is conducted by the researchers alone (based on target audience member feedback and characteristics). Future studies should compare similar types of message-seeding strategies via established opinion leaders with more traditional message tailoring approaches to further assess their potential variant influence on health outcomes.

In addition, findings revealed that disseminating health messages through opinion leaders may influence the further dissemination and amplification of those messages, meaning that there is potential for rapid diffusion of information via social media interventions. Findings of the *Tracx* analysis indicated that initial mommy blog message posts were shared on a variety of social networking sites, including Facebook, Twitter, Instagram, and other blogs, with Instagram being the most widely used. However, data from the blog reader survey (RQ2) indicated that participants varied widely in terms of self-reported preferences for how they would like to receive environmental breast cancer risk information in the future. Overall, these findings suggest that disseminating health information through blogs may have an extensive reach. Information may be encountered by individuals in one social media platform and then shared with others across other (perhaps more preferred) social media. Our *Tracx* data on message reach showed that some form of the initial seeded intervention messages from the participating bloggers was shared well beyond their initial blog reader audience.

Moreover, this study found that reach, or rapid diffusion of information, extended beyond online communities and platforms to individuals' personal networks. Our data showed that mommy bloggers' posts spurred interaction about breast cancer risk among women and their daughters, specifically. This finding is an important reminder that health information encountered via social media is often shared in close relationships, which may reinforce the impact of the information on health behaviors. This finding is consistent with Rogers' diffusion of innovation theory [14] and associated research that illustrates a 2-step flow of information from media to interpersonal conversations. Future research would benefit from the use of social network analysis to explore the spread of message dissemination among bloggers, their readers, and other online or face-to-face social network members.

The findings also prompt questions about how to best prepare women who use blogs and other social media for conversations with their daughters about environmental breast cancer risk. Communication competence is particularly of concern when mothers talk to their daughters about risk-related topics since these conversations can increase psychological and physiological distress for daughters [3,50]. Other factors, including culture, relational history, and age or maturity, make these conversations even more complex to navigate. Future studies should consider messages that incorporate tailored guidance for engaging in family conversations about breast cancer risk.

Moreover, the findings indicated that the intervention message had a modest impact on key study measures of exposure, satisfaction with the message, intentions to act on the information, and likelihood of sharing the information with others. One intervening variable was participant's recall of the message. Although we know that individuals who recalled the message were more exposed and satisfied with them, and they had stronger intentions to act on the information than the comparison group, it is unknown whether they were recalled from the original blogger post, comments from online or face-to-face social network members, or another social networking site (eg, Facebook and Instagram). Future researchers should take this into account when attempting similar types of seeded messaging strategies. Although researchers can track where information has been reposted or replied to using programs such as *Tracx*, it is often difficult to assess (outside of self-reports) where people encountered specific intervention messages within the social media landscape or how these transfer to interpersonal conversations. In addition, since 42% of women in the comparison group had a family history of breast cancer, it is likely that these individuals had a keen interest in breast cancer prevention and could have been prone to seek this information on blogs (increasing the chance for contamination between the treatment and comparison groups).

The findings also suggest that Instagram was the most popular *other* social media platform for blog readers. This is likely because of Instagram's ability to conveniently share or obtain breast cancer information on a variety of other social media channels (such as Facebook and Twitter). Future research should examine how Instagram (and similar platforms) may enhance the reach and impact of future interventions that attempt to seed similar online opinion leaders with evidence-based breast cancer and environmental risk and prevention messages.

Limitations

There are a number of limitations in this study. First, the study relied on a convenience sample of participants using a quasi-experimental design as opposed to a randomized treatment or control group study. One of the reasons for this was the desire to work with naturally occurring bloggers and their established online social networks of readers. However, the lack of randomized groups limits the generalizability of claims regarding the efficacy of the seeded mommy blogger intervention used in this study.

Additionally, the bloggers and blog readers tended to be highly educated overall. Less than 10% of the intervention group or comparison group members had a high school education or less. The researchers are in the process of analyzing qualitative interview data from a subset of ethnic minority and lower socioeconomic status respondents who participated in this study to assess how intervention messages could be tailored in more culturally sensitive and educationally appropriate ways in future research.

Yet, the findings provide insights into the value of this approach while highlighting consideration for future research. One future direction would be to systematically assess what factors are associated with greater likelihood of recall among participants in this type of intervention. The education level was controlled in this study, but other factors associated with education, such

XSL•FO
RenderX

as health (or electronic health) literacy, could influence message recall. To explore the impact of such factors on participant recall, studies might employ interviews and/or other qualitative methods to see how user-centered design can be integrated with scientifically accurate information to develop more memorable messages.

In terms of other limitations, blog readers typically follow more than one blog, and this might have led to contamination between the intervention and comparison groups. However, we attempted to correct this by asking readers to report up to 5 blogs they visited most frequently in the past year. This allowed us to discover individuals in the comparison group who saw the intervention message. Furthermore, we were limited by the number of survey items we could reasonably include, which contributed to not being able to use larger multidimensional measures of the key variables of interest. In future research, we hope to capture more nuanced aspects of complex variables, such as behavioral intentions to act on environmental breast cancer guidance, as well as finding ways to better distinguish between people who actively disseminate intervention messages and those who are more passive recipients of these messages (ie, lurkers).

Finally, the use of longitudinal research designs is important when assessing health knowledge and behavior change that may result from a social media intervention such as the one reported here. Given that the reader survey was administered over a 2-month period following the start of the mommy blogger intervention, it is difficult to know about the long-term effects of exposure to these types of messages. Future studies need to find creative ways to measure variables longitudinally in an environment that is dynamic in terms of online network membership and the stability of specific social media channels (eg, social media platforms may decline in popularity over time and users may shift to newer platforms). This presents numerous challenges in terms of deciding the best channel(s) for health intervention messages, including those regarding environmental breast cancer risk. Moreover, future researchers should consider assessing mommy bloggers' satisfaction regarding their participation in terms of serving as opinion leaders in similar interventions.

Conclusions

The study results showed promise for the utilization of online opinion leaders such as mommy bloggers as influential channels for the dissemination of both evidence-based and user-generated environmental breast cancer risk messages. Moreover, this study shed light on how blog readers share environmental breast cancer risk and prevention messages across different social media platforms and the preferred channels for receiving this type of information. Results also revealed the potential reach of disseminating health risk information via bloggers. Mommy bloggers may be especially primed to disseminate health risk information geared toward mothers, and this channel may effectively serve as a catalyst for mother-daughter communication about breast cancer risk reduction. Mommy bloggers may also aid interventionists in enhancing the messaging, delivery, and impact of environmental breast cancer risk information.

Acknowledgments

This research was supported by the BCERP Centers grant number R21 ES027417 from the NIEHS and the NCI, NIH, Department of Health and Human Services. This research is solely the responsibility of the authors and does not necessarily represent the official views of the NIEHS or NCI.

Conflicts of Interest

None declared.

References

1. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. CA Cancer J Clin 2017 Nov;67(6):439-448 [FREE Full text] [doi: 10.3322/caac.21412] [Medline: 28972651]
2. Bylund CL, Fisher CL, Brashers D, Edgerson S, Glogowski EA, Boyar SR, et al. Sources of uncertainty about daughters' breast cancer risk that emerge during genetic counseling consultations. J Genet Couns 2012 Apr;21(2):292-304. [doi: 10.1007/s10897-011-9400-y] [Medline: 21833819]
3. Fisher CL, Maloney E, Glogowski E, Hurley K, Edgerson S, Lichtenthal WG, et al. Talking about familial breast cancer risk: topics and strategies to enhance mother-daughter interactions. Qual Health Res 2014 Apr;24(4):517-535. [doi: 10.1177/1049732314524638] [Medline: 24633365]
4. Fisher CL. Coping Together, Side by Sidenriching Mother-Daughter Communication across the Breast Cancer Journey. New York, NY: Hampton Press; May 7, 2014:295.
5. Fisher CL, Wolf BM, Fowler C, Canzona MR. Experiences of "openness" between mothers and daughters during breast cancer: implications for coping and healthy outcomes. Psychooncology 2017 Nov;26(11):1872-1880. [doi: 10.1002/pon.4253] [Medline: 27530810]
6. Sinicrope PS, Brockman TA, Patten CA, Frost MH, Vierkant RA, Petersen LR, et al. Factors associated with breast cancer prevention communication between mothers and daughters. J Womens Health (Larchmt) 2008;17(6):1017-1023 [FREE Full text] [doi: 10.1089/jwh.2007.0497] [Medline: 18554093]

XSL•FO
RenderX

7. Sinicrope PS, Patten CA, Clark LP, Brockman TA, Rock EE, Frost MH, et al. Adult daughters' reports of breast cancer risk reduction and early detection advice received from their mothers: an exploratory study. Psychooncology 2009 Feb;18(2):169-178 [FREE Full text] [doi: 10.1002/pon.1393] [Medline: 18636437]
8. Hiatt RA, Brody JG. Environmental determinants of breast cancer. Annu Rev Public Health 2018 Apr 1;39:113-133. [doi: 10.1146/annurev-publhealth-040617-014101] [Medline: 29328875]
9. Interagency Breast Cancer and Environmental Research Coordinating Committee (IBCERCC). 2013 Feb. Breast Cancer and the Environment Prioritizing Prevention: Summary of Recommendations of the Interagency Breast Cancer and Environmental Research Coordinating Committee (IBCERCC) URL: https://www.niehs.nih.gov/about/assets/docs/summary_of_recs_508.pdf [accessed 2018-12-03] [WebCite Cache ID 74NPk9Qss]
10. Breast Cancer and the Environment Research Program. Breast Cancer and the Environment Research Program URL: https://bcerp.org/ [accessed 2018-09-30] [WebCite Cache ID 72pF0GNTu]
11. Walsh MC, Trentham-Dietz A, Schroepfer TA, Reding DJ, Campbell B, Foote ML, et al. Cancer information sources used by patients to inform and influence treatment decisions. J Health Commun 2010 Jun;15(4):445-463. [doi: 10.1080/10810731003753109] [Medline: 20574881]
12. Kulkarni S, Lewis K, Adams SA, Brandt HM, Lead JR, Ureda JR, et al. A comprehensive analysis of how environmental risks of breast cancer are portrayed on the internet. Am J Health Educ 2018;49(4):222-233 [FREE Full text] [doi: 10.1080/19325037.2018.1473182] [Medline: 30079123]
13. Quinn EM, Corrigan MA, McHugh SM, Murphy D, O'Mullane J, Hill AD, et al. Who's talking about breast cancer? Analysis of daily breast cancer posts on the internet. Breast 2013 Feb;22(1):24-27. [doi: 10.1016/j.breast.2012.05.001] [Medline: 22683246]
14. Rogers EM. Diffusion of Innovations. 5th edition. New York, NY: Free Press; Aug 16, 2003:576.
15. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
16. Silk KJ, Perrault EK, Neuberger L, Rogers A, Atkin C, Barlow J, et al. Translating and testing breast cancer risk reduction messages for mothers of adolescent girls. J Health Commun 2014;19(2):226-243. [doi: 10.1080/10810730.2013.811322] [Medline: 24111511]
17. Quinn EM, Corrigan MA, McHugh SM, Murphy D, O'Mullane J, Hill AD, et al. Breast cancer information on the internet: analysis of accessibility and accuracy. Breast 2012 Aug;21(4):514-517. [doi: 10.1016/j.breast.2012.01.020] [Medline: 22349349]
18. Friedman DB, Toumey C, Porter DE, Hong J, Scott GI, Lead JR. Communicating with the public about environmental health risks: a community-engaged approach to dialogue about metal speciation and toxicity. Environ Int 2015 Jan;74:9-12. [doi: 10.1016/j.envint.2014.09.015] [Medline: 25305416]
19. Burke-Garcia A, Kreps GL, Wright KB. Perceptions about disseminating health information among mommy bloggers: quantitative study. JMIR Res Protoc 2018 Apr 24;7(4):e116 [FREE Full text] [doi: 10.2196/resprot.7764] [Medline: 29691204]
20. Pettigrew S, Archer C, Harrigan P. A thematic analysis of mothers? Motivations for blogging. Matern Child Health J 2016 May;20(5):1025-1031. [doi: 10.1007/s10995-015-1887-7] [Medline: 26662282]
21. Ho J. Functional themes across blogs, wikis and social networking sites relating to pregnancy. Electronic J Health Inform(2) 2011 Jan;6(2):e17 [FREE Full text]
22. 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 Jan;22(1):243-256. [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
23. 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;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
24. Rutten LJ, Squiers L, Hesse B. Cancer-related information seeking: hints from the 2003 Health Information National Trends Survey (HINTS). J Health Commun 2006;11(Suppl 1):147-156. [doi: 10.1080/10810730600637574] [Medline: 16641080]
25. Fox S, Duggan M. Pew Research Center Internet & Technology. 2013 Jan 15. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [WebCite Cache ID 72pFbwzGt]
26. Cobb NK, Poirier J. Effectiveness of a multimodal online well-being intervention: a randomized controlled trial. Am J Prev Med 2014 Jan;46(1):41-48 [FREE Full text] [doi: 10.1016/j.amepre.2013.08.018] [Medline: 24355670]
27. Fox S. Pew Research Center Internet & Technology. 2011 May 12. The Social Life of Health Information, 2011 URL: http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/ [WebCite Cache ID 72pFlpy4G]
28. Keating DM, Rains SA. Health blogging and social support: a 3-year panel study. J Health Commun 2015;20(12):1449-1457. [doi: 10.1080/10810730.2015.1033119] [Medline: 26332404]
29. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
30. Granovetter MS. The strength of weak ties. Am J Sociol 1973 May;78(6):1360-1380 [FREE Full text]

XSL•FO
RenderX

31. Walther JB. Relational aspects of computer-mediated communication: experimental observations over time. Organ Sci 1995 Apr;6(2):186-203. [doi: 10.1287/orsc.6.2.186]
32. Wellman B, Salaff J, Dimitrova D, Garton L, Gulia M, Haythornthwaite C. Computer networks as social networks: Collaborative work, telework, and virtual community. Annu Rev Sociol 1996 Aug;22(1):213-238. [doi: 10.1146/annurev.soc.22.1.213]
33. Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. Commun Res 2009 Nov 25;37(1):105-132. [doi: 10.1177/0093650209351512]
34. Rains SA. Language style matching as a predictor of perceived social support in computer-mediated interaction among individuals coping with illness. Commun Res 2015 Jan 13;43(5):694-712. [doi: 10.1177/0093650214565920]
35. Lu AS. An experimental test of the persuasive effect of source similarity in narrative and nonnarrative health blogs. J Med Internet Res 2013;15(7):e142 [FREE Full text] [doi: 10.2196/jmir.2386] [Medline: 23887135]
36. McQueen A, Kreuter MW, Kalesan B, Alcaraz KI. Understanding narrative effects: the impact of breast cancer survivor stories on message processing, attitudes, and beliefs among African American women. Health Psychol 2011 Nov;30(6):674-682 [FREE Full text] [doi: 10.1037/a0025395] [Medline: 21895370]
37. eMarketer. 2010 Nov 22. Understanding How New Moms Share URL: https://www.emarketer.com/Article/Understanding-How-New-Moms-Share/1008056 [accessed 2018-09-30] [WebCite Cache ID 72pG3GOOD]
38. Lupton D, Pedersen S, Thomas GM. Parenting and digital media: from the early web to contemporary digital society. Sociol Compass 2016 Aug 2;10(8):730-743. [doi: 10.1111/soc4.12398]
39. Morrison A. Autobiography in real time: a genre analysis of personal mommy blogging. Cyberpsychology 2010 Jan;4(2):- [FREE Full text]
40. Nardi BA, Schiano DJ, Gumbrecht M. Blogging as social activity, or, would you let 900 million people read your diary? Proceedings of the 2004 Association for Computing Machinery conference on Computer Supported Cooperative Work; 2004 Nov. In: Proceeding CSCW '04 Proceedings of the 2004 ACM conference on Computer supported cooperative work. 2004 Nov Presented at: Association for Computing Machinery conference on Computer Supported Cooperative Work; November 6-10, 2004; Chicago, Illinois, USA p. 222-231. [doi: 10.1145/1031607.1031643]
41. Fox S, Lenhart A. Pew Research Center Internet & Technology. 2006 Jul 19. Bloggers URL: http://www.pewinternet.org/2006/07/19/bloggers/ [WebCite Cache ID 72pHpWyDQ]
42. Lopez LK. The radical act of 'mommy blogging': redefining motherhood through the blogosphere. New Media Soc 2009 Jul 21;11(5):729-747. [doi: 10.1177/1461444809105349]
43. Petersen EJ. Mommy bloggers as rebels and community builders. J Motherhood Initiative Res Commun Involvement 2015;6(1):1-22 [FREE Full text]
44. 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]
45. Chang YT, Yu H, Lu HP. Persuasive messages, popularity cohesion, and message diffusion in social media marketing. J Bus Res 2015 Apr;68(4):777-782. [doi: 10.1016/j.jbusres.2014.11.027]
46. Strekalova YA, Krieger JL. Beyond words: amplification of cancer risk communication on social media. J Health Commun 2017 Oct;22(10):849-857. [doi: 10.1080/10810730.2017.1367336] [Medline: 28956723]
47. Bissonnette-Maheux V, Dumas AA, Provencher V, Lapointe A, Dugrenier M, Straus S, et al. Women's perceptions of usefulness and ease of use of four healthy eating blog characteristics: a qualitative study of 33 French-Canadian women. J Acad Nutr Diet 2017 Jul;118(7):1220-1227. [doi: 10.1016/j.jand.2017.08.012] [Medline: 29107587]
48. Fisher CL, Roccotagliata T, Rising CJ, Kissane DW, Glogowski EA, Bylund CL. "I Don't Want to Be an Ostrich": managing mothers' uncertainty during BRCA1/2 genetic counseling. J Genet Couns 2017 Jun;26(3):455-468. [doi: 10.1007/s10897-016-9998-x] [Medline: 27473644]
49. Fisher CL. Coping with breast cancer across adulthood: emotional support communication in the mother–daughter bond. J Appl Commun Res 2010 Nov;38(4):386-411. [doi: 10.1080/00909882.2010.513996]
50. Berlin KL, Andreotti C, Yull F, Grau AM, Compas BE. Mother-daughter communication about breast cancer risk: interpersonal and biological stress processes. J Behav Med 2013 Jun;36(3):328-339. [doi: 10.1007/s10865-012-9418-0] [Medline: 22569774]
51. Perrault EK, Silk KJ. Testing the effects of the addition of videos to a website promoting environmental breast cancer risk reduction practices: are videos worth it? J Appl Commun Res 2014 Jan 1;42(1):20-40 [FREE Full text] [doi: 10.1080/00909882.2013.854400] [Medline: 25143661]
52. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012 May 31;7:50 [FREE Full text] [doi: 10.1186/1748-5908-7-50] [Medline: 22651257]
53. Fishbein M. A reasoned action approach to health promotion. Med Decis Making 2008 Nov;28(6):834-844 [FREE Full text] [doi: 10.1177/0272989X08326092] [Medline: 19015289]
54. Fishbein M, Cappella JN. The role of theory in developing effective health communications. J Commun 2006 Aug;56(s1):S1-S17. [doi: 10.1111/j.1460-2466.2006.00280.x]
55. Greenwood S, Perrin A, Duggan M. Pew Research Center Internet & Technology. 2016 Nov 11. Social Media Update 2016 URL: http://www.pewinternet.org/2016/11/11/social-media-update-2016/ [WebCite Cache ID 72pI58afA]

XSL•FO
RenderX

56. Kratzke C, Wilson S, Vilchis H. Reaching rural women: breast cancer prevention information seeking behaviors and interest in internet, cell phone, and text use. J Community Health 2013 Feb;38(1):54-61. [doi: 10.1007/s10900-012-9579-3] [Medline: 22706805]
57. The Motherhood. Influencer Marketing, Handcrafted URL: https://www.themotherhood.com [accessed 2018-12-07] [WebCite Cache ID 74UQ53mzx]
58. Breast Cancer and the Environment Research Program. Your Daughter and Breast Cancer: Reducing Her Risk Now URL: https://bcerp.org/wp-content/uploads/2016/10/1386_NIEHS_BCERP_Mother_Toolkit_Final_PC.pdf [accessed 2018-09-30] [WebCite Cache ID 72pIiMV2x]
59. Tracx. Navigate the social jungle with Tracx URL: https://www.tracx.com [accessed 2018-12-03] [WebCite Cache ID 74NPNqG4h]

Abbreviations

ANCOVA: analysis of covariance
BCERP: Breast Cancer and the Environment Research Program
BPA: bisphenol A
H1: first hypothesis
H2: second hypothesis
IBCERCC: Interagency Breast Cancer and Environmental Research Coordinating Committee
NCI: National Cancer Institute
NIEHS: National Institute of Environmental Health Sciences
NIH: National Institutes of Health
RQ: research question

Edited by G Eysenbach; submitted 07.10.18; peer-reviewed by AA Dumas, T Ginossar, S Santarossa; comments to author 15.11.18; revised version received 11.12.18; accepted 14.12.18; published 07.03.19.

Please cite as:
Wright K, Fisher C, Rising C, Burke-Garcia A, Afanaseva D, Cai X
Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention
J Med Internet Res 2019;21(3):e12441
URL: https://www.jmir.org/2019/3/e12441/
doi:10.2196/12441
PMID:30843866

©Kevin Wright, Carla Fisher, Camella Rising, Amelia Burke-Garcia, Dasha Afanaseva, Xiaomei Cai. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Characterizing Media Content and Effects of Organ Donation on a Social Media Platform: Content Analysis

Xiaoya Jiang[1], MA; Wenshi Jiang[2], MS; Jiawei Cai[2], BSc; Qingdong Su[3], MS; Zhigang Zhou[4], MD; Lingnan He[1,5], PhD; Kaisheng Lai[6], PhD

[1]School of Communication and Design, Sun Yat-Sen University, Guangzhou, China

[2]Intelligence Sharing for Life Science Research Institute, Shenzhen, China

[3]923 Hospital of Chinese People's Liberation Army, Nanning, China

[4]The First People's Hospital of Kunming, Kunming, China

[5]Guangdong Key Laboratory for Big Data Analysis and Simulation of Public Opinion, Guangzhou, China

[6]School of Journalism and Communication, Jinan University, Guangzhou, China

Corresponding Author:
Kaisheng Lai, PhD
School of Journalism and Communication
Jinan University
Number 601 Whampoa Avenue West
Guangzhou, 510632
China
Phone: 86 020 38374980
Email: kaishenglai@126.com

Abstract

Background: The lack of organ donors has become a barrier for the development of organ transplantation programs, and many countries are currently facing a severe shortage of deceased organs. Media campaigns on social media have the potential to promote organ donation. However, little is known about what kind of media content is the most appropriate for this purpose.

Objective: This study aimed to analyze media posts regarding organ donation on Weibo, a social media platform, and to identify the media themes that are most advantageous in promoting public awareness and attitudes concerning organ donation.

Methods: Based on 16 million social media users' posts randomly extracted from January 1 to December 31, 2017, 1507 reposts of 141 distinct media posts relevant to organ donation were found. We analyzed the media posts' themes and examined their effects in promoting public awareness about organ donation by comparing the number of reposts and comments they prompted. The themes' impact on attitude toward organ donation was gauged using the comments indicating support and intentions for organ donation.

Results: Overall, 5 major themes were identified from the media posts, among which "organ donation behaviors" constituted the highest proportion (58/141, 41.13%). However, themes of "statistical descriptions of organ donation" and "meaningfulness of donation" were the most influential in promoting awareness on organ donation: approximately 3 of 10 commenters for the former theme and 2 of 10 commenters for the latter expressed intentions to become organ donors. These two themes, along with "meaningfulness of organ donation for society," a subtheme of "meaningfulness of donation," were the most effective for evoking support and intentions for donation.

Conclusions: A discrepancy was revealed between the media themes that were the most salient on the media agenda and those that were the most effective in increasing organ donation awareness and intentions on social media. These findings provide guidance for campaigns on organ donation. The results also suggest the potential of campaigns on social media for promoting prosocial health behaviors and highlight the importance of strategic message design for serving this goal.

(J Med Internet Res 2019;21(3):e13058) doi:10.2196/13058

KEYWORDS

organ donation; media; media effects; social media

XSL•FO
RenderX

Introduction

Background

Organ transplantation has brought hope to people with illness previously considered incurable [1]. However, the donated organ is a prerequisite of organ transplantation, and the need for organ donation has increased globally in the past years [2]. In China, for every 30 patients requiring organ transplantation, only one receives the organ [3]. Given the alarming extent of shortage of organ donation, measures have been actively taken to promote donation among citizens through policy making and simplification of procedure. In 2013, the National Health Commission of China announced the Provisions on the Human Organ Procurement and Allocation to provide guidance on organ donations [4]. Three years later, a fast channel for organ donation registration was also provided on a major e-commerce platform, in which citizens could become donor volunteers with just a few clicks [5]. Despite these efforts, the number of available organs is still below the requisite. Medical officials and professionals indicated that the plausible reasons for this shortage include the public's insufficient awareness of and motivation for organ donation [3].

Role of Social Media in Organ Donation

Social media plays an indispensable role in the public health landscape [6] and has the potential to leverage public engagement for organ donation. As many people are becoming accustomed to assessing health-related information online [6], social media has been widely used for communicating health information [7]. Individuals log on to social media sites to interact with each other for a variety of health issues, ranging from cancer [8] and e-cigarettes [9] to the role of virtual reality in health care [10]. Regarding organ donation, social media has also been broadly adopted for communication at the hospital, community, and grassroot levels [11].

Despite functioning as a venue for information exchange among peers, social media is a platform for promoting health regimens [6]. Information exchange usually occurs via a top-down approach and through campaigns. For example, the state health departments in the United States actively disseminate health information to citizens via Facebook and Twitter [12]. In addition, the media plays an indelible role in leveraging health topics in the social media arena. The media has been providing information on health topics such as lung cancer treatment [8] and antibiotic usage [13] on social media. The media has also been used for educational and promotional purposes for various health-related issues [2], including tobacco use [14], alcohol consumption [15], and HIV prevention [16]. With many news institutions using social media as a platform for disseminating their reports [17] and numerous news consumers habitually assessing media contents via social media [18], media campaigns on social media may play an increasingly important part in promoting public health regimens.

Effects of Organ Donation in the Media

In the context of organ donation, the public tends to resort to the media and social media sites for information, indicating the plausibility of promoting this health behavior via media campaigns on social media. A thematic analysis of family dyads' discussion on organ donation exemplified this predilection by revealing that the media was the most important source of information for decision making in family dyads [19]. Previous research regarding media effects also indicates that exposure to media contents may shape people's perceptions of organ donation. For example, agenda-setting studies have shown that the media could leverage the public's attention to the issues it covers [20]; research on framing effects has shown evidence of media's impact on people's attitudes [21]. Additionally, empirical studies revealed that the public would be particularly susceptible to media influence when their relevant knowledge level was low [22], which suggests that the media may exert a particularly pronounced influence on the citizens' attitudes toward organ donation, as their understanding of organ donation is currently in a nascent state [3]. Several existing media campaigns' effects have indeed been promising. For instance, the news reports about World Transplant Games caused noticeable surges in organ donation in many regions [23], media coverage of cornea donation boosted donation intentions in Korea [24], and media campaigns elevated eye donation awareness in India and Ethiopia [25,26].

Besides the media's active role in promoting organ donation, the social media is an important arena for this purpose; for instance, health educators have been communicating about organ transplantation on social media sites [27]. Previous research suggested that several characteristics of social media makes it advantageous for media campaigns. Not constrained by traditional media readership of certain demographic groups, social media allows for dissemination of information to numerous citizens with various demographic characteristics and geographical locations [11,28]. Furthermore, when exposed to social media contents, people are more inclined to conduct additional information search [29,30] and express sentiments [9], which may further amplify the media campaigns' impacts. In sum, the media effects and wide usage of social media suggest the potential of media campaigns on social media sites.

Despite the plausibility, the media's actual effects of organ donation campaigns on social media sites remain unknown. Additionally, effective media campaigns require careful selection of media contents; however, existing knowledge about what kind of media posts are most suitable for organ donation is inadequate. Organ donation cannot be increased by merely distributing media posts, as the effect of persuasive messages may not always meet the message designers' expectations, and some messages might even induce backfire effects [31]. For example, research regarding health messages have cautioned that the presence of smoking cues in antismoking messages would weaken their persuasive effects [32] and that the success of cancer campaigns is largely contingent on the nature of the campaigns [30]. Effectiveness of media campaigns varies by the type of issues included and the specific campaign strategies employed, and existing media campaigns for organ donation have, in general, yielded mixed results, despite a few successful cases [2].

Therefore, careful selection of media contents is crucial for organ donation campaigns, and identifying the most appropriate media contents is a critical step for this purpose. Previous studies

XSL•FO
RenderX

have yielded informative findings; for example, gain rather than loss frames [33] and presentation of identifying information of organ recipients instead of donors [34] would induce more favorable reactions to organ donation. However, media posts in a natural environment are complex and nuanced, which are beyond the scope of the frameworks of these existing studies. To address this concern, media posts on organ donation on social media could provide valuable resources. An analysis of the media contents may help in profiling the existing media themes, and the themes' effects could be assessed by approaches such as examining the repost frequencies [35]. Additionally, as the media posts and users' reactions occur naturally, they could serve as a simulation of actual media campaigns. Spotting media themes that are most motivating for organ donation among the existing posts is also beneficial, as it allows for effective media campaigns while maintaining the existing journalistic practices.

Research Questions

This study addressed two research questions:

1. What are the themes of the media contents regarding organ donation on social media?
2. What effects do different media themes have on people's attitudes toward organ donation, and which themes are the most effective in promoting organ donation?

Methods

Data Collection

A total of 16 million Weibo users' posts were randomly extracted from January 1 to December 31, 2017, through Weibo's application programming interface. Weibo is the Chinese equivalent of Twitter, with an enormous amount of media posts and 37.6 million monthly active users in 2017 [36]. As such, it is a good source for observing the diffusion and effects of media posts on social media. From this dataset, 7046 posts were extracted via key word searches including combinations of "organ," "donation," and "shortage," (eg, organ donation and organ shortage) and the names of particular types of organs (eg, organ, liver, and liver donation). The posts were further manually filtered to include only reposts that contain both the content of the posts and a distinguishable media source. This procedure yielded 1507 reposts, which served as the dataset for further investigation.

Data Analysis

Primarily, we identified major themes in the media posts. The 1507 reposts were attributed to 141 distinct media posts. In line with previous research, a direct content-analysis approach was adopted to analyze these posts [37]. One researcher screened the posts and proposed a codebook for categorizing them based on media themes. Another researcher subsequently evaluated the categorizations and provided suggestions. Consensus between the two researchers was reached for all themes. Following the procedure used by a previous study [37], the two researchers then coded the first 30 pieces of the media posts, which constitutes approximately 20% of the total. Disparities in coding were discussed and resolved. If one post pertained to two themes, it was labeled according to the more salient one. Lastly, one of the researchers coded the remining posts. Media themes of the posts were identified, and all the media posts were classified according to their themes.

When browsing media posts, social media users can repost with a click and sometimes add comments along with the reposts. Adopting the methodology of a previous study [35], we used the reposting frequency of a post and the number of comments it received as proxy measures of issue awareness promoted by the post. The numbers of posts and comments received for each media theme were counted for this purpose. To gauge a particular theme's contribution to organ donation awareness, the repost/post ratio and comment/post ratio were calculated by dividing the total number of reposts and comments pertaining to each theme by the total number of posts. For example, if 58 posts regarding "organ donation behavior" gained 295 reposts and 37 comments, the repost/post ratio would be 295/58, indicating that every post evoked approximately 5 reposts on an average, and the comment/post ratio would be 37/58, suggesting that every post induced approximately 0.64 comments.

In addition, the extent to which a media theme motivated organ donation was assessed based on the number of its comments that expressed prodonation attitudes and stated donation intentions. Using a similar procedure for the media content analysis, the comments that indicated support and intention for organ donation were coded and counted for each theme. The number of comments pertaining to a theme that showed prodonation attitudes and intentions was divided by the total number of comments pertaining to the theme, yielding a prodonation/comment ratio and intention/comment ratio, respectively. For instance, the media posts under the theme of "organ donation behavior" received 37 comments, of which 7 contained prodonation attitudes and 3 showed donation intentions, which makes the prodonation/comment ratio 7/37 and the donation/comment ratio 3/37. The results indicate that 18.92% comments pertaining to this theme expressed support for organ donation and 8.11% showed donation intentions. Finally, the media themes' effectiveness in promoting organ donation was assessed and compared based on these indicators.

Results

Media Contents

Five major themes were derived from the 141 media posts: organ donation behaviors, issues and policies regarding organ donation, meaningfulness of organ donation, statistical descriptions of organ donation, and organ donation practice (Table 1).

XSL•FO
RenderX

Table 1. Coded themes or subthemes of media posts for organ donation.

Theme and subtheme	Number of posts	Definition	Example
Organ donation behaviors	58	The text pertains to individual donors' or their families' behavior related to donation or the donation process	A father donated his beloved son's organ after the son's death in a car accident
Meaningfulness of organ donation	28	The text pertains to the meaningfulness generated by organ donation	A mother heard her deceased son's heartbeat when she met her son's organ recipient
Meaningfulness for donors	10	The text pertains to the meaningfulness of organ donations for the donors	A mother heard her deceased son's heartbeat when she met her son's organ recipient
Meaningfulness for recipients	4	The text pertains to the meaningfulness of organ donations for the recipients	An organ recipient was grateful to the donor's families and called them "mom and dad"
Meaningfulness for society	14	The text pertains to the meaningfulness of organ donations for society	Organ donation extends our lives
Issues and policies regarding organ donation	28	The text pertains to the state and the problems facing organ donation or policies that address these problems	The "organ donation coordinator" was employed to help alleviate the organ donation shortage
Statistical descriptions of organ donation	21	The text pertains to the statistics of registered organ donors	In 2016, the number of registered organ donation volunteers reached 104,528 in China
Organ donation practice	6	The text pertains to the implementation and procedure of organ donation	A well-known medical professional performed an organ donation transplantation for a patient while on a visit to Beijing for a major conference

Among them, the largest proportion of the posts addressed "organ donation behaviors" (58/141, 41.13%). This theme refers to an individual donor or his family's engagement with organ donation (eg, "The father donated his beloved son's organ after the son passed away in a car accident"). "Issues and policies regarding organ donation" and "meaningfulness of organ donation," each constituted 19.86% (28) of the posts. Upon close analysis, we found that three subthemes of "meaningfulness of organ donation" were observed, based on the subjects involved in organ donation: the meaningfulness of donation for the donors, recipients, and society. The remaining two themes pertained to "statistical descriptions of organ donation" (21, 14.89%) and "organ donation practice" (6, 4.26%).

Sources of the media posts were also identified and classified according to three commonly recognized media source categories in China: official media, which was responsible to the government institutions; market-oriented media, which is largely influenced by market competition; and self-media, which is initiated by individual persons or organizations [38,39]. The analysis revealed that the three types of media, in general, have taken equally active roles in disseminating organ donation information. Specifically, market-oriented media produced 34.75% of the posts, official media had a marginally lower proportion of 34.04%, and self-media produced 31.20% of the total.

Media Effects

Of the five themes, "statistical descriptions of organ donation" had the highest repost/post ratio of 37.24%, indicating that the media posts in this theme triggered 37.24 reposts on an average, followed by "meaningfulness of organ donation," with a repost/post ratio of 14.04%. It is also noteworthy that among its three subthemes, "meaningfulness for recipients" induced a high repost/post ratio of 27.75%. By contrast, the ratios for "organ donation practice" and "issues and policies regarding organ donation" ranked low at 1.33% and 1.04%, respectively. Regarding the comment/post ratio, "statistical descriptions of organ donation" had a value of 7.71%, which was the highest, and "organ donation behaviors" had the lowest ratio of 0.64% (Table 2).

XSL•FO
RenderX

Table 2. Posts, reposts, and comments across media themes.

Media themes and subthemes	Number of posts	Number of reposts	Number of comments	Repost/post ratio (%)	Comment/post ratio (%)
Organ donation behaviors	58	295	37	5.09	0.64
Meaningfulness of organ donation	28	393	42	14.04	1.5
Meaningfulness for donors	10	198	24	19.80	2.4
Meaningfulness for recipients	4	111	11	27.75	2.75
Meaningfulness for society	14	84	7	6.00	0.5
Issues and policies regarding organ donation	28	29	4	1.04	0.14
Statistical descriptions of organ donation	21	782	162	37.24	7.71
Organ donation practice	6	8	0	1.33	0

Table 3. Prodonation attitudes and donation intentions expressed across media themes.

Media themes and subthemes	Number of prodonations	Prodonation/comment ratio (%)	Number of donation intentions	Donation/comment ratio (%)
Organ donation behaviors	7	18.92	3	8.11
Meaningfulness of organ donation	16	38.10	9	21.43
Meaningfulness for donors	7	29.17	4	16.67
Meaningfulness for recipients	4	36.36	1	9.10
Meaningfulness for society	5	71.43	4	57.14
Issues and policies regarding organ donation	1	25.00	0	0.00
Statistical descriptions of organ donation	85	52.47	50	30.86
Organ donation practice	N/A[a]	N/A	N/A	N/A

[a]N/A: not applicable.

Among the 1507 reposts, 245 contained comments, of which 109 expressed prodonation attitudes and 62 comments contained clear donation intentions. These social media users showed positive attitudes toward organ donations with both straightforward claims such as “I support organ donation” and more nuanced tones such as “It shows the advancement of our society.” There were typical comments regarding organ donation intentions, such as “If I meet the requirements, I would like to donate my organs.” The media themes were compared based on the two indicators for prodonation attitudes and donation intentions. The analysis showed that “statistical descriptions of organ donation” generated the highest prodonation/comment ratio of 52.47% and intention/comment ratio of 30.86%, suggesting that 52.47% of the comments under this theme contained prodonation sentiments and 30.86% included donation intentions. “Meaningfulness of donation” had the second highest prodonation/comment ratio of 38.10% and a donation/comment ratio of 21.43%. In addition, one subtheme of “meaningfulness of donation”—“meaningfulness for society”—had a prodonation/comment ratio of 71.43% and a donation/comment ratio of 57.14%. In comparison, organ donation behaviors had both the lowest prodonation/comment ratio of 18.92% and the lowest donation/comment ratio of 8.11% (Table 3).

The effectiveness of campaigns would be further enhanced if the target audience could be profiled. Specifically, it would be informative to acknowledge people with characteristics that are more likely to transmit the organ donation posts and who are more inclined to develop prodonation attitudes after media exposure. Therefore, we took a further step to characterize the people who reposted, commenters, donation supporters, and self-reported donor volunteers by retrieving and analyzing information on their sex and age. The social media users’ characteristics in our sample were compared with those of the original 16 million social media users to ensure that the observed difference in the dataset of organ donation (ie, more young people supported organ donation than the old) was not derived from overrepresentation or underrepresentation of certain groups. The analysis showed that female individuals were more likely to comment about organ donation than male individuals and that the sex distribution of commenters in the datasets of organ donation and the dataset of 16 million users was significantly different (χ^2_1=23.0; P<.001). Moreover, individuals younger than 30 years of age were found to be more likely to repost after controlling for the age distribution in the original dataset (χ^2_1=54.0; P<.001). For donation intentions, 62 people’s sex was identifiable, and among them, more female (n=45) than male individuals (n=17) expressed willingness to become donors. Subsequently, we examined the characteristics of self-proclaimed donors via similar approaches. The analysis suggested that young people (aged less than 30 years) were more likely to donate. However, these differences were not statistically significant after controlling for the sex and age distributions in the original dataset.

XSL•FO
RenderX

Discussion

Principal Results

With social media and media, in general, taking increasingly active roles in the public health arena [2,6], this study empirically analyzed the effects of media for organ donation on a social media platform. The analysis identified media themes that were most effective in promoting organ donation and revealed a discrepancy between the themes that received the most media spotlight (ie, "organ donation behaviors") and themes that were most impactful for public awareness and attitudes of organ donation (ie, "statistical descriptions of organ donation" and "meaningfulness of organ donation"). As organ donation is emerging as a concern in many countries such as China [3], the United States [34], and India [25], these findings provide practical guidance for promoting organ donation via campaigns on social media. Furthermore, this study highlighted the importance of strategic message design in promoting health regimens via campaigns on social media and explored methodological approaches to quantify the media effects by using the existing online behavioral data.

With respect to evoking the public's donation intentions, "statistical descriptions of organ donation" was found to have a dominant advantage: 3 of 10 commenters were motivated to become donors. "Meaningfulness of organ donation" and one of its subthemes, "meaningfulness for society," also generated a significant amount of donation intentions. These two major themes were observed in the other aspects as well (eg, prodonation attitudes). Their particular effectiveness in provoking donation intentions may stem from social norms, which could be promoted by the media, and influence people's behavioral intentions [40,41]. The media posts of "statistical description of organ donation" might have activated a descriptive social norm, which pertains to commonly adopted behaviors [41]. In addition, "meaningfulness for society" might be connected to injunctive norms, which concerns behaviors that people think they ought to exhibit [41,42]. The notable effectiveness of "meaningfulness for society" could also be explained: As potential organ donation supporters tend to be prosocial [43], addressing social meaningfulness would further activate their motivations to donate.

In sharp contrast with these two themes' impacts, "organ donation behaviors" had the weakest effect on promoting organ donation intentions. This result is consistent with previous experimental findings: As compared to identifying information regarding the organ recipient, presenting information of the deceased donors helped less in motivating people to register for donations, as it may induce thoughts of death rather than saving lives [34]. This study's findings lent more empirical support to this tendency and cautioned against media's overemphasis on the group of organ donors, which is a disadvantage not only when compared to media themes regarding organ recipients, but also when evaluated across the other themes.

Combining the findings regarding the salience and effects of the media themes, a gap was observed between the media themes that received the most media attention and those that were the most advantageous in promoting organ donation. This discrepancy indicates an unrealized potential for the media to increase organ donations; media campaigns might fulfill this goal by emphasizing donation trends and meaningfulness while curtailing the number of reports on ordinary individuals' donation behaviors. Furthermore, the themes that are advantageous in promoting organ donation awareness, attitudes, and intentions also provide templates for message designs for future media campaigns.

An exploratory analysis was also conducted to characterize social media users who were actively involved in the issue of organ donation after exposure to the relevant media contents. The analysis of characteristics of commenters revealed patterns including one showing that female individuals more actively commented on organ donation than male individuals. This finding is consistent with previous insights that female individuals had higher level of engagement with social media [10]. No significant difference based on sex and age was observed for the group of potential donors after considering sex and age distributions in the original dataset. This might be due to the limited number of commenters who expressed willingness to donate. Future studies may further explore the characteristics of such critical groups with larger datasets.

Overall, this study examined the media effects of a health issue on social media. Previous researchers pointed out that social media has become an increasingly influential platform for health-related campaigns and a rich mine for research on health-related behaviors [6]. Echoing with these proposals, studies have analyzed social media posts' role in health-related issues such as cancer awareness [7,30] and e-cigarette flavors [35]. Building on previous insights, this study further explored social media in the context of organ donation. Although many existing studies addressed peer interactions and public discourses regarding health issues on social media [10,35] or investigated media discourses without referring to specific platforms [44,45], this study examined the media representation of organ donation on a social media site. The analysis revealed major media themes of an important health issue and yielded practical implications for future media campaigns. Moreover, this study assessed social media posts' influence in additional aspects. Previous research has modeled the information propagation of social media posts based on the frequency of reposts [35], whereas this study went further by incorporating the posts' influence on attitudes and behavioral intentions. Admittedly, the operationalizations employed are not necessarily perfect; nonetheless, this approach was developed based on previous research frameworks [35]. In addition, it serves as an initial step that may inspire further methodological enhancement and suggests a route for creating more comprehensive and informative indicators for the effects on social media. Lastly, this study undertook a step to characterize the target audience for future campaigns, in response to previous researchers' suggestions for determining the critical group [7,46]. Subsequent studies may profile the target audience using a larger sample and additional perspectives. Besides the demographic characteristics, the social media users' preferred topics and previously expressed sentiments may also serve as aspects of their profile, and the abundant data on social media allow for investigation on this domain.

XSL•FO
RenderX

Limitations

The quantity of the data for this analysis was not large, partly due to general insufficient awareness about organ donation [3]. Future research may track the media contents and their effects as the organ donation issue gains more importance in the media and public agenda. In addition, this study was conducted in the context of China; future studies are needed to examine the themes' effects in other geographical and cultural contexts. This study also found it difficult to establish a causal relationship in a natural environment, and it is possible that a social media user's expressed endorsement of organ donation originated from sources other than media exposure. Nevertheless, this bias might not be pronounced, as it is likely to exist for each of the media themes, and the conclusions were based on the themes' relative effects rather than absolute effects. Another limitation is that not all people who reposted and developed donation intentions after the media exposure would write comments to express their thoughts; thus, the donation/comment ratio might be an underestimated indicator. On the other hand, expressing intent to donate might not be equal to the act of registering for donation, which may produce overestimation in the observation. Nevertheless, previous empirical studies have shown that behavioral intention is associated with actual behaviors [47]. In addition, assessing people's organ donation attitudes via comments have certain advantages; compared to other research approaches like surveys, comments are helpful in avoiding socially desirable answers. For example, in a survey conducted in 2015, 89.9% respondents expressed pro-organ donation attitudes and 42.2% expressed donation intention [48]. Future studies should develop methods that allow for gauging donation intentions more accurately or comparing and quantifying the biases pertinent to different research approaches.

Acknowledgments

This work was partially supported by the Major Project of the National Social Science Foundation of China (Award #18ZDA165).

Conflicts of Interest

None declared.

References

1. Linden PK. History of solid organ transplantation and organ donation. Crit Care Clin 2009 Jan;25(1):165-184. [doi: 10.1016/j.ccc.2008.12.001] [Medline: 19268801]
2. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. Lancet 2010 Oct 9;376(9748):1261-1271 [FREE Full text] [doi: 10.1016/S0140-6736(10)60809-4] [Medline: 20933263]
3. The Xinhua News Agency. The gift of life? The improvements and the concerns behind the increased organ donation rates [in Chinese] URL: http://www.xinhuanet.com/politics/2017-03/31/c_1120735285.htm [accessed 2019-03-06] [WebCite Cache ID 76fUME6VO]
4. The Beijing News. Provision on the administration of human organ procurement and allocation: clarifying provision of organs [in Chinese] URL: http://www.bjnews.com.cn/news/2019/01/28/543583.html [accessed 2019-03-06] [WebCite Cache ID 76fdEccXJ]
5. Health Chengdu. Becoming organ donation volunteer on Zhifubao with one click, are you going to volunteer? [in Chinese] URL: http://www.sohu.com/a/122856037_456094 [accessed 2019-03-06] [WebCite Cache ID 76fVEIZ7A]
6. Centola D. Social media and the science of health behavior. Circulation 2013 May 28;127(21):2135-2144 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.101816] [Medline: 23716382]
7. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: examples of cancer awareness via Twitter. JMIR Public Health Surveill 2016;2(1):e17 [FREE Full text] [doi: 10.2196/publichealth.5205] [Medline: 27227152]
8. Sutton J, Vos SC, Olson MK, Woods C, Cohen E, Gibson CB, et al. Lung cancer messages on Twitter: content analysis and evaluation. J Am Coll Radiol 2018 Jan;15(1 Pt B):210-217. [doi: 10.1016/j.jacr.2017.09.043] [Medline: 29154103]
9. Martinez LS, Hughes S, Walsh-Buhi ER, Tsou M. "Okay, We Get It. You Vape": an analysis of geocoded content, context, and sentiment regarding E-cigarettes on Twitter. J Health Commun 2018 Jul 06;23(6):550-562. [doi: 10.1080/10810730.2018.1493057] [Medline: 29979920]
10. Keller MS, Park HJ, Cunningham ME, Fouladian JE, Chen M, Spiegel BMR. Public perceptions regarding use of Virtual Reality in health care: a social media content analysis using Facebook. J Med Internet Res 2017 Dec 19;19(12):e419 [FREE Full text] [doi: 10.2196/jmir.7467] [Medline: 29258975]
11. Henderson ML, Clayville KA, Fisher JS, Kuntz KK, Mysel H, Purnell TS, et al. Social media and organ donation: Ethically navigating the next frontier. Am J Transplant 2017 Nov;17(11):2803-2809 [FREE Full text] [doi: 10.1111/ajt.14444] [Medline: 28744966]
12. Harris JK, Mueller NL, Snider D. Social media adoption in local health departments nationwide. Am J Public Health 2013 Sep;103(9):1700-1707. [doi: 10.2105/AJPH.2012.301166] [Medline: 23865660]
13. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. Am J Infect Control 2010 Apr;38(3):182-188 [FREE Full text] [doi: 10.1016/j.ajic.2009.11.004] [Medline: 20347636]

XSL•FO
RenderX

14. Wakefield MA, Durkin S, Spittal MJ, Siahpush M, Scollo M, Simpson JA, et al. Impact of tobacco control policies and mass media campaigns on monthly adult smoking prevalence. Am J Public Health 2008 Aug;98(8):1443-1450. [doi: 10.2105/AJPH.2007.128991] [Medline: 18556601]
15. Anderson P, Chisholm D, Fuhr DC. Effectiveness and cost-effectiveness of policies and programmes to reduce the harm caused by alcohol. Lancet 2009 Jun 27;373(9682):2234-2246. [doi: 10.1016/S0140-6736(09)60744-3] [Medline: 19560605]
16. Noar SM, Palmgreen P, Chabot M, Dobransky N, Zimmerman RS. A 10-year systematic review of HIV/AIDS mass communication campaigns: Have we made progress? J Health Commun 2009 Feb;14(1):15-42. [doi: 10.1080/10810730802592239] [Medline: 19180369]
17. Hong S. Online news on Twitter: Newspapers' social media adoption and their online readership. Information Economics and Policy 2012 Mar;24(1):69-74. [doi: 10.1016/j.infoecopol.2012.01.004]
18. Pentina I, Tarafdar M. From "information" to "knowing": Exploring the role of social media in contemporary news consumption. Computers in Human Behavior 2014 Jun;35:211-223. [doi: 10.1016/j.chb.2014.02.045]
19. Morgan SE, Harrison TR, Long SD, Afifi WA, Stephenson MT, Stephenson MS, et al. Family discussions about organ donation: how the media influences opinions about donation decisions. Clin Transplant 2005 Oct;19(5):674-682. [doi: 10.1111/j.1399-0012.2005.00407.x] [Medline: 16146561]
20. McCombs M, Shaw D. The agenda-setting function of mass media. Public Opin Q 1927;36(2):176-187. [doi: 10.1086/267990]
21. Chong D, Druckman JN. Framing theory. Annu Rev Polit Sci 2007 Jun;10(1):103-126. [doi: 10.1146/annurev.polisci.10.072805.103054]
22. Schuck AR, de Vreese CH. Between risk and opportunity: news framing and its effects on public support for EU enlargement. Eur J Commun 2006 Mar 01;21(1):5-32. [doi: 10.1177/0267323106060987]
23. Slapak M. The effect of The World Transplant Games on transplant rates in five continents. Ann Transplant 2004;9(1):46-50 [FREE Full text] [Medline: 15478890]
24. Bae H, Brown WJ, Kang S. Social influence of a religious hero: The late Cardinal Stephen Kim Sou-hwan's effect on cornea donation and volunteerism. Journal of Health Communication 2010 Nov 17;16(1):62-78. [doi: 10.1080/10810730.2010.529489]
25. Krishnaiah S, Kovai V, Nutheti R, Shamanna BR, Thomas R, Rao GN. Awareness of eye donation in the rural population of India. Indian J Ophthalmol 2004 Mar;52(1):73-78 [FREE Full text] [Medline: 15132388]
26. Hussen MS, Gebreselassie KL, Woredekal AT, Adimassu NF. Willingness to donate eyes and its associated factors among adults in Gondar town, North West Ethiopia. BMC Ophthalmol 2017 Oct 2;17(1):178. [doi: 10.1186/s12886-017-0577-1]
27. Ruck JM, Henderson ML, Eno AK, Van Pilsum Rasmussen SE, DiBrito SR, Thomas AG, et al. Use of Twitter in communicating living solid organ donation information to the public: An exploratory study of living donors and transplant professionals. Clin Transplant 2018 Nov 13;33(1):e13447. [doi: 10.1111/ctr.13447]
28. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]
29. Glynn RW, Kelly JC, Coffey N, Sweeney KJ, Kerin MJ. The effect of breast cancer awareness month on internet search activity--a comparison with awareness campaigns for lung and prostate cancer. BMC Cancer 2011;11:442 [FREE Full text] [doi: 10.1186/1471-2407-11-442] [Medline: 21993136]
30. Vraga EK, Stefanidis A, Lamprianidis G, Croitoru A, Crooks AT, Delamater PL, et al. Cancer and social media: A comparison of traffic about breast cancer, prostate cancer, and other reproductive cancers on Twitter and Instagram. J Health Commun 2018 Jan 09;23(2):181-189. [doi: 10.1080/10810730.2017.1421730] [Medline: 29313761]
31. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. Misinformation and its correction: Continued influence and successful debiasing. Psychol Sci Public Interest 2012 Dec;13(3):106-131. [doi: 10.1177/1529100612451018] [Medline: 26173286]
32. Lee S, Cappella JN, Lerman C, Strasser AA. Effects of smoking cues and argument strength of antismoking advertisements on former smokers' self-efficacy, attitude, and intention to refrain from smoking. Nicotine & Tobacco Research 2012 Sep 04;15(2):527-533. [doi: 10.1093/ntr/nts171]
33. Reinhart AM, Marshall HM, Feeley TH, Tutzauer F. The persuasive effects of message framing in organ donation: The mediating role of psychological reactance. Communication Monographs 2007 Jun;74(2):229-255. [doi: 10.1080/03637750701397098]
34. Harel I, Kogut T, Pinchas M, Slovic P. Effect of media presentations on willingness to commit to organ donation. Proc Natl Acad Sci USA 2017 May 01;114(20):5159-5164. [doi: 10.1073/pnas.1703020114]
35. Zhou J, Zhang Q, Zeng DD, Tsui KL. Influence of flavors on the propagation of e-cigarette–related information: Social media study. JMIR Public Health Surveill 2018 Mar 23;4(1):e27. [doi: 10.2196/publichealth.7998]
36. Sina Weibo. 2017 Weibo user report [in Chinese] URL: http://data.weibo.com/report/reportDetail?id=404 [accessed 2019-03-06] [WebCite Cache ID 76fc2n7IB]
37. Hand RK, Kenne D, Wolfram TM, Abram JK, Fleming M. Assessing the viability of social media for disseminating evidence-based nutrition practice guideline through content analysis of Twitter messages and health professional interviews: An observational study. J Med Internet Res 2016 Nov 15;18(11):e295. [doi: 10.2196/jmir.5811]

XSL•FO
RenderX

38. Wang H, Sparks C, Huang Y, Lv N. A content analysis of the journalism models in China traditional media [in Chinese]. Journal of International Communication 2017;39(6):105-123. [doi: 10.13495/j.cnki.cjjc.2017.06.006]
39. Long Q, Li Y. From propaganda to Hegemony: Communication models of new party media in the context of social media [in Chinese]. Journal of International Communication 2017;39(2):52-65. [doi: 10.13495/j.cnki.cjjc.2017.02.003]
40. Yanovitzky I, Stryker J. Mass media, social norms, and health promotion eff. Communication Research 2001 Apr 01;28(2):208-239. [doi: 10.1177/009365001028002004]
41. Cialdini RB, Reno RR, Kallgren CA. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. Journal of Personality and Social Psychology 1990;58(6):1015-1026. [doi: 10.1037/0022-3514.58.6.1015]
42. Hyde M, White K. To be a donor or not to be? Applying an extended theory of planned behavior to predict posthumous organ donation intentions. J of Appl Soc Psychol 2009;39(4):880-900. [doi: 10.1111/j.1559-1816.2009.00464.x]
43. Conesa C, Ríos A, Ramírez P, Rodríguez MM, Rivas P, Canteras M, et al. Psychosocial profile in favor of organ donation. Transplant Proc 2003 Jun;35(4):1276-1281. [Medline: 12826136]
44. Macdonald S, Cunningham Y, Patterson C, Robb K, Macleod U, Anker T, et al. Mass media and risk factors for cancer: the under-representation of age. BMC Public Health 2018 Apr 26;18(1):490 [FREE Full text] [doi: 10.1186/s12889-018-5341-9] [Medline: 29695238]
45. Leroy F, Brengman M, Ryckbosch W, Scholliers P. Meat in the post-truth era: Mass media discourses on health and disease in the attention economy. Appetite 2018 Jun;125:345-355. [doi: 10.1016/j.appet.2018.02.028]
46. Beretta V, Maccagnola D, Cribbin T, Messina E. An interactive method for inferring demographic attributes in Twitter. 2015 Presented at: Proceedings of the 26th ACM Conference on Hypertext & Social Media; Sept 2015; Cyprus p. 113-122. [doi: 10.1145/2700171.2791031]
47. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychological Bulletin 2006;132(2):249-268. [doi: 10.1037/0033-2909.132.2.249]
48. Zhang HC, Zheng JN, Liu WH, Ding J, Zhang LQ. Investigation and strategic analysis of public attitudes and willingness towards organ donation in Jiangsu. Journal of Nanjing Medical University (Social Sciences) 2015;5:339-343 [FREE Full text]

Edited by G Eysenbach; submitted 09.12.18; peer-reviewed by D Arigo, C Laurent, D Carvalho; comments to author 17.01.19; revised version received 11.02.19; accepted 03.03.19; published 12.03.19.

Please cite as:
Jiang X, Jiang W, Cai J, Su Q, Zhou Z, He L, Lai K
Characterizing Media Content and Effects of Organ Donation on a Social Media Platform: Content Analysis
J Med Internet Res 2019;21(3):e13058
URL: http://www.jmir.org/2019/3/e13058/
doi:10.2196/13058
PMID:30860489

©Xiaoya Jiang, Wenshi Jiang, Jiawei Cai, Qingdong Su, Zhigang Zhou, Lingnan He, Kaisheng Lai. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Health-Related Internet Usage and Design Feature Preference for E-Mental Health Programs Among Men and Women

Rachel Smail-Crevier[1], BSc, MSc; Gabrielle Powers[1], BA; Chelsea Noel[1], BSc; JianLi Wang[1,2], MSc, PhD

[1]Work & Mental Health Research Unit, The Royal's Institute of Mental Health Research, University of Ottawa, Ottawa, ON, Canada

[2]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
JianLi Wang, MSc, PhD
Work & Mental Health Research Unit
The Royal's Institute of Mental Health Research
University of Ottawa
1145 Carling Ave
Ottawa, ON, K1Z 7K4
Canada
Phone: 1 613 722 6521 ext 6057
Email: jianli.wang@theroyal.ca

Abstract

Background: Major depressive episodes (MDEs) are prevalent in the workplace and affect workers' health and productivity. Therefore, there is a pressing need for innovation in the prevention of MDEs in the workplace. Electronic mental (e-mental) health programs are a cost-effective approach toward the self-management of stress and emotional issues. E-mental health dropout rate, MDE prevalence, and symptoms greatly vary by sex and age. Thus, the development and implementation of e-mental health programs for the prevention of MDEs need to be examined through a sex and age lens to enhance program use and effectiveness.

Objective: This study aimed to examine design feature preferences based on sex and age for an e-mental health program targeted toward depression prevention.

Methods: Household residents across Canada were contacted using the random digit dialing method. 500 women and 511 men who were 18 years and older and who were at high risk of having MDEs were interviewed. Internet use was assessed using questions from the 2012 Canadian Internet Use Survey conducted by Statistics Canada, and preferred design features of e-mental health program questions were developed by the BroMatters team members. The proportions of likely use of specific features of e-mental health programs in women were estimated and compared with those in men using chi-square tests. The comparisons were made overall and by age groups.

Results: Men (181/511, 35.4%) and women (211/500, 42.2%) differed significantly in their likelihood of using an e-mental health program. Compared with men (307/489, 62.8%), women (408/479, 85.2%) were more likely to use the internet for medical or health-related information. Women were more likely to use the following design features: practices and exercises to help reduce symptoms of stress and depression (350/500, 70.7%), a self-help interactive program that provides information about stress and work problems (302/500, 61.8%), the ability to ask questions and receive answers from mental health professionals via email or text message (294/500, 59.9%), and to receive printed materials by mail (215/500, 43.4%). Men preferred to receive information in a video game format (156/511, 30.7%). Younger men (46/73, 63%) and younger women (49/60, 81%) were more likely to access a program through a mobile phone or an app, and younger men preferred having access to information in a video game format.

Conclusions: Factors such as sex and age influenced design feature preferences for an e-mental health program. Working women who are at high risk for MDEs preferred interactive programs incorporating practice and exercise for reducing stress, quality information about work stress, and some guidance from professionals. This suggests that sex and age should be taken into account when designing e-mental health programs to meet the needs of individuals seeking help via Web-based mental health programs and to enhance their use.

(J Med Internet Res 2019;21(3):e11224) doi:10.2196/11224

XSL•FO
RenderX

KEYWORDS

occupational health; workplace; mental health; preventive health program; depression; internet

Introduction

Major depressive episodes (MDEs) are prevalent in the workplace and affect workers' health and productivity. In the United States, workers with depression cost an estimated US $44.01 billion per year in lost productivity [1]. One of the severe consequences of having an MDE is suicide [2]. Given the significant disease burden, there is a pressing need for innovation in the prevention of MDEs in the workplace. A number of approaches have been proposed and investigated to help workers. Studies in the workplace have focused on providing treatment services to employees with depression [3], including cognitive behavioral therapy (CBT) conducted by professionals, diagnosis of depression, enhanced primary care delivery, treatment by medical doctors, integrated care management, and worksite stress reduction [3-6]. At the population level, despite the significant increase in antidepressant use in the past two decades [7], there has been no measurable change in the prevalence of MDEs in different countries [8-11]. A recent systematic review concluded that there is insufficient evidence to determine which individual-based interventions improve depression management in the workplace [12]. In contrast, another recent meta-analysis concluded that secondary interventions involving CBT-based stress management and tertiary interventions with specific focus on work such as exposure therapy and CBT-based and problem-focused return-to-work programs improved both symptoms of common mental disorders and occupational outcomes [13]. To date, most workplace responses to the burden of mental health problems have been reactive, with the implementation of workplace interventions only being considered once a worker is symptomatic or even on sick leave [14]. Evidence suggests that various mental health problems can be prevented [15]. As a result, policy makers and researchers have begun to consider strategies aimed at early prevention [15,16], for example, identifying and treating asymptomatic persons who have already developed risk factors but in whom the condition has not become clinically apparent (high-risk individuals). From a public health perspective, universal interventions are attractive for their ability to reach more working adults and because they have the ability to reach specified groups and individuals without screening them, an exercise that has shown to be costly [17,18]. In terms of depression prevention, one cost-effective strategy for helping those who are at high risk of depression is an electronic mental (e-mental) health program [19,20].

In the past decade, there has been considerable interest in the delivery of e-mental health programs, that is, delivering mental health services through the internet or mobile apps. Web-based interventions and mobile apps can be utilized for self-help or for treatment purposes [21]. E-mental health programs are either guided or unguided and draw on the internet and other media, including video, phone, and apps for mobile phones and tablets [22], and generally address 4 areas of mental health service delivery: (1) information provision; (2) screening, assessment, and monitoring; (3) intervention; and (4) social support (sometimes concurrently) [19]. The wide use of internet and mobile devices has laid a strong foundation for eHealth development. In 2011, 83% of Canadians reported using Web-based services [23], and in 2014, 99% of young Canadians reported having access to the internet outside of school [24]. The number of individuals using the internet to search for medical or health-related information increased from 57.9% in 2005 to 69.9% in 2009 [25]. Over 60% of internet users use wireless handheld devices to access the internet [23]. With such a large portion of Canada's population online and with many young adults using the internet for health-seeking purposes, e-mental health could be an effective approach toward the self-management of stress and emotional issues and improving service accessibility, geographic barriers, and anonymity [19].

The development and implementation of e-mental health programs for the prevention of MDEs need to be examined through a sex and age lens. The prevalence of MDEs in women is about two times higher than that in men [26]; however, the Canadian national data show that in 2009, 76% of the total number of suicides were committed by men [27]. Due to the existence of gender norms and the desirability of perceived masculinity, men are less likely than women to seek help and to disclose depressive symptoms and often delay seeking help until symptoms become severe [28,29]. Men are socialized to be emotionally stoic and exemplify traditional masculine characteristics such as independence, self-reliance, and dominance [30]. Men are concerned over the perceived negative judgments from family and friends if they access treatment for depression. Furthermore, it has been found that being male significantly increases the risk of treatment dropout for self-guided Web-based interventions for depression, highlighting differences in treatment compliance and willingness to complete Web-based interventions [31]. These sex-related help-seeking behaviors and social norms may influence men's and women's preference for the content and design features of e-mental health programs and the likely use of the program to deal with work-related stress. The use of eHealth programs may also be related to age. For instance, some studies have demonstrated that young adults are more likely to use eHealth programs than the elderly [32], and other studies have shown that younger age is related to low treatment adherence [31]. Given the lack of consensus and knowledge regarding the sex and age differences related to preferred design features and the likely use of e-mental health programs, and treatment adherence, it is imperative to explore design features that may increase their use and thereby, their effectiveness.

In 2015, we initiated a research project called BroMatters, which aimed to develop and evaluate an e-mental health program to be used by working men who are at high risk of having an MDE. To guide the development of the e-mental health program, we conducted a national survey in 2015 among working men who were at high risk of having an MDE, regarding their acceptance and preferred design features of e-mental health programs [33]. In 2016, as a separate study, we repeated this survey in working

XSL•FO
RenderX

women who were at high risk of having an MDE. The objective of this analysis was to examine the sex and age differences in the acceptance and preferred design features of the e-mental health program in male and female workers who are at high risk of having an MDE.

Methods

Participants

The 2 surveys employed the same methodology as described in a previous publication [33]. The target population for the 2 surveys were individuals who (1) were working at the time of the survey, (2) were aged 18 years or older, (3) did not have an MDE in the past 12 months, (4) had no language barrier to either English or French, and (5) were at high risk of having an MDE based on sex-specific risk prediction algorithms [34]. The random digit dialing method was used to contact household residents across the country. Interviews were conducted by the Bureau d'intervieweurs professionels (BIP) located in Montreal. BIP has access to household telephone and validated cell phone numbers. The household contact was asked to retrieve or provide contact information (eg, the first name) of the household residents who were currently working. If there was more than one potentially eligible individual in the same household, one was randomly selected. Once the participant was fully informed about the objectives and procedures of the study, oral consent and continuation of telephone interview were deemed as adequate consent for participation. Data were collected using computer-assisted telephone interviews, which were completed by trained telephone interviewers in participants' language of preference (English or French). Data collection occurred over a period of 9 months between March and December 2015. A minimum of 9 callback attempts were spaced over weekdays for the duration of the data collection period. Participants received a Can $20 incentive for each completed interview as a token of appreciation. Detailed call composition for the men's survey can be found in our previous publication [33]. In the women's survey, 47,555 phone numbers were called. A majority of the calls (46,300/47,555, 97.36%) were not valid (not in service, fax or modem, answering machine, language barriers, ineligibility, duplications, and refusal before eligibility was assessed). Among 1255 eligible women, 755 (60.1%) were excluded from the study (prolonged absence, incomplete questionnaires, scheduled callbacks not within data collection period, and refusal after eligibility was verified). The data collection for women occurred between January and April 2016. For the 2 surveys, 500 eligible women and 511 eligible men were interviewed. The studies were approved by the Conjoint Health Research Ethics Review Board of the University of Calgary.

Measurement

The *sex-specific multivariable risk prediction algorithms for major depression* were administered to estimate the risk (probability) of having an MDE in the next 4 years for each participant [34]. The risk prediction models were designed to be used by individuals who did not have an MDE in the past 12 months. On the basis of participant's exposure to a key set of risk factors (predictors) in the model, the algorithm can generate the absolute risk and probability of having an MDE in the next 4 years, analogous to the Framingham risk prediction algorithm for coronary heart disease [35,36]. The risk prediction algorithms for MDEs were developed and validated using data from 4737 Canadian men and 5864 Canadian women who were aged 18 years or older and who did not have an MDE in the past 12 months [34]. The risk prediction algorithms include age, personal and family history of an MDE, childhood trauma, and ongoing stress and life events. There are also sex-specific predictors in the models (details regarding sex-specific predictors in the models may be found in our previous publication [34]). The predictive power of the risk prediction algorithms was measured by C statistics. The algorithm for men had a C statistic of 0.7953 and the C statistic of the algorithm for women was 0.7667 [34]. The models had excellent calibration with data, as indicated by the Hosmer-Lemeshow test and visual comparison between the predicted and observed risks by decile risk groups [34]. In our study, ≥6.51% and ≥11.19% were defined as high risk for men and women, respectively, which represents the top 2 decile risk groups in the Canadian male and female populations.

Internet use was assessed using questions from the 2012 Canadian Internet Use Survey conducted by Statistics Canada, including use of the internet in the past 12 months, number of hours spent on the internet in the past week, use of the internet to search for medical- and health-related information, and perceived importance of the health information on the internet for decision making.

Preferred design features of e-mental health program questions were developed by the BroMatters team members. Participants were asked "We want to hear your opinion about e-mental health programs for dealing with work and stress issues. Electronic health (eHealth) is defined as...For the following features, please indicate how likely it is that you would use them." For 17 design features, participants were asked how likely they were to use a feature and answered on a 5-point Likert scale ranging from very likely to very unlikely. Open-ended questions were asked about any other features they may want in an e-mental health program, whether the participant and his or her coworkers would use an e-mental health program to deal with work stress, and what makes it difficult to use an e-mental health program. At the end of this set of questions, the participants were asked "would you or your co-workers use an eHealth program to deal with work stress?" The participants answered *yes*, *maybe*, or *no*. For eligible participants, administering the questions and instruments took an average of 22 min to complete.

Data Analysis

A total of 500 men and 511 women were included in the data analysis. Responses including *do not know* and *refuse to answer* were excluded from the data analysis. There were no missing data for demographic characteristics or general and health-related internet usage. The percentage of missing data (ie, do not know and refuse to answer) for items regarding preferred design features of an e-mental health program ranged from 0% to 2%. Demographic and socioeconomic characteristics of participants were tabulated. The proportions of general and health-related internet usage were compared between women

XSL•FO
RenderX

and men using a chi-square test. The proportions of likely use of specific features of e-mental health programs in women were estimated and compared with those in men using chi-square tests. The comparisons were made overall and by age groups. As men and women were compared in 17 specific design features, the Bonferroni correction was used, and the significance level was set at .003.

Results

Demographic Characteristics

Table 1 presents the demographic characteristics of the participants included in this study. Compared with men, women were slightly older; were more likely to be divorced, separated, or widowed; and were more likely to have a lower personal annual income, to have a part-time job, and to be from small- or mid-sized organizations. Finally, a higher proportion of women obtained a higher level of education (partial or completed university), as compared with men.

General and Health-Related Internet Usage

Table 2 presents the general and health-related internet usage for men and women. Men and women did not differ significantly in terms of the time they spent online or the means by which they accessed the internet. For instance, 35.8% (175/489) of men and 38.0% (182/479) of women indicated that within the past 12 months, they had spent less than 5 hours per week online. Similarly, the majority of men (408/489, 83.4%) and women (413/ 479, 86.2%) accessed the internet using various devices (mobile phone, tablet, etc). Both men (256/307, 83.4%) and women (334/408, 81.9%) believed that being able to access health resources on the internet was important or very important. Women (408/479, 85.2%) compared with men (307/489, 62.8%) were significantly more likely to use the internet for medical or health-related information. Conversely, men (231/307, 75.2%) were significantly more likely to state that the internet was useful in helping them make decisions compared with women (277/408, 67.9%).

When participants were asked “would you or your co-workers use an eHealth program to deal with work stress?, women were more likely to report *yes* than men (P=.04). Among 511 men, 35.4% (181) reported *yes*, 37.2% (190) reported *maybe*, and 21.9% (112) reported *no*, whereas among 500 women, 42.2% (211) reported *yes*, 34.0% (170) reported *maybe*, and 16.4% (82) reported *no*.

Preferred Design Features

Table 3 contains the preferred design features in an e-mental health program for men and women. Men and women differed significantly in terms of their e-mental health design feature preferences. Compared with men, a higher proportion of women stated that they were likely to use design features such as practices and exercises to help reduce symptoms of stress and depression (350/495, 70.7%), a self-help interactive program that provides information about stress and work problems (302/489, 61.8%), the ability to ask questions and receive answers from mental health professionals via email or text message (294/491, 59.9%), and to receive printed materials by mail (215/495, 43.4%). Compared with women (104/496, 21.0%), a higher proportion of men (156/509, 30.7%) preferred having some of the program information about ways of dealing with stress and work-related issues to be delivered in a video game format. There were no significant differences between men and women for the remaining design features.

Table 4 presents the preferred design features of an e-mental health program for men and women by age group. Age differences were found in 3 design features in men and women. Compared with older men, men in the youngest age group (36/73, 49%) preferred to receive information on how to deal with work and stress issues in video game format. Men aged between 50 and 64 years (61/140, 43.6%) preferred to receive printed materials by mail rather than having to print online materials oneself, compared with younger men and men over the age of 65 years. Younger men also preferred accessing a program through a smartphone or as an app, with 63% (46/73) of 18- to 29-year olds stating that they would likely use this feature. Younger women also preferred accessing a program through a smartphone or as an app (P<.001), considering that of 489 women, 81% (49/60) of 18 to 29 year olds, 64.2% (163/254) of 30 to 49 year olds, 36% (59/162) of 50 to 64 year olds, and 15% (2/13) of women aged 65 years and older said that they were likely to access a program in this manner. There were no other significant age differences in men or women regarding e-mental health design feature preferences.

Sensitivity analyses were conducted in participants who reported likely use of an e-mental health program to deal with work stress and in those with a history of an MDE. In participants who reported likely use of an e-mental health program, information delivered in a video game format and receiving printed materials by mail were still preferred among men (69/179, 38.0%) and women (115/210, 54.8%), respectively. In addition, men preferred to receive information about anger management (98/180, 55%) and women preferred to chart and track mood (133/209, 63.6%). There were no significant differences between men and women for the remaining design features. In participants with a history of an MDE, women were more likely to use a self-help interactive program (139/218, 63.8%), to receive information about improving sleep hygiene (153/216, 70.8%), and to receive printed materials by mail (91/216, 42.1%), and men were more likely to receive information in a video game format (43/130, 33.1%).

XSL•FO
RenderX

Table 1. Demographic and socioeconomic characteristics of the participants who were at high risk of major depression classified by sex.

Variable	Men (n=511)	Women (n=500)	*P* value
Age (years), mean (SD)	42.0 (12.2)	44.3 (15.0)	.002
Marital status, n (%)			**<.001**
Married or common-law	389 (76.1)	342 (69.0)	
Divorced, separated, or widowed	22 (4.3)	68 (13.7)	
Single	100 (19.6)	86 (17.3)	
Personal income, n (%)			**<.001**
<Can $30,000	60 (12.1)	130 (27.3)	
Can $30,000-<$60,000	152 (30.6)	216 (45.4)	
Can $60,000-<$80,000	98 (19.7)	69 (14.5)	
Can $80,000+	187 (37.6)	61 (12.8)	
Education level, n (%)			**<.001**
<High school	43 (8.4)	12 (2.4)	
High school	107 (20.9)	81 (16.3)	
College	167 (32.7)	188 (37.7)	
University or higher	194 (38.0)	217 (43.6)	
Employment, n (%)			**.98**
Employee	413 (81.5)	396 (81.0)	
Self-employed	92 (18.2)	91 (18.6)	
Family business no pay	2 (0.4)	2 (0.4)	
Job type, n (%)			**<.001**
Full time	434 (84.9)	340 (68.0)	
Part time	37 (7.2)	130 (26.0)	
Seasonal	18 (3.5)	10 (2.0)	
Contract	19 (3.7)	18 (3.6)	
Other	3 (0.6)	2 (0.4)	
Size of company or worksite, n (%)			**.01**
<50	276 (54.2)	290 (59.7)	
50-499	149 (29.3)	148 (30.4)	
>500	84 (16.5)	48 (9.9)	
Provinces, n (%)			**.16**
British Columbia	33 (7.1)	58 (13.2)	
Alberta	52 (11.2)	51 (11.6)	
Saskatchewan	18 (3.9)	15 (3.4)	
Manitoba	20 (4.3)	15 (3.4)	
Ontario	170 (36.6)	161 (36.7)	
Quebec	141 (30.4)	109 (24.8)	
New Brunswick	12 (2.6)	8 (1.8)	
Nova Scotia	12 (2.6)	13 (2.9)	
Newfoundland	4 (0.9)	7 (1.6)	
Prince Edward Island	2 (0.4)	2 (0.5)	
Language, n (%)			**.31**
English	374 (73.2)	380 (76.0)	

XSL•FO
RenderX

Variable	Men (n=511)	Women (n=500)	*P* value
French	137 (26.8)	120 (24.0)	
Work function impairment, n (%)			**.91**
None	314 (62.9)	296 (61.4)	
Mild	146 (29.3)	150 (31.1)	
Moderate	36 (7.2)	34 (7.1)	
Severe	3 (0.6)	2 (0.4)	

Table 2. General and health-related internet usage during the past 12 months in men and women who were at high risk of major depression.

Internet use	Men		Women		*P* value
	n (%)	95% CI	n (%)	95% CI	
Use internet for personal use (n=1011)	489 (95.7)	93.5-97.2	479 (95.8)	93.6-97.2	.93
Hours of internet use each week (n=968)					**.35**
<5 hours	175 (35.8)	31.6-40.2	182 (38.0)	33.7-42.4	
5-9 hours	137 (28.0)	24.2-32.2	144 (30.1)	26.1-34.3	
10-19 hours	113 (23.1)	19.6-27.1	102 (21.3)	17.8-25.2	
20-29 hours	39 (8.0)	5.9-10.7	35 (7.3)	5.3-10.1	
30-39 hours	11 (2.3)	1.2-4.0	8 (1.7)	0.8-3.3	
>40 hours	14 (2.9)	1.7-4.8	6 (1.3)	0.6-2.8	
Do not know	0 (0)	0-0	2 (0.42)	0.1-1.7	
Access internet with a smart phone, tablet, or other device (n=968)	408 (83.4)	79.9-86.5	413 8(6.2)	82.8-89.0	.23
Used internet for medical or health-related information (n=968)	307 (62.8)	58.4-67.0	408 (85.2)	82.1-88.4	<.001
How useful was the internet in helping you make a decision (n=715)					**.04**
Not useful	32 (10.4)	7.5-14.4	41 (10.1)	7.5-13.4	
Unsure	41 (13.4)	10.0-17.7	88 (21.6)	17.8-25.8	
Useful	231 (75.2)	70.1-79.8	277 (67.9)	63.2-72.3	
How important is it for you to be able to access health resources on the internet (n=715)					**.22**
Not important	31 (10.1)	7.2-14.0	32 (7.8)	5.6-10.9	
Unsure	20 (6.5)	4.2-9.9	41 (10.1)	7.5-13.4	
Important	256 (83.4)	78.8-87.2	334 (81.9)	78.0-85.3	

XSL•FO
RenderX

Table 3. Proportions of preferred design features of e-mental health program in men and women who were at high risk of major depression.

Feature	Men		Women		*P* value
	n (%)	95% CI	n (%)	95% CI	
Practice and exercise to reduce stress	303 (59.5)	55.2-63.7	350 (70.7)	66.5-74.6	<.001
Information about improving sleep hygiene	313 (61.3)	56.9-65.4	329 (66.9)	62.6-70.9	.09
Educational materials	295 (57.8)	53.5-62.1	313 (63.8)	59.4-67.9	.002
Self-help interactive program	244 (47.8)	43.5-52.2	302 (61.8)	57.4-66.0	<.001
Setting personal goals and tracking them	277 (54.6)	50.3-58.9	296 (59.9)	55.5-64.2	.02
Ask questions and receive answers from mental health professionals	253 (49.7)	45.4-54.1	294 (59.9)	55.5-64.1	.005
Direct referral to health professional to deal with work and stress issues in person	263 (51.7)	47.3-56.0	292 (59.5)	55.0-63.7	.004
Being able to access a program via a smartphone or as an app	264 (52.0)	47.6-56.3	273 (55.8)	51.4-60.2	.39
Watching videos online on how to deal with work and stress issues	272 (53.3)	49.0-57.6	271 (54.5)	50.1-58.9	.58
Receiving printed materials by mail	167 (32.7)	28.7-36.9	215 (43.4)	39.1-47.9	<.001
Chart and track mood	217 (42.6)	38.4-47.0	237 (48.1)	43.7-52.5	.15
Access by phone to a trained coach to help with work stress	211 (41.3)	37.1-45.6	235 (48.1)	43.6-52.5	.02
Information about anger management	211 (41.5)	37.2-45.8	190 (38.3)	34.1-42.7	.17
A risk calculator predicting future risk of having major depression	215 (42.6)	38.3-46.9	184 (38.1)	33.9-42.5	.13
Online chat room	129 (25.3)	21.7-29.3	124 (25.0)	21.3-29.0	.25
Online peer connection	140 (27.5)	23.7-31.5	136 (27.8)	24.0-31.9	.02
Information delivered in a video game format	156 (30.7)	26.8-34.8	104 (21.0)	17.6-24.8	.002

Table 4. Proportions of preferred design features of e-mental health program in high-risk men by age groups.

Feature	18-29 years		30-49 years		50-64 years		>65 years		*P* value
	n (%)	95% CI	n (%)	95% CI	n (%)	95% CI	n (%)	95% CI	
Information delivered in video game format (n=509)	250 (49.3)	37.8-60.9	151 (29.7)	24.6-35.3	120 (23.7)	17.3-31.6	108 (21.4)	6.0-54.0	.001
Print materials by mail (n=511)	105 (20.6)	12.6-31.6	159 (31.3)	26.2-37.0	222 (43.6)	35.5-52.0	73 (14.3)	2.9-48.0	.001
Access program via smartphone or app (n=508)	320 (63.0)	51.2-73.5	281 (55.3)	49.4-61.1	211 (41.7)	33.7-50.2	145 (28.6)	9.6-60.1	.003

Discussion

Principal Findings

The data from the 2 national surveys demonstrate that men and women who were at high risk of developing depression did not differ in terms of their general internet usage. However, women were more likely to use the internet for health- and mental health–related information and demonstrated greater acceptability of the use of e-mental health programs for work stress. Although there were similarities among men and women in terms of preferred e-mental health design features, important sex differences emerged. Women preferred design features that were interactive. The only design feature that was preferred by men, as compared with women, was receiving information in a video game format. Age differences in preferred e-mental health design features were minimal, irrespective of sex.

There were some broad similarities among men and women with respect to design feature preference. Men and women differed significantly in their preference for only 5 of the 17 design features. Information regarding sleep hygiene, information about work stress, and practice and exercise to reduce stress were the top 3 features endorsed by men and

XSL•FO
RenderX

women. It is unsurprising that design features, which function to provide health information, were preferred among men and women. Previous research demonstrates that seeking Web-based health information is highly prevalent among internet users [37] and obtaining information is a primary motive for internet use [38].

Notable differences in e-mental health design feature preference were found between men and women. Compared with men, women preferred 5 design features for an e-mental health program including practice and exercise to reduce stress, quality information about work stress, self-help interactive programs, being able to ask questions and receive answers, and receiving printed materials. Men preferred only 1 feature as compared with women—being able to receive information in a game format. These results were replicated in sensitivity analyses among men and women with a history of MDE; men were more likely to use a feature that delivered information in a video game format, whereas women preferred to use a self-help interactive program.

Our findings are in accordance with previous studies, which demonstrate that men and women do not differ in their self-reported general internet usage [39], although women are more likely to use the internet for health-related information [40,41], as compared with men. Previous studies have demonstrated greater acceptability, attitudes, help-seeking, and use of e-mental health programs [30,42] as well as more favorable attitudes toward seeking psychological professional help through face-face, email interaction, and Web-based counseling [43] among women. This study likewise found that women demonstrated greater acceptability of the use of e-mental health programs for work stress.

Comparison and Interpretation

To our knowledge, no studies have examined sex or age differences in specific e-mental health design feature preference. Thus, direct comparisons with previous studies on sex and age differences in preferred design features are not possible. However, previous research examining health-related internet use has demonstrated that women are more likely to be interactive Web-based health users [40], which is consistent with the results of this study. In comparison with women, men preferred to receive information in a video game format, a finding that is consistent with general video game use across genders, as studies show that men play twice as many video games per week as women [44].

These differences may be useful to consider when designing effective eHealth interventions for mental health. For example, low adherence is a significant problem for most eHealth programs, and this issue may be exacerbated among population groups who are at risk for MDEs but are not currently experiencing an episode [1]. The incorporation of interactive features or gamification strategies to increase adherence may be a useful method of increasing adherence among the overall target audience. Gamification may be especially useful for increasing adherence in men, who are receptive to receiving information in a video game format but also face substantial barriers to seeking mental health treatment [45,46].

A number of factors likely contribute to the observed sex differences in patterns of health information seeking, eHealth acceptability, and design feature preference for an e-mental health program. Compared with men, women may be more likely to search for Web-based mental health information because they have stronger positive beliefs regarding the benefits of Web-based health searches [47] and because it is an efficient method of searching [48]. According to some authors, this can be explained from a social role perspective [49-51]. As women often have a multitasking agenda, especially in middle age, efficiency and convenience may be more highly regarded. Fewer men reported that they would use an eHealth program to manage work-related stress. This reflects the tendency for men to be less likely than women to seek either formal or informal help for mental health issues [30]. As some authors have noted, study participants with a history of seeking Web-based health information have overcome some practical and stigma-related barriers in seeking Web-based psychological help [52]. As such, greater acceptability of e-mental health programs among women may be partially attributed to their greater use of the internet for health and mental health information. Women also tend to have greater psychological openness, which describes the ability to acknowledge a psychological problem and a need for help. Conversely, men have more difficulty identifying feelings of distress as emotional problems [53]. In addition, a number of perceived barriers, such as internalized and treatment stigma [45], masculine norms, communication barriers, the role of self-help strategies, and perceptions of mental health, have been shown to prevent men from accessing Web-based health resources [46]. Other barriers, such as privacy; ease of navigation; personal relevance; and lack of personal interaction, time, and knowledge, may prevent men from using eHealth programs for depression [33]. Thus, there are probable barriers other than attitudes toward help seeking that inhibit men from seeking Web-based help.

The majority of design features preferred among women were interactive in nature. Furthermore, women were more likely to use a guided feature in which they could ask questions and receive answers from a mental health professional. This reflects women's use of eHealth programs for social motives and enjoyment, in addition to information seeking [48], and their willingness to speak to a mental health professional [43]. Men preferred to receive information in a video game format, compared with women. Men enjoy gaming and seek out game-play for social situations more so than women [54]; consequently, they may find eHealth information more interesting and engaging in this format.

Among men and women who were likely to use an e-mental health program for work stress, few differences in design feature preference were evident. Most notably, women no longer preferred interactive features as compared with men. Men who are likely to use an e-mental health program may have greater acceptability of eHealth programs, resulting in a greater likelihood of using a variety of e-mental health features. Factors such as perceived internet skill and perceived credibility of Web-based information predict the use of Web-based interactive features and may also differ among men who are likely to use an e-mental health program for work stress [55].

XSL•FO
RenderX

To our knowledge, there are no studies that have examined specific design feature preferences of e-mental health programs across age; however, the general use of mobile phones and apps is consistent with our findings: younger individuals are more likely to own a mobile phone [56] and to download and use health-related apps [57] than older individuals. Accordingly, tailoring eHealth interventions to participants' age may improve the efficacy of eHealth interventions.

Limitations

First, the findings from this study are limited by the survey data's reliance on self-reporting. As such, reporting and recall biases are possible and causal inferences cannot be drawn. Second, the surveys in men and in women were not conducted during the same time. The impact of timing of the survey on the results is not clear. Third, both surveys were conducted in Canada; thus, precautions need to be taken when extrapolating the results to other regions. Fourth, the response rate for women was relatively low, as compared with men. Among 1255 eligible women, 755 (60.1%) were excluded from the study because of prolonged absence, incomplete questionnaires, scheduled callbacks not within data collection period, and refusal after eligibility was verified. This may have reflected systematic bias, wherein females with negative attitudes toward psychological treatment or e-mental health have declined to respond to the survey. Furthermore, income level was used to calculate the risk of depression in women; thus, women with a lower income level may be overrepresented in this sample. Fifth, women showed greater acceptability toward the use of e-mental health programs, which could influence design feature preference. Likewise, previous treatment experience was not examined in this study and may have an effect on e-mental health acceptability [58] as well as design feature preference. Sixth, design feature preference was measured by the use of a single item. The use of more items may have generated a score that is more reliable. Finally, although studies have found that treatment acceptability is important to consider as it may improve both treatment adherence [59] and overall outcome [60], and although intention is believed to be the best proximal predictor of actual behavior, it may not translate directly to the actual use of internet technology use [61].

Conclusions

E-mental health programs can play an important role in the prevention of workplace depression. To enhance the program's acceptance, adherence, and effectiveness, researchers and program developers should account for patients' preferences and needs in the design process, including any sex- or age-based differences. This research found that both men and women highly endorsed an intervention containing quality information about stress reduction, work-related stress, and sleep hygiene. However, sex differences emerged. Most notably, women preferred interactive programs that incorporated exercises to reduce stress and Web-based guidance, whereas men preferred receiving information in a video game format. Tailored programs, such as those that involve interactive exercises and guidance or are presented in a video game format, could be beneficial in enhancing e-mental health use in women and men, respectively. Future studies should evaluate the effectiveness of e-mental health programs from sex and age perspectives and identify innovative ways for enhancing the use of e-mental health programs in working men and women at high risk of developing depression.

Acknowledgments

The authors thank E Jones, R Huntley, J Thomas, M Uppal, S McKinney, and NG Tafreshi for their assistance in data collection. This work was supported by the Department of Psychiatry, Faculty of Medicine, University of Calgary, Mathison Centre for Mental Health Research and Education, Hotchkiss Brain Institute, University of Calgary, and HBI-Pfizer Research Award.

Conflicts of Interest

None declared.

References

1. Van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
2. Bostwick JM, Pankratz VS. Affective disorders and suicide risk: a reexamination. Am J Psychiatry 2000 Dec;157(12):1925-1932 [FREE Full text] [doi: 10.1176/appi.ajp.157.12.1925] [Medline: 11097952]
3. Lo Sasso AT, Rost K, Beck A. Modeling the impact of enhanced depression treatment on workplace functioning and costs: a cost-benefit approach. Med Care 2006 Apr;44(4):352-358. [doi: 10.1097/01.mlr.0000204049.30620.1e] [Medline: 16565636]
4. Furukawa TA, Horikoshi M, Kawakami N, Kadota M, Sasaki M, Sekiya Y, GENKI Project. Telephone cognitive-behavioral therapy for subthreshold depression and presenteeism in workplace: a randomized controlled trial. PLoS One 2012;7(4):e35330 [FREE Full text] [doi: 10.1371/journal.pone.0035330] [Medline: 22532849]
5. Mino Y, Babazono A, Tsuda T, Yasuda N. Can stress management at the workplace prevent depression? A randomized controlled trial. Psychother Psychosom 2006;75(3):177-182. [doi: 10.1159/000091775] [Medline: 16636633]
6. Martin A, Sanderson K, Cocker F. Meta-analysis of the effects of health promotion intervention in the workplace on depression and anxiety symptoms. Scand J Work Environ Health 2009 Jan;35(1):7-18 [FREE Full text] [Medline: 19065280]

XSL•FO
RenderX

7. Patten SB, Williams JV, Lavorato DH, Fiest KM, Bulloch AG, Wang J. Antidepressant use in Canada has stopped increasing. Can J Psychiatry 2014 Nov;59(11):609-614 [FREE Full text] [doi: 10.1177/070674371405901107] [Medline: 25565477]
8. Jorm AF. Why hasn't the mental health of Australians improved? The need for a national prevention strategy. Aust N Z J Psychiatry 2014 Sep;48(9):795-801. [doi: 10.1177/0004867414546387] [Medline: 25147289]
9. Jorm AF, Patten SB, Brugha TS, Mojtabai R. Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. World Psychiatry 2017 Feb;16(1):90-99 [FREE Full text] [doi: 10.1002/wps.20388] [Medline: 28127925]
10. Mojtabai R, Jorm AF. Trends in psychological distress, depressive episodes and mental health treatment-seeking in the United States: 2001-2012. J Affect Disord 2015 Mar 15;174:556-561. [doi: 10.1016/j.jad.2014.12.039] [Medline: 25556674]
11. Patten SB, Williams JV, Lavorato DH, Bulloch AG, Wiens K, Wang J. Why is major depression prevalence not changing? J Affect Disord 2016 Jan 15;190:93-97. [doi: 10.1016/j.jad.2015.09.002] [Medline: 26485311]
12. Furlan AD, Gnam WH, Carnide N, Irvin E, Amick BC, DeRango K, et al. Systematic review of intervention practices for depression in the workplace. J Occup Rehabil 2012 Sep;22(3):312-321. [doi: 10.1007/s10926-011-9340-2] [Medline: 22161150]
13. Joyce S, Modini M, Christensen H, Mykletun A, Bryant R, Mitchell PB, et al. Workplace interventions for common mental disorders: a systematic meta-review. Psychol Med 2016 Mar;46(4):683-697. [doi: 10.1017/S0033291715002408] [Medline: 26620157]
14. Harvey SB, Henderson M. Occupational psychiatry. Psychiatry 2009 May;8(5):174-178. [doi: 10.1016/j.mppsy.2009.03.011]
15. 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]
16. Hickie IB. Preventing depression: a challenge for the Australian community. Med J Aust 2002 Oct 07;177(Suppl):S85-S86. [Medline: 12358562]
17. Tan L, Wang MJ, Modini M, Joyce S, Mykletun A, Christensen H, et al. Preventing the development of depression at work: a systematic review and meta-analysis of universal interventions in the workplace. BMC Med 2014 May 09;12:74 [FREE Full text] [doi: 10.1186/1741-7015-12-74] [Medline: 24886246]
18. Lynch FL, Hornbrook M, Clarke GN, Perrin N, Polen MR, O'Connor E, et al. Cost-effectiveness of an intervention to prevent depression in at-risk teens. Arch Gen Psychiatry 2005 Nov;62(11):1241-1248. [doi: 10.1001/archpsyc.62.11.1241] [Medline: 16275811]
19. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 01;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]
20. Musiat P, Tarrier N. Collateral outcomes in e-mental health: a systematic review of the evidence for added benefits of computerized cognitive behavior therapy interventions for mental health. Psychol Med 2014 Nov;44(15):3137-3150. [doi: 10.1017/S0033291714000245] [Medline: 25065947]
21. 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]
22. 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]
23. Statistics Canada. 2013. Individual Internet use and e-commerce URL: http://www.statcan.gc.ca/daily-quotidien/131028/dq131028a-eng.pdf [accessed 2018-05-31] [WebCite Cache ID 6zpjFNSIy]
24. Steeves V. Media Smarts. 2014. Young Canadians in a Wired World, Phase III: Life Online URL: http://mediasmarts.ca/ycww/life-online [accessed 2018-11-01] [WebCite Cache ID 6zpkQ2Y9c]
25. Statistics Canada. 2010. Internet use at home, by Internet activity and urban or rural distribution URL: https://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/comm29a-eng.htm [accessed 2018-05-31] [WebCite Cache ID 6zpjuodr8]
26. Patten SB, Williams JV, Lavorato DH, Wang JL, McDonald K, Bulloch AG. Descriptive epidemiology of major depressive disorder in Canada in 2012. Can J Psychiatry 2015 Jan;60(1):23-30 [FREE Full text] [doi: 10.1177/070674371506000106] [Medline: 25886546]
27. Navaneelan T. Statistics Canada. Suicide rates: An overview URL: https://www.statcan.gc.ca/pub/82-624-x/2012001/article/11696-eng.htm [accessed 2018-05-31] [WebCite Cache ID 6zpk5WF0q]
28. Swami V. Mental health literacy of depression: gender differences and attitudinal antecedents in a representative British sample. PLoS One 2012;7(11):e49779 [FREE Full text] [doi: 10.1371/journal.pone.0049779] [Medline: 23166769]
29. Judd F, Komiti A, Jackson H. How does being female assist help-seeking for mental health problems? Aust N Z J Psychiatry 2008 Jan;42(1):24-29. [doi: 10.1080/00048670701732681] [Medline: 18058440]
30. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. Am Psychol 2003 Jan;58(1):5-14. [Medline: 12674814]
31. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726. [doi: 10.1017/S0033291715000665] [Medline: 25881626]
32. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]

XSL•FO
RenderX

33. Wang J, Lam RW, Ho K, Attridge M, Lashewicz BM, Patten SB, et al. Preferred features of e-mental health programs for prevention of major depression in male workers: results from a Canadian national survey. J Med Internet Res 2016;18(6):e132 [FREE Full text] [doi: 10.2196/jmir.5685] [Medline: 27267782]
34. Wang JL, Manuel D, Williams J, Schmitz N, Gilmour H, Patten S, et al. Development and validation of prediction algorithms for major depressive episode in the general population. J Affect Disord 2013 Oct;151(1):39-45 [FREE Full text] [doi: 10.1016/j.jad.2013.05.045] [Medline: 23790813]
35. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Feb 12;117(6):743-753 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.699579] [Medline: 18212285]
36. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. Circulation 2010 Apr 20;121(15):1768-1777 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.849166] [Medline: 20404268]
37. 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;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
38. Papacharissi Z, Rubin AM. Predictors of internet use. J Broadcast Electron Media 2010 Jun 07;44(2):175-196. [doi: 10.1207/s15506878jobem4402_2]
39. Powell J, Clarke A. Internet information-seeking in mental health: population survey. Br J Psychiatry 2006 Sep;189:273-277. [doi: 10.1192/bjp.bp.105.017319] [Medline: 16946364]
40. Kummervold PE, Chronaki CE, Lausen B, Prokosch H, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
41. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. Int J Med Inform 2006 Jan;75(1):8-28. [doi: 10.1016/j.ijmedinf.2005.07.032] [Medline: 16125453]
42. 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 [FREE Full text] [doi: 10.1093/her/cyl062] [Medline: 16880222]
43. Tsan JY, Day SX. Personality and gender as predictors of online counseling use. J Technol Hum Serv 2008 Sep 08;25(3):39-55. [doi: 10.1300/J017v25n03_03]
44. Greenberg BS, Sherry J, Lachlan K, Lucas K, Holmstrom A. Orientations to video games among gender and age groups. Simul Gaming 2008 Jul 25;41(2):238-259. [doi: 10.1177/1046878108319930]
45. 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]
46. Ellis LA, Collin P, Hurley PJ, Davenport TA, Burns JM, Hickie IB. Young men's attitudes and behaviour in relation to mental health and technology: implications for the development of online mental health services. BMC Psychiatry 2013;13:119 [FREE Full text] [doi: 10.1186/1471-244X-13-119] [Medline: 23601273]
47. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
48. Bidmon S, Terlutter R. Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. J Med Internet Res 2015;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]
49. Marton C, Wei Choo C. A review of theoretical models of health information seeking on the web. J Doc 2012 Apr 20;68(3):330-352. [doi: 10.1108/00220411211225575]
50. Young R. The household context for women's health care decisions: impacts of UK policy changes. Soc Sci Med 1996 Mar;42(6):949-963. [Medline: 8779006]
51. Yoo E, Robbins LS. Understanding middle-aged women's health information seeking on the web: a theoretical approach. J Am Soc Inf Sci 2008 Feb 15;59(4):577-590. [doi: 10.1002/asi.20766]
52. Wallin EE, Mattsson S, Olsson EM. The preference for internet-based psychological interventions by individuals without past or current use of mental health treatment delivered online: a survey study with mixed-methods analysis. JMIR Ment Health 2016 Jun 14;3(2):e25 [FREE Full text] [doi: 10.2196/mental.5324] [Medline: 27302200]
53. Mackenzie CS, Gekoski WL, Knox VJ. Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. Aging Ment Health 2006 Nov;10(6):574-582. [doi: 10.1080/13607860600641200] [Medline: 17050086]
54. Lucas K. Sex differences in video game play: a communication-based explanation. Communication Research 2004 Oct 01;31(5):499-523. [doi: 10.1177/0093650204267930]
55. Chung DS. Interactive features of online newspapers: identifying patterns and predicting use of engaged readers. Comput Mediat Commun 2008;13(3):658-679. [doi: 10.1111/j.1083-6101.2008.00414.x]
56. Poushter J. Pew Research Center. 2016. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies URL: http://www.pewresearch.org/wp-content/uploads/sites/2/2016/02/

XSL•FO
RenderX

pew_research_center_global_technology_report_final_february_22__2016.pdf [accessed 2019-01-25] [WebCite Cache ID 75hBepG00]
57. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
58. Musiat P, Goldstone P, Tarrier N. Understanding the acceptability of e-mental health--attitudes and expectations towards computerised self-help treatments for mental health problems. BMC Psychiatry 2014;14:109 [FREE Full text] [doi: 10.1186/1471-244X-14-109] [Medline: 24725765]
59. Santana L, Fontenelle LF. A review of studies concerning treatment adherence of patients with anxiety disorders. Patient Prefer Adherence 2011;5:427-439 [FREE Full text] [doi: 10.2147/PPA.S23439] [Medline: 21949606]
60. Swift JK, Callahan JL, Cooper M, Parkin SR. The impact of accommodating client preference in psychotherapy: a meta-analysis. J Clin Psychol 2018 Nov;74(11):1924-1937. [doi: 10.1002/jclp.22680] [Medline: 30091140]
61. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]

Abbreviations

BIP: Bureau d'intervieweurs professionels
CBT: cognitive behavioral therapy
eHealth: electronic health
e-mental health: electronic mental health
MDE: major depressive episode

Edited by G Eysenbach; submitted 04.06.18; peer-reviewed by L Bijker, S Dawadi, F Kählke; comments to author 13.09.18; revised version received 07.11.18; accepted 10.12.18; published 18.03.19.

Please cite as:
Smail-Crevier R, Powers G, Noel C, Wang J
Health-Related Internet Usage and Design Feature Preference for E-Mental Health Programs Among Men and Women
J Med Internet Res 2019;21(3):e11224
URL: http://www.jmir.org/2019/3/e11224/
doi:10.2196/11224
PMID:30882361

©Rachel Smail-Crevier, Gabrielle Powers, Chelsea Noel, JianLi Wang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Do Search Engine Helpline Notices Aid in Preventing Suicide? Analysis of Archival Data

Qijin Cheng[1,2], PhD; Elad Yom-Tov[3,4], PhD

[1]Department of Social Work, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Microsoft Research, Herzeliya, Israel

[4]Faculty of Industrial Engineering and Management, Technion, Haifa, Israel

Corresponding Author:
Elad Yom-Tov, PhD
Microsoft Research
13 Shenkar Street
Herzeliya, 46733
Israel
Phone: 972 747111359
Email: eladyt@yahoo.com

Abstract

Background: Search engines display helpline notices when people query for suicide-related information.

Objective: In this study, we aimed to examine if these notices and other information displayed in response to suicide-related queries are correlated with subsequent searches for suicide prevention rather than harmful information.

Methods: Anonymous suicide-related searches made on Bing and Google in the United States, the United Kingdom, Hong Kong, and Taiwan in a span of 10 months were extracted. Descriptive analyses and regression models were fit to the data to assess the correlation with observed behaviors.

Results: Display of helpline notices was not associated with an observed change in the likelihood of or future suicide searches (*P*=.42). No statistically significant differences were observed in the likelihood of people making future suicide queries (both generally and specific types of suicide queries) when comparing search engines in locations that display helpline notices versus ones that do not. Pages with higher rank, being neutral to suicide, and those shown among more antisuicide pages were more likely to be clicked on. Having more antisuicide Web pages displayed was the only factor associated with further searches for suicide prevention information (hazard=1.18, *P*=.002).

Conclusions: Helpline notices are not associated with harm. If they cause positive change in search behavior, it is small. This is possibly because of the variability in intent of users seeking suicide-related information. Nonetheless, helpline notice should be displayed, but more efforts should be made to improve the visibility and ranking of suicide prevention Web pages.

(J Med Internet Res 2019;21(3):e12235) doi:10.2196/12235

KEYWORDS

search engines; suicide

Introduction

Suicide prevention has been recognized as a global imperative by the World Health Organization [1]. Systematic reviews of suicide prevention strategies found that restricting access, physically and cognitively, to means of suicide is crucial for suicide prevention [2,3]. With the rapid increase in internet usage globally, Web-based search engines have become a crucial gateway for people to access information on suicide methods or places to access the means for suicide [4-6]. However, search engines can also facilitate people to find helping resources or access to Web-based support [7-9].

Since 2010, search engine providers have been gradually adding a box with information of the contact details of a suicide helpline when suicide-related queries are made [10]. Hereafter, we will refer this intervention as *helpline notice* or *the notice*. It is hoped that the prominent placement of helpline notices can attract suicidal individuals to utilize intervention resources rather than to further look for harmful information. However, other than the special placement of helpline notices, search engine

XSL•FO
RenderX

providers use an automated algorithm to rank search results to satisfy the user's information need by matching a document to a user's search query and preference [11]. In other words, even though a helpline notice is displayed at the top of a search result page, information about suicide methods might also be shown on the search results page when a person queries for, for example, *how to kill myself*. In a situation like this, will the user read the helpline notice, harmful information about suicide methods, or both?

Although examination of the information returned by search engines when searching suicide-related queries has been conducted in multiple countries [5,7,8,12-18] to the best of our knowledge, no research exists on whether helpline notices successfully and sufficiently direct people to read information on prevention rather than harmful information, such as information about suicide methods or encouragement to suicidal behavior.

In view of the research gap and social significance, this study was conducted with an aim to answer the following research questions:

1. To what extent is the display of helpline notices associated with search engine user's clicks on different types of suicide-related information?
2. Beside the presence of helpline notices, what other search characteristics, such as search queries, search timing (ie, when a search was made), and display availability of different types of information predict whether or not users would click through different types of suicide-related information?
3. Would the presence of helpline notices and the presence of antisuicide information in the search results prevent people from further searching for prosuicide information? Would such presence facilitate further searching for suicide prevention information?

Methods

Study Regions

The United States (US), the United Kingdom (UK), Hong Kong (HK), and Taiwan (TW) were chosen as the study regions for 2 major reasons. First, both Google and Bing search engines are freely accessible in the 4 regions. At the time of data collection, Google provided helpline notices in all 4 regions, whereas Bing provided such notices only in the US and UK. Therefore, a comparative analysis can be performed stratifying regions and search engines. Second, US and UK can be viewed as representatives of the English-speaking cultures, whereas HK and TW are viewed as representatives of the Chinese-speaking cultures, which allows for a cross-cultural comparison. Mainland China, although population wise is more comparable with UK and US, was not chosen because it implements rather strict censorship on the internet, and Google search is blocked there.

Aside from the abovementioned differences, the 4 jurisdictions share commonalities that allow for a comparison: they all have very high internet penetration rates (87% to 95%) and relatively high freedom of speech, where suicide is not criminalized, and no specific legal statute regulates suicide-related information. The 4 jurisdictions report different patterns regarding suicide: the age-standardized suicide rate in the UK (6.2 per 100,000) is lower than the other 3, followed by HK (8.8 per 100,000), US (12.1 per 100,000), and TW (13.1 per 100,000). The latter 3 had comparable rates [1] of suicide. All the 4 jurisdictions have suicide prevention hotlines available 24/7 for free for nation-wide callers. The types of suicide-related information which can be accessed by search engines have been examined individually in each of the 4 regions [7,8,12,14] (see details in the Multimedia Appendix 1). An initial comparison found that people in HK might be exposed to less prosuicide information through Web-based search than the other 3 regions. We noted, however, that those studies did not use consistent assessment protocols.

Informed consent was waived by the IRB. This study was approved by the IRBs of The University of Hong Kong (for labeling of data) and by the Technion (for the offline analysis of archival data).

Data Collection

We used data from 2 primary sources:

Dataset 1—All searches made on the Bing search engine by people in the 4 regions between November 2016 and August 2017 (a total of 10 months). For each user we collected an anonymous user identifier, the region from where the user made the search, time and date, the text of the query, and the pages shown to people in response to the query and which pages they clicked on.

Dataset 2—A sample of anonymized logs from consenting users of a widely distributed Web browser add-on toolbar associated with the Microsoft Internet Explorer Web browser. Only visits to search engines were logged from these data. For each user in these data, we collected an anonymous user identifier, the region from where the user made the search, time and date, and the text of the query. These data allow cross-search engine comparisons.

The data were filtered to include only suicide-related queries. These queries were identified in 2 ways: English-language queries were those that triggered a helpline notice by Bing during June 2017 and Chinese-language queries were those queries that triggered helpline notices by Google in HK and TW. We noted that, to the best of our knowledge, helpline notices are triggered for a predefined list of query terms.

Owing to the different population sizes and to obtain a large enough sample from all 4 regions from dataset 1, we focused on those users who made suicide queries from US and UK in June 2017, and between January and June 2017 for HK and TW. Data from outside these dates were used for analysis of searches by users before and after they made suicide searches. In dataset 2, we focused on users who made suicide queries during June 2017 and tested for prior and post searches using data from May 1, 2017, to August 31, 2017.

Suicide-related queries in this study can be grouped into 2 categories: (1) queries which are associated with suicide prevention resources (referred as *suicide prevention queries*

XSL•FO
RenderX

hereafter, eg, *helpline*, *lifeline*, *national suicide*, *crisis*, *hotline*, *chat*, *Samaritans*, or the telephone numbers of well-known local suicide prevention hotlines) and (2) queries associated with suicidal thoughts, behaviors, or suicide methods (referred as *suicide method queries* hereafter, eg, *poison*, *pills*, *hanging*, or *pesticide* but excluding the irrelevant phrases such as *poison ivy*). For both these query types, the Chinese-language equivalents were also included.

During data collection, we noted that the number of suicide-related searches and the number of frequently displayed Web pages varied in the 4 regions. To make a balanced comparison, in each country we selected 100 Web pages most often displayed to people in response to suicide-related queries. As there were only 91 Web pages displayed 10 or more times in HK, all of the 91 pages were assessed.

Content Coding

Following previous examinations on Web pages' bias toward suicide [13-14], we classified the Web pages into 6 types: antisuicide (ie, about suicide prevention or discouraging suicide), prosuicide (ie, the main content is encouraging suicide, romanticizing suicide, or introducing detailed suicide methods), neutral (ie, factual information about suicide incidents or statistics without notable bias), mixed (ie, mixed prosuicide and antisuicide information in the same Web page), not a suicide site (using *suicide* as a metaphor or joke, eg, animal suicide and political suicide), or error (ie, the Web page or major content of the page cannot be found). One modification on the previous commonly used coding frame is that we split the code Neutral/Mixed to 2 codes Neutral and Mixed because we suspect the 2 types of contents may have different impacts on vulnerable individuals. Examples of the 6 types of Web pages can be found in the Multimedia Appendix 1.

The assessment was performed by 6 trained coders. The training was provided by the first author and included guided reading of a booklet of recommendations on how to responsibly represent suicide-related information in mass media and Web-based media [19], and group exercise with coding 20 sample Web pages together. After training, each region's dataset was assigned to 2 coders to conduct the classification independently. The codes agreed by the both coders were assigned as the final classification results. In total, 87% to 95% of each region's data received agreed coding in the first round. For the remaining disagreed items, a third coder (ie, the first author) joined in the 2 coders to discuss the items until arriving at a consensus.

Data Analysis

Descriptive analysis was first conducted to summarize the proportion of different types of Web pages being displayed in the search results in each study region. To answer Q1 & Q2, we used both descriptive statistics of dataset 2 and a logistic regression model fit to the data from dataset 1, with whether a Web page was clicked as the dependent variable (DV). The independent variables (IVs) included 3 types: (1) characteristics of a specific search, including when and where the search was made, whether or not helpline notice was displayed in the search results, and fractions of displayed Web pages in each type (calculated as the number of a certain type of pages being displayed divided by the number of total displayed pages) in the search results; (2) characteristics of a specific Web page in the search results, including ranking of the Web page, and what bias toward suicide the Web page shows; and (3) characteristics of the same user's previous search within 7 days, including the total number of search queries made in 7 days, and the average number of search queries made every day. The latter was included to represent the level of Web-based activity of individual users. A total of 20,077,272 observations were included in this analysis.

To answer Q3, a hazard model was fit to the data from dataset 1, with 2 DVs: whether in the next 7 days the same user searched for (1) suicide method queries or (2) suicide prevention queries. The IVs included fractions of each type of page being displayed, fractions of each type of page being clicked, whether or not the helpline notice is being displayed, and the terms used in initial search queries. As only people who were active for at least 7 consecutive days were eligible for this analysis, a total of 154,286 observations were included in this analysis.

Results

Figure 1 shows the distribution of displayed pages in the 4 regions of Bing search in terms of their bias toward suicide. As the figure shows, the vast majority of pages frequently shown to users are antisuicide pages across the 4 regions. However, in TW the frequency of displayed Web pages contained a higher proportion of prosuicide (13.8%) but lower proportion of antisuicide information (15.0%) compared with the other 3.

Figure 1 also shows the click-to-display ratio by region and type of pages. The ratio is calculated as follows: (% clicks/%displays−1)×100. Despite a smaller proportion in frequently displayed pages, prosuicide and mixed pages were much more likely to be clicked in all but the UK. Conversely, although antisuicide pages were more likely to be displayed, their click-to-display ratio appeared to be low across the 4 regions and were even negative in TW. In addition, neutral pages were more likely to be clicked in HK and TW but not so in UK and US when being displayed. Keywords from queries associated with the highest probability of clicking on a specific page type are shown in the Multimedia Appendix 1.

An analysis of the association between clicks on helpline notices and other behaviors on the search page shows that 83.7% of people who clicked on the notice also clicked on other links compared with 91.5% of those who did not click on the helpline notice (χ^2=59.3, n=49,950, P<.001). Clicks on helpline notices are not associated with a change in the likelihood of future suicide searches (40.1% when not clicked, 38.6% when they are, χ^2=0.7, P=.42).

XSL•FO
RenderX

Figure 1. Characteristics of displayed and clicked Web pages in response to suicide-related search queries in the 4 study regions. Unlabeled pages are pages displayed fewer than 10 times in the study period. HK: Hong Kong; TW: Taiwan; UK: United Kingdom; US: United States of America.

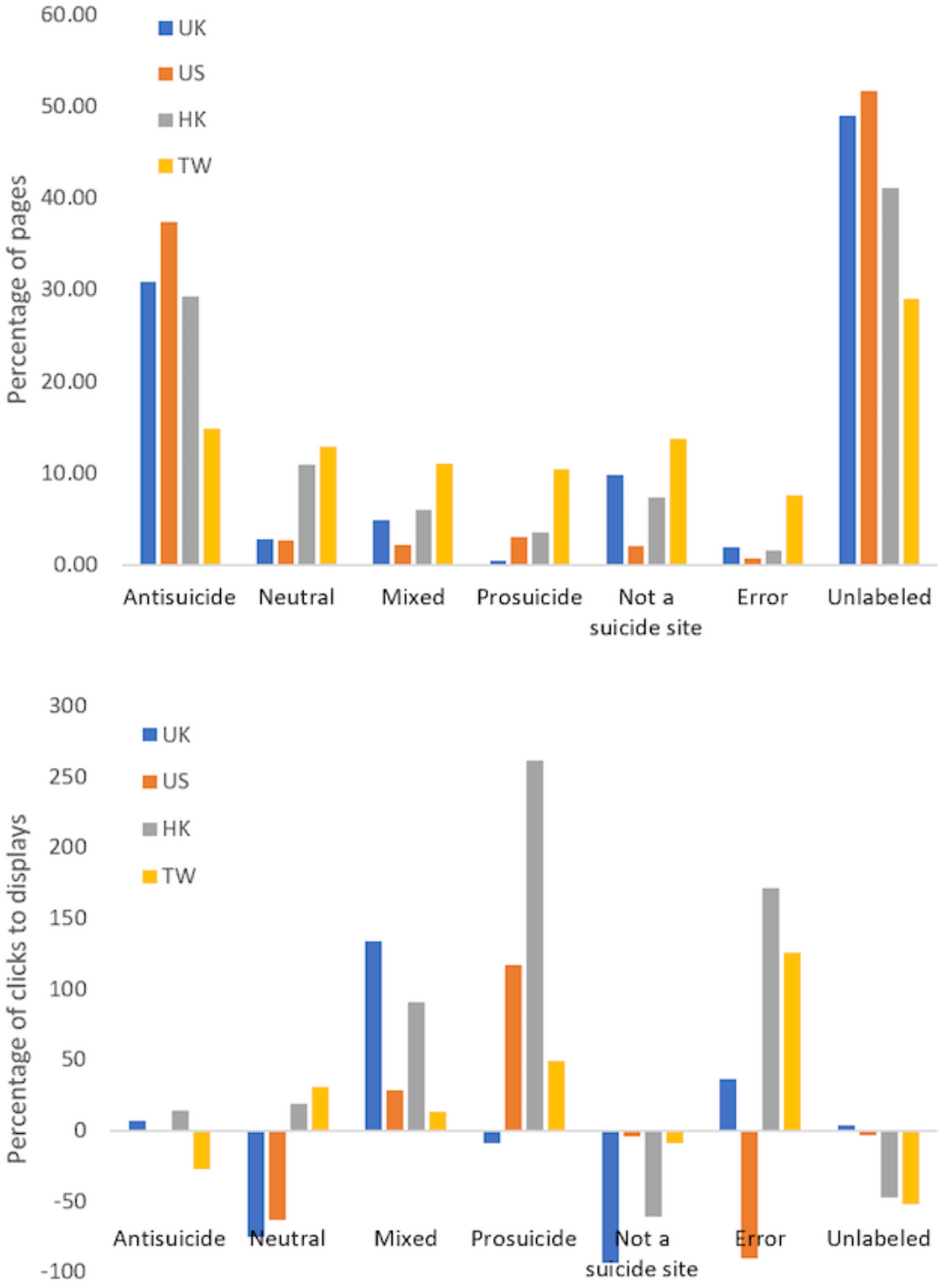


Table 1 shows the logistic regression model parameters for predicting whether a user would click on a page given its parameters and those of the user and their past search behaviors. As the table shows, helpline notice displays are not statistically significantly associated with a change in click behavior. A page with higher rank being neutral to suicide and having more antisuicide pages displayed in the same results list were more likely to be clicked on. That page's probability to be clicked would be even higher when a search was made in later time of a day (as the slope is positive) and an early day of a week (negative slope, where days of the week are coded from Sunday (1) to Saturday (7)). In terms of regions, the click behavior in UK differed from that of the other regions. Previous levels of search activity did not show significant influence on one's click behavior in the model.

XSL•FO
RenderX

Table 1. Parameters of the logistic regression model to predict clicks on a page. The United States has been used as the baseline region in the cross-region model, and neutral pages as the baseline for webpage types.

Parameters	Estimation (SE)	*P* value
Search results characteristics		
Is helpline displayed	9.333 (34)	>.99
Fraction of antisuicide displayed	0.817 (0.037)	<.001
Fraction of neutral displayed	0.664 (0.109)	<.001
Fraction of mixed displayed	−0.851 (0.072)	<.001
Fraction of prosuicide displayed	−0.766 (0.053)	<.001
Page characteristics		
Rank	0.205 (0.002)	<.001
Antisuicide	−1.034 (0.033)	<.001
Prosuicide	−1.073 (0.039)	<.001
Mixed	−1.32 (0.038)	<.001
Time		
Hour of the day	−0.006 (0.001)	<.001
Day of the week	0.008 (0.003)	.003
Region		
United Kingdom	−0.882 (0.015)	<.001
Hong Kong	−10.026 (34)	>.99
Taiwan	−10.1 (34)	>.99
Previous search characteristics		
Number of previous queries	0.000 (0.000)	.46
Previous queries per day	0.000 (0.000)	.004

Table 2 shows the hazard model parameters for predicting searches for suicide method queries in the next 7 days. As the table shows, people who were exposed to more suicide-related pages were associated with higher likelihood of further searching for suicide method queries. Among terms, initial search queries of *suicide* and *kill (myself)* were associated with higher risk to further search suicide methods, as opposed to *suicide hotline*. Some terms, however, were less expected in their association with increased or decreased hazards. For example, people who queried for *kill yourself* or *kill myself* were less likely to further search for suicide methods. The model for future suicide prevention queries found only one statistically significant variable—the fraction of antisuicide pages displayed (hazard=1.182, *P*=.002).

Using dataset 2, we compared user behaviors, stratified by region and search engine. Table 3 shows the probability that users will make future suicide-related queries, future suicide method queries, or future helpline queries, given their region and search engine (*P* values are calculated using the chi-square test). As shown in the table, no statistically significant difference can be found among the users of any of the 4 regions. The finding is consistent with the hazard model's results.

XSL•FO
RenderX

Table 2. Hazard model parameters for predicting future suicide methods searches. The following are attributes with statistically significant results (P<.05).

Parameters	Hazard ratio	*P* value
Fraction of antisuicide displayed	1.0131	<.001
Fraction of neutral displayed	1.0811	<.001
Fraction of mixed displayed	1.0702	<.001
Fraction of prosuicide displayed	1.0428	.002
Number of previous suicide queries	1.0009	<.001
Term: *suicide*	1.4137	<.001
Term: *kill*	2.4258	<.001
Term: *suicide hotline*	0.8088	<.001
Term: *to kill*	1.2231	<.001
Term: *kill yourself*	0.5616	<.001
Term: *suicidal*	1.3115	<.001
Term: *kill myself*	0.5247	<.001

Table 3. The probability of further searching for overall suicide queries, suicide method queries, and suicide prevention queries by region and search engine. *P* values calculated using the chi-square test.

Country	All suicide queries			Suicide methods queries			Suicide prevention queries		
	Bing	Google	*P* value	Bing	Google	*P* value	Bing	Google	*P* value
United Kingdom	0.457	0.508	.27	0.009	0.008	.91	0.009	0.004	.50
United States	0.467	0.498	.20	0.008	0.014	.27	0.012	0.016	.46
Hong Kong	0.696	0.750	.67	0.261	0.357	.46	0.087	0.036	.44
Taiwan	0.801	0.755	.31	0.635	0.612	.66	0.000	0.005	.36

Discussion

As Web-based search engines are commonly used by people as the gateway to information, it is imperative to examine how the information returned by search engines influences people's behaviors for the purpose of suicide prevention. To the best of our knowledge, this study is the first to empirically examine the helpline notice intervention's impacts, doing so across geographic regions, languages, and search engines.

Our results suggested that helpline notices do no harm, as it was neither associated with greater clicks on prosuicide pages nor with further search for suicide method information. However, the display of helpline notices was not associated with a higher observed likelihood of clicking on or further searching for suicide prevention information. Although a previous study found that the prominent placement of the helpline notice draws people's attention [20], our results suggest that more views do not result in more click-throughs. This might be related to the fact that the helpline notices provide a helpline number, using which people can directly call the helpline without clicking on search results. Alternatively, people searching for suicide-related information may be doing so for a variety of reasons. There might be only a small group of users whose searching behaviors were affected by the display of helpline notice.

Compared with the display of the helpline notice, the types of Web pages displayed in the search results and the search queries people made appear to be more predictive of what people read and further search for. We posit that search queries may reflect the users' suicide risk intention and psychological needs [21]. For example, those who searched for the suicide hotline might be willing to seek help and would less likely further search for suicide methods. However, it seems to be surprising that those who initially searched for *kill yourself* or *kill myself* were also less likely to later search for suicide methods. A potential explanation is that those people might be satisfied by the information they found in their initial search or have taken actions to harm themselves after the initial search, which reduced their likelihood of further searching.

More displays of antisuicide Web pages in the search results were found to be the only factor associated with further searching for suicide prevention information, after controlling for what search queries were used initially. The result suggests that, when a person is exposed to more antisuicide information, regardless of their initial mental status and needs, their interests in suicide prevention information might be stimulated or their information need might be unsatisfied, leading to additional queries. The finding might also be explained by the planned behaviors theory that people's behavior is influenced by what they perceive as the social norm [22,23]. When more antisuicide information is displayed, it may create an impression that suicide prevention is a widely accepted norm and attract people to learn more about it.

XSL•FO
RenderX

People were exposed mostly to antisuicide pages in the results displayed by Bing. We note that, among the 4 study regions, people in TW saw more prosuicide and fewer antisuicide pages in their search results, and their antisuicide pages got the lowest click-to-display ratio. The findings are consistent with another study examining what suicide-related information was shown in the first 3-page search results in TW [14]. Coincidently, TW also reports the highest suicide rate among the 4. A previous study found that Web-based search for a new suicide method in TW was associated with more suicide incidents in the following days [24]. The special phenomenon in TW and its underlying reasons deserve more investigations. Anecdotal observations of the prosuicide pages reveal that many of them were related to a book— *The Complete Manual of Suicide* —which described suicide methods in details and rated each method by its level of painfulness, probability of failure, and other dimensions. By contrast, the promotion of this book only appeared once in HK search results. The book was originally published in Japan and soon translated into Chinese by a TW publisher and sold in both TW and HK. Although the book was later banned from sale in TW and HK, its content is still accessible on the Web and might have caused greater impact in TW, given TW's close cultural relation with Japan.

Some limitations of the study should be noted. First, we only selected 4 regions for this study, which does not allow us to generalize our results to the entire world. In addition, the included search queries were those that triggered helpline notices. There might be other search queries used for finding suicide information that were absent from the list. Nonetheless, the list included hundreds of terms in English and Chinese, which should have covered the most frequently used ones. Another limitation is that we used clicking on and further searching for certain types of information to indicate behavioral change. Other scenarios such as calling a helpline or acting on suicidal intentions without further searching were not addressed by these measures. Finally, users may be issuing their queries from multiple devices, on multiple services, or from shared devices. This can cause our analysis of future and prior searches to be affected, mostly in a way which underestimates the likelihood of future (or prior) searches.

This study suggested that helpline notices were not associated with harmful outcomes such as clicking on more prosuicide search results or searching for more prosuicide information. Given that a previous study found that people's attention is focused on these notices, it should thus be continued so as to raise public awareness that suicide is preventable and local prevention resources are available. Other than the helpline notice, having more antisuicide pages being displayed in Web-based search results and in higher ranking may facilitate people to access more suicide prevention information. Suicide prevention organizations and other public health institutes, especially those in TW, should increase the antisuicide information's visibility and attractiveness. In addition, health policymakers should be aware of the potential impacts of Web-based search and initiate directives or guidelines for appropriate Web-based information dissemination.

Conflicts of Interest

EYT is an employee of Microsoft, owner of Bing. QC declares no conflict of interest.

Multimedia Appendix 1

1. Background info of the four study regions. 2. Examples of six types of webpages. 3. Terms associated with the highest probability of clicking on a specific page type. 4. Screenshots of help notices from Hong Kong and from the US.

[PDF File (Adobe PDF File), 322KB - jmir_v21i3e12235_app1.pdf]

References

1. World Health Organization. World Health Organization. 2014. Preventing Suicide: A Global Imperative URL: https://www.who.int/mental_health/suicide-prevention/world_report_2014/en/ [accessed 2019-01-30] [WebCite Cache ID 75oW4br08]
2. Mann JJ, Apter A, Bertolote J, Beautrais A, Currier D, Haas A, et al. Suicide prevention strategies: a systematic review. J Am Med Assoc 2005 Oct 26;294(16):2064-2074. [doi: 10.1001/jama.294.16.2064] [Medline: 16249421]
3. Zalsman G, Hawton K, Wasserman D, van Heeringen K, Arensman E, Sarchiapone M, et al. Suicide prevention strategies revisited: 10-year systematic review. Lancet Psychiatry 2016 Jul;3(7):646-659. [doi: 10.1016/S2215-0366(16)30030-X] [Medline: 27289303]
4. Tsai CW, Gunnell D, Chou YH, Kuo CJ, Lee MM, Chen YY. Why do people choose charcoal burning as a method of suicide? An interview based study of survivors in Taiwan. J Affect Disord 2011 Jun;131(1-3):402-407. [doi: 10.1016/j.jad.2010.12.013] [Medline: 21236495]
5. Biddle L, Gunnell D, Owen-Smith A, Potokar J, Longson D, Hawton K, et al. Information sources used by the suicidal to inform choice of method. J Affect Disord 2012 Feb;136(3):702-709. [doi: 10.1016/j.jad.2011.10.004] [Medline: 22093678]
6. Gunnell D, Bennewith O, Kapur N, Simkin S, Cooper J, Hawton K. The use of the Internet by people who die by suicide in England: a cross sectional study. J Affect Disord 2012 Dec 10;141(2-3):480-483 [FREE Full text] [doi: 10.1016/j.jad.2012.04.015] [Medline: 22609195]

XSL•FO
RenderX

7. Biddle L, Derges J, Mars B, Heron J, Donovan JL, Potokar J, et al. Suicide and the Internet: changes in the accessibility of suicide-related information between 2007 and 2014. J Affect Disord 2016 Jan 15;190:370-375. [doi: 10.1016/j.jad.2015.10.028] [Medline: 26546772]
8. Recupero PR, Harms SE, Noble JM. Googling suicide: surfing for suicide information on the Internet. J Clin Psychiatry 2008 Jun;69(6):878-888. [Medline: 18494533]
9. Mars B, Heron J, Biddle L, Donovan JL, Holley R, Piper M, et al. Exposure to, and searching for, information about suicide and self-harm on the Internet: Prevalence and predictors in a population based cohort of young adults. J Affect Disord 2015 Oct 01;185:239-245 [FREE Full text] [doi: 10.1016/j.jad.2015.06.001] [Medline: 26150198]
10. Cohen N. New York Times. 2010 Apr 5. 'Suicide' query prompts Google to offer hotline URL: https://www.nytimes.com/2010/04/05/technology/05google.html?ref=todayspaper [accessed 2019-01-30] [WebCite Cache ID 75oWhjM6g]
11. Carmel D, Yom-Tov E, Darlow A, Fine S. What makes a query difficult? In: 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. 2006 Presented at: SIGIR '06; 2006; Seattle, Washington p. 390-397. [doi: 10.1145/1148170.1148238]
12. Cheng Q. Online suicide-related information in Hong Kong: perspectives from youth suicide prevention. J Youth Stud 2011;14(27):74-85 [FREE Full text]
13. Cheng Q, Fu KW, Yip P. A comparative study of online suicide-related information in Chinese and English. J Clin Psychiatry 2011 Mar;72(3):313-319. [doi: 10.4088/JCP.09m05440blu] [Medline: 20868633]
14. Chen YY, Hung GC, Cheng Q, Tsai CW, Wu KC. Searching for suicide-related information on Chinese websites. Psychiatry Res 2017 Dec;258:506-510. [doi: 10.1016/j.psychres.2017.08.087] [Medline: 28886904]
15. Thornton L, Handley T, Kay-Lambkin F, Baker A. Is a person thinking about suicide likely to find help on the internet? An evaluation of google search results. Suicide Life Threat Behav 2017 Dec;47(1):48-53. [doi: 10.1111/sltb.12261] [Medline: 27111845]
16. Szumilas M, Kutcher S. Teen suicide information on the internet: a systematic analysis of quality. Can J Psychiatry 2009 Sep;54(9):596-604. [doi: 10.1177/070674370905400904] [Medline: 19751548]
17. Westerlund M, Hadlaczky G, Wasserman D. The representation of suicide on the Internet: implications for clinicians. J Med Internet Res 2012 Sep 26;14(5):e122 [FREE Full text] [doi: 10.2196/jmir.1979] [Medline: 23010086]
18. Krysinska K, Andriessen K. On-line support and resources for people bereaved through suicide: what is available? Suicide Life Threat Behav 2010 Dec;40(6):640-650. [doi: 10.1521/suli.2010.40.6.640] [Medline: 21198333]
19. Yip P, Cheng Q. Centre for Suicide Research and Prevention. 2018. Recommendations on Suicide Reporting & Online Information Dissemination for Media Professionals Hong Kong SAR2015 URL: https://csrp.hku.hk/media/ [accessed 2019-01-30] [WebCite Cache ID 75oXRjKCw]
20. Youngmann B, Yom-Tov E. Anxiety and information seeking: Evidence from large-scale mouse tracking. In: World Wide Web Conference. 2018 Presented at: WWW '18; April 23-27, 2018; Lyon, France p. 753-762. [doi: 10.1145/3178876.3186156]
21. Wong PW, Fu KW, Yau RS, Ma HH, Law YW, Chang SS, et al. Accessing suicide-related information on the internet: a retrospective observational study of search behavior. J Med Internet Res 2013 Jan 11;15(1):e3 [FREE Full text] [doi: 10.2196/jmir.2181] [Medline: 23305632]
22. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
23. Montaño DE, Kasprzyk D. Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In: Glanz K, Rimer BK, Kasisomayajula V, editors. Health Behavior: Theory, Research, And Practice (Jossey-Bass Public Health). San Francisco, CA: Jossey-Bass; 2015:95.
24. Chang SS, Kwok SS, Cheng Q, Yip PS, Chen YY. The association of trends in charcoal-burning suicide with Google search and newspaper reporting in Taiwan: a time series analysis. Soc Psychiatry Psychiatr Epidemiol 2015 Sep;50(9):1451-1461. [doi: 10.1007/s00127-015-1057-7] [Medline: 25859754]

Abbreviations

DV: dependent variable
HK: Hong Kong
IVs: independent variables
TW: Taiwan
UK: United Kingdom
US: United States

XSL•FO
RenderX

Edited by G Eysenbach; submitted 17.09.18; peer-reviewed by J Tana, R Morriss; comments to author 06.12.18; revised version received 17.12.18; accepted 03.01.19; published 26.03.19.

Please cite as:
Cheng Q, Yom-Tov E
Do Search Engine Helpline Notices Aid in Preventing Suicide? Analysis of Archival Data
J Med Internet Res 2019;21(3):e12235
URL: http://www.jmir.org/2019/3/e12235/
doi:10.2196/12235
PMID:30912753

©Qijin Cheng, Elad Yom-Tov. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient Rounds With Video-Consulted Relatives: Qualitative Study on Possibilities and Barriers From the Perspective of Healthcare Providers

Christina Østervang[1*], RN, MScN; Lene Vedel Vestergaard[1*], RN, MScN; Karin Brochstedt Dieperink[1,2,3*], RN, MCN, PhD; Dorthe Boe Danbjørg[3,4,5*], RN, MScN, PhD

[1]Department of Oncology, Odense University Hospital, Odense, Denmark

[2]The Danish Knowledge Centre for Rehabilitation and Palliative Care (REHPA), Odense University Hospital, Nyborg, Denmark

[3]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[4]Department of Heamatology, Odense University Hospital, Odense, Denmark

[5]Centre for Innovative Medical Technology, Odense University Hospital, Odense, Denmark

[*]all authors contributed equally

Corresponding Author:
Christina Østervang, RN, MScN
Department of Oncology
Odense University Hospital
Sdr. Boulevard 29, Indg. 98
Odense, 5000
Denmark
Phone: 45 21757009
Email: christina.oestervang.nielsen@rsyd.dk

Abstract

Background: In cancer settings, relatives are often seen as a resource as they are able to support the patient and remember information during hospitalization. However, geographic distance to hospitals, work, and family obligations are reasons that may cause difficulties for relatives' physical participation during hospitalization. This provided inspiration to uncover the possibility of telehealth care in connection with enabling participation by relatives during patient rounds. Telehealth is used advantageously in health care systems but is also at risk of failing during the implementation process because of, for instance, health care professionals' resistance to change. Research on the implications for health care professionals in involving relatives' participation through virtual presence during patient rounds is limited.

Objective: This study aimed to investigate health care professionals' experiences in using and implementing technology to involve relatives during video-consulted patient rounds.

Methods: The design was a qualitative approach. Methods used were focus group interviews, short open interviews, and field observations of health care professionals working at a cancer department. The text material was analyzed using interpretative phenomenological analysis.

Results: Field observational studies were conducted for 15 days, yielding 75 hours of observation. A total of 14 sessions of video-consulted patient rounds were observed and 15 pages of field notes written, along with 8 short open interviews with physicians, nurses, and staff from management. Moreover, 2 focus group interviews with 9 health care professionals were conducted. Health care professionals experienced the use of technology as a way to facilitate involvement of the patient's relatives, without them being physically present. Moreover, it raised questions about whether this way of conducting patient rounds could address the needs of both the patients and the relatives. Time, culture, and change of work routines were found to be the major barriers when implementing new technology involving relatives.

Conclusions: This study identified a double change by introducing both new technology and virtual participation by relatives at the same time. The change had consequences on health care professionals' work routines with regard to work load, culture, and organization because of the complexity in health care systems.

(J Med Internet Res 2019;21(3):e12584) doi:10.2196/12584

XSL•FO
RenderX

KEYWORDS

telehealth; family; relatives; cancer; technology; qualitative research

Introduction

Background

When a patient receives a cancer diagnosis, it will not only affect the patient but also the entire family surrounding the patient [1,2]. Relatives are often seen as a resource in connection with the course of the illness, both when the patients are at home and during admission to the hospital [1,3,4].

However, as Kahriman and Zaybak found in their study, relatives may feel that they are taking on a great responsibility of caring for the patient and all the practicalities related to the illness. As a result of that responsibility, relatives may experience a burden that can have an impact on their physical, emotional, and psychological health [5]. One way to support relatives is to educate them to support the patient [1,6,7]. A systematic review investigated how technology could be a support intervention for relatives of patients with cancer. It found that educational websites and smartphone apps had shown great potential to help ease the family's burden and experiences during the patient's trajectory [8]. A randomized controlled trial by Collinge et al evaluated a multimedia instructional program for family caregivers and found that by using technology both self-efficacy and satisfaction in caregiving was enhanced [9]. Furthermore, Fuentes et al tested a mobile system that maps relatives' social network to prevent social isolation and found that the tool was perceived valuable [10]. However, the systematic review states that many technological solutions to support relatives already exist but minimal research focusing on improving active participation from relatives through a 2-way face-to-face communication has been published and is leaving a gap for further research [8]. During hospitalization, many patients request their relatives to actively participate when decisions about treatment and care are made, which often happens during patient rounds [11]. Relatives can support the patient by remembering and understanding the information given by nurses and physicians [3]. Direct involvement of the relatives may not only result in reducing the burden they feel, but their participation will also bring positive aspects for the patient by reducing frequency and length of hospitalizations [1,2,12]. However, work, family obligations, and distance to the hospital are reasons that cause difficulties for relatives to be present at the hospital and therefore complicate their participation in patient rounds [11,13]. Rising et al found that by using videoconferencing platforms, health care professionals were able to facilitate relative's participation during patient rounds remotely [11]. This inspired us to look at the possibilities that telehealth care provides to enable increased relative participation for hospitalized patients.

Telehealth has been a priority in many countries for years and represents solutions that could provide some of the answers to the challenges that health care systems are facing [14]. These challenges include demographic changes with more elderly patients living with chronic conditions and long distances to hospitals [11,13-16]. In cancer care, the use of telehealth has, as mentioned above, shown beneficial aspects not only as support interventions for relatives but also as a positive addition to promote adherence in treatment and care, improving patient outcomes [17,18]. In spite of these experiences using telehealth, numerous barriers must be considered if the implementation is to succeed [14]. A systematic review shows that the barriers of implementation are often directly related to the culture among the health care professionals. Resistance to changes in working procedures along with their unwillingness to invest time in training in the new workflows were significant barriers [19]. According to Ross et al, potential barriers should be identified early in the process when planning implementation of changes. This will allow formulation of strategies to prevent resistance to change [20]. It is essential that there is a *sense of urgency* if the change is to be carried out among the health care professionals, in addition to a coalition of efficient people who can guide, coordinate, and communicate the change [21]. Although these factors are present, there is still a big risk of the implementation failing [19]. Many studies describe impacts and experiences of telehealth used as part of the treatment among patients and health care professionals [22-24], but in-depth research that explores the implications for health care professionals by involving relatives through a virtual presence is limited [11,13,15].

Objective

The aim of this study was to investigate health care professionals' experiences in using and implementing technology to involve relatives during video-consulted patient rounds.

Methods

Study Design

This is a qualitative study inspired by a phenomenological and hermeneutical position, investigating the experiences of health care professionals participating in video-consulted patient rounds with relatives. Qualitative research is characterized by collection, organization, and interpretation of textual data stemming from statements, conversation, and behavior [25]. The research design is in accordance with the consolidated criteria for reporting qualitative research appropriate for investigating the meanings of the social phenomena experienced and told by individuals, which in this study concerned video-consulted patient rounds [26]. The chosen methods are field observational studies, short open interviews, and focus groups. To obtain insight into the health care professionals' perspective, interpretative phenomenological analysis (IPA; Textbox 1) is used for data analysis [27]. The data from the field observations made it possible to organize the focus group interviews using questions and tasks, related and relevant to specific observations made in connection with video-consulted patient rounds. The combination of the 2 methods also made it possible to strengthen the analysis by validating the field notes with quotes.

XSL•FO
RenderX

Textbox 1. An excerpt of the 6 steps in interpretative phenomenological analysis (IPA) from 1 case, showing the emergence of a main theme.

Step 1: Reading and rereading

- Authors read each case several times

Step 2: Initial noting

- Describe frustrations about the arranged set of time
  - To be a slave of time
  - Time takes control
  - Have to get enough time to be well-prepared
  - A set of time is predetermined that is difficult

Step 3: Developing emergent themes

- Perspective of time
- To be prepared
- Structure of the day

Step 4: Searching for connections across emergent themes—superordinate themes

- Do not wish to be controlled by time
- Content and structure for the conversation

Step 5: Moving to next case

- Authors clear their minds before moving to the next case

Step 6: Looking for patterns across cases

- Just one more thing on the to-do list

Setting and Technology

The study was conducted at a cancer department in the Region of Southern Denmark. The cancer department is a part of a large hospital with 42 departments and 8700 employees. The bed unit where the study was conducted treats 1300 patients per year. This study is part of a larger study, which also includes the perspective of relatives.

For this study, eligible patients were approached at the ward during admission by the first or second author to assess their interest in participating in video-consulted patient rounds. The patients were under active medical treatment owing to their cancer diagnosis and were aged from 61 to 86 years. They had been preselected by the staff nurse, nurses, and physicians at the department based on the inclusion criteria: patients able to hear, able to understand and speak Danish, who are expected to be hospitalized for more than 2 days, and who have relatives with internet connection to their smartphone, tablet, or personal computer.

The connection was provided with the use of the Cisco Jabber app. Staff at the hospital used an administration tool to set up a Cisco Jabber guest link for the relative. The relatives received the link by email. The link was connected to an ad hoc virtual meeting room, which meant that a link could be created for each relative. It required internet access and one of the following devices available to activate the link and participate in video-consulted patient rounds: personal computer, tablet, or smartphone. The staff used a Jabber unified communication client on a tablet. The video conversation was encrypted and complied with the security requirements in line with the Danish legislation concerning management of personal information. The relatives were educated in both oral and written formats on how to use the technology and were also provided with a support line.

Recruitment

Participants were purposively recruited from the cancer department. Participants for the field observational study were nurses and physicians working at the cancer department, both those who carried out the video-consulted patient rounds and those who participated in discussions about it.

At the morning conferences, the attending nurses and physicians were orally informed about which patients had a video-consulted patient round arranged for that day. They were informed by the head nurse that either the first or the second author would be present for the purpose of observation. For the focus group interviews, the inclusion criteria were more rigorous. The inclusion criteria were as follows: health care professionals employed at the cancer department and registered nurses or physicians who had worked with video-consulted patient rounds with relatives once or several times. The exclusion criterion was as follows: physicians and nurses who were employed temporarily. In total, 12 participants were identified and asked

XSL•FO
RenderX

by email if they would take part in one of the focus group interviews. Participants who agreed to participate were included in the focus group interviews, which were carried out in January 2018. Descriptive statistics such as professional background, experience, age, and gender were collected.

Data Collection

Data collection was carried out by the first and second authors, both having many years of experience in nursing but no prior connection to the cancer department.

Field Observational Studies

Field observational studies were carried out for 3 days a week, for a total of 5 weeks from October 2017 to January 2018. The first and second authors were present at the department for 4 to 5 hours per day.

As described by Green and Thorogood, the field observational studies allowed us to directly obtain knowledge about what participants do and what they say they do, in connection with video-consulted patient rounds [28]. Furthermore, it provided the opportunity to conduct short open interviews in the field, posing a few open questions to physicians and nurses [29]. The questions were determined and verified as part of discussions among the research group and also based upon the field observations. The field observational studies were carried out with the acceptance of the department's management. The nurses and physicians who were involved in the video-consulted patient rounds in different ways, such as planning, scheduling, and discussions of content in the actual consultations, were observed. The observations were carried out in 2 ways inspired by James Spradley's description of *moderate participation* and *passive participation* [30]. *Moderate participation* was carried out as the first and second authors conducted observations of all sessions with video-consulted patient rounds. Before and after the patient rounds, both authors carried out small open interviews, gaining understanding of the participants' thoughts and experiences in addition to the observations. In addition, *passive participation* was applied in relation to video-consulted patient rounds. First and second authors listened and observed the work of health care professionals by being present in the conference room, the hallway, the offices, and in the hospital rooms. As Spradley outlines, these observations gave firsthand knowledge of and insight into verbal and nonverbal statements and actions [30]. For each day, field notes were taken as keywords and were later in the same day transcribed into continuous text to secure correct recall [28].

Focus Groups

In addition to the field observational studies, focus groups were conducted in January 2018. We chose focus groups as they allow the researcher to obtain knowledge from the interactions between the participants, and we wanted to mobilize associations where the dynamics between the participants contribute to creation of narratives and discussions about the use of video-consulted patient rounds [28]. Focus groups were built around the specific topic, video-consulted patient rounds, and the interaction in the group facilitated discussions about the use of technology to enable the participation of relatives [28]. A semistructured interview guide with open questions and tasks guided the focus groups. Each focus group interview was split into 2 parts. First, the participants were asked to write down 3 positive and negative thoughts about video-consulted patient rounds with relatives and afterward they discussed their opinions. Second, participants were introduced to quotes spoken by relatives concerning video-consulted patient rounds and the following discussions started with the participants' thoughts about the quotes.

Special effort was put into creating the groups, as groups that were too homogeneous risked a lack of interaction and groups that were too heterogeneous risked larger disagreements [31]. The first author facilitated the sessions located in a conference room at the hospital. The second author was present as an observer, writing field notes and validating the content of the discussion.

Ethical Considerations

In accordance with the Helsinki Declaration and the Ethical Guidelines for Nursing Research in the Nordic countries, the participants were informed both orally and in writing, and we obtained written consent [32]. All participants received an information letter describing the aim and focus of the project. They were informed about their right to withdraw from the study, at any time, without consequences, and that their data would be anonymized. Furthermore, participants agreed to show respect and confidentiality about statements made during the focus group. Pictures taken during the sessions were only used with written consent from the participants. According to Danish legislation at the time of this study, the study did not need ethical approval or approval from the National Committee on Health Research Ethics. The study is registered with the Danish Data Protection Agency (17/43851). The data are stored in SharePoint (Microsoft Corporation).

Data Analysis

In accordance with the qualitative research approach, the transcripts were analyzed using the detailed 6 step guide for IPA provided by Smith and Osborn [33] as shown in Textbox 1. IPA is phenomenological in that it concerns exploring experiences in its own terms and adds the ideographic and hermeneutical philosophy to interpret a small sample size [33].

Going through the initial focus group data and initial observational data required different processes as the interpretations were being derived from different positions. Data from focus group interviews emerged directly from the participant's spoken words, whereas data from the field observations occurred as a product of the first and second authors' observations and preunderstandings. Therefore, the first and second authors handled steps 1 to 4 in the IPA process separately for each dataset. This analytical process involved reading the transcripts several times, followed by open coding with a focus on the descriptive comments, leading to superordinate themes [33]. To ensure that the participants' experiences were adequately represented in the themes, data were continuously checked throughout the entire process. The first and second author read the transcripts separately, followed by discussions to identify themes and codes in the data. From the list of superordinate themes, 3 main themes were derived

XSL•FO
RenderX

from patterns across the focus group interviews combined with the field observations. Synthesizing the data at step 6 was considered to supplement and validate the results of the analysis by underlining equivalence and differences. To ensure identification of convergence and divergence in the data, themes were arranged using NVivo 11 (QSR International).

Results

Description of the Participants

In total, 12 health care professionals were approached and we obtained informed consent from 9 to participate in the focus group interviews. The 3 health care professionals who did not attend the focus groups had busy schedules. The population of participants consisted of 7 nurses and 2 physicians, 3 men and 6 women aged between 24 and 60 years, and their professional experience ranged between 6 months and 25 years (Table 1).

Field observational studies were conducted for a total of 15 days, which yielded a total of 75 hours of observation. A total of 8 short open interviews were carried out with physicians, nurses, and staff from the department's management. In total, 14 sessions of video-consulted patient rounds were observed, and 15 pages of field notes written. On the basis of IPA, 3 main themes were derived from patterns across the focus group interviews, field notes, and transcripts of short open interviews. The themes are as follows (Table 2):

1. Relatives can qualify the conversation.
2. Is It a patient round or a family round?
3. Just one more thing on the to-do list.

Relatives Can Qualify the Conversation

This theme was derived from consistent expressions by physicians and nurses concerning relatives' positive impact during patient rounds, which was a universal assumption at the department. Health care professionals described the relatives' involvement as a resource for obtaining useful information, helping the patient to remember important information and to follow up on, for example, changes in the medical treatment:

> *The physician talks about new medication, the patient looks at his daughter at the screen, he looks confused. The daughter says that she will pick it up at the pharmacy and bring it to his home after discharge.* [Field note, December 13, 2017]

Furthermore, health care professionals experienced great willingness by the relatives to participate and take responsibility for the information given at the patient rounds. Health care professionals experienced that using telehealth to enable participation by relatives reduced the number of misunderstandings during the hospitalization and reduced stress among both patients and relatives:

> *Many misunderstandings are reduced or prevented when relatives participate* [Nurse, 26 years]

The technology made it possible for the relatives to extend the conversation further by making corrections, for example, of what was realistic after discharge. In this context, health care professionals experienced that the visual contact was an improvement compared with a phone, because of the possibility to visualize the patient's home:

> *The relative could walk around in the patient's living room with the camera. That way we are able to see and direct focus on the challenges with for example mobilization after discharge.* [Nurse, 43 years]

The technology made it possible to deliver information related to treatment and care to the relative on the same day they were given to the patient, instead of waiting for several days until the relative could be physically present or reached by phone:

> *Specific details about the discharge are arranged, and the relative contributes with clarifying information, which the patient was not able to remember. The patient looks at her daughter at the screen and smiles, she looks relieved.* [Field note, November 20, 2017]

Health care professionals described how the technology made it possible to involve cross-sectional collaborators, for instance, rehabilitation homes, simultaneously with the participation of the relative:

> *The staff at the rehabilitation home was able to see him and talk to him. Moreover, they met his wife virtually and talked to her. It sounded like some of the wife's concerns disappeared.* [Nurse, 54 years]

Health care professionals experienced that this way of preparing discharges could minimize anxiety related to the discharge for both the patient and the relative.

Is it a Patient Round or a Family Round?

In general, health care professionals experienced video consultation as a possible way of accommodating participation by both patients and relatives in treatment and care.

> *This is a good way of getting knowledge about the relatives too.* [Short open interview with manager, January 6, 2018]

Although the health care professionals recognized the opportunities in involving both relatives and patients in the consultation, they also emphasized that they unintentionally directed their focus toward the tablet screen and indirectly the relative, rather than focusing on the patient. They did not feel comfortable using the new technology, mostly because they felt challenged by the amount of attention they had to direct at the relative behind the screen:

> *I think that it's difficult to turn my focus away from the screen, suddenly you are at risk of forgetting the patient. It's important for me that we talk directly to the patient… but of course we should listen to the relative too.* [Nurse, 26 years]

XSL•FO
RenderX

Table 1. Sociodemographics of health care professionals (n=9) in a Danish study of video-consulted patient rounds at a cancer department.

Sociodemographics	Statistics
Profession, n	
Nurse	7
Physician	2
Age (years)[a] , mean	
Nurse	37
Physician	46
Professional experience (years)[b]	
<5 years	4
>5 years	5
Number of participations in patient rounds, mean	2

[a]Range: 24 to 60 years.

[b]Range: 4 months to 22 years.

Table 2. From superordinate themes to 3 main themes in a Danish study investigating telehealth in patient rounds.

Examples of superordinate themes	Main themes
First theme	
Easy access to information	Relatives can qualify the conversation
Recognizable technology	Relatives can qualify the conversation
Reduce misunderstandings	Relatives can qualify the conversation
Second theme	
Content of the conversation is important	Is it a patient round or a family round?
Expectations from the relatives	Is it a patient round or a family round?
The presence of the screen	Is it a patient round or a family round?
Third theme	
Being controlled by time is stressful	Just one more thing on the to-do list!
New workflows are difficult	Just one more thing on the to-do list!
High workload	Just one more thing on the to-do list!

In relation to the screen, the health care professionals stressed that it would require a period of time to learn how to use the technology, which they perceived as a new coplayer. All participants agreed that the definition of the content in the conversation had great importance. That way they were able to establish a regulatory framework that still made it possible for them to plan their work, in spite of this new way to do patient rounds:

> *This is new, and we have to learn to use it, and to define the content of the conversation. We also have to acknowledge that the rounds are held with focus on the patient.* [Physician, 43 years]

The health care professionals expressed the importance of matching expectations with relatives before their participation. This way, it would be clearly stated what the health care professionals were able to offer.

Health care professionals found that the solution was time-consuming and that it could result in them taking care of the needs expressed by relatives too, which they did not find themselves capable of:

> *The conference room is filled; an interdisciplinary meeting is held. Some are standing up. The discussion is concerning the selection of patients to participate in video consulted patient rounds. One physician says: this way of doing patient rounds only takes the relatives into account. Not the patients. Another physician says, it could be time-consuming involving the family. Nurses and managerial staff are silent.* [Field note, December 19, 2017]

Nurses and physicians experienced that for the video-consulted patient rounds to be a success, they must have a clearly defined structure and definition of content in the conversation, which everybody in the department agrees to follow.

Just One More Thing on the To-Do List

The majority of the health care professionals found the technology easy to use and recognizable from their everyday

XSL•FO
RenderX

lives, using FaceTime or Skype. However, when image or sound quality failed one or more times during some sessions, it felt stressful and caused inconvenient disruptions to the conversation.

These interruptions were time-consuming and required health care professionals to have technical skills to be able to resolve the issues. This was a contributory reason for not arranging this type of patient rounds often.

> *Even though the technical solution is easy... The fact that I know from the beginning of the day that I'm responsible for the preparations - it's just one more thing to do in a busy schedule.* [Nurse, 25 years]

Video-consulted patient rounds required a new way of organizing work, which was perceived as difficult and resistance to change was seen in more situations:

> *A health care professional says at the morning conference, that it would be nice if work was like in the old days. There's silence in the room. The health care professional continues; "in the old days I only had to do patient rounds, not all these new things". The majority of the health care professionals nod their heads, the managerial staff too.* [Field note, November 20, 2017]

The nurses experienced a great responsibility for the technical setup to be ready at the agreed time. The responsibility was not only for the tablet to be ready but also to ensure that the patient was ready, the relative had appeared live on the screen, and let the physicians know that everything was well-prepared:

> *You feel a great responsibility for everything to be ready, you don't want to waste anyone's time by them having to wait for you* [Nurse, 25 years]

Health care professionals also expressed that if a consistent work structure regarding the video-consulted patient rounds was established, it could free up time. They predicted that it could reduce the number of phone calls from relatives in the evening shift and this would result in freeing up more time to care for the patient:

> *Because relatives can participate and therefore inform the whole family, it could save the nurse on the evening shift many phone calls, where she retells the written journal text from earlier that day.* [Nurse, 60 years]

In both focus groups, health care professionals stressed that changing workflows was difficult because of their already high workload and in spite of introducing an *easy-to-use technology*, it required them to prioritize their assignments differently during the day.

Discussion

Principal Findings

This study identified a number of possibilities and barriers related to adopting a telehealth solution for including relatives. The overall finding was that health care professionals experienced the technology as a way to facilitate involvement of the patient's family, although the relatives were not physically present. Moreover, it raised questions about whether this way of doing patient rounds could embrace the needs of both the patients and relatives. Time and change of work routines involving new technology were recognized as major barriers when implementing new technology involving relatives.

Family-Centered Care

One of this study's main findings was health care professionals agreeing on the positive impact of relatives participating in the patient rounds, which they experienced as a qualification of the conversation. Their virtual presence was able to reduce misunderstandings in the information given about treatment and care, benefitting the patient, the family, and the health care professionals. This is consistent with a study concluding that by having the family join the patient rounds via telemedicine, resolutions of an issue where family input is very important can be expedited [16]. However, another finding in our study showed that health care professionals discussed where the focus should lie, with the patient or with the relative. This statement was met with different attitudes. Some considered the relative equal to the patient, allowing them to take active part in the conversation, whereas others prioritized the perspective of the patient. In line with these findings, a systematic review found that family-centered decisions are highly dependent on how health care professionals recognize the patient and the family as one unit of care, in addition to the communication and attitudes used in their presence [34]. The benefits of using video-consulted patient rounds were the possibilities of providing more family-centered care connected to an increasing understanding of the whole family, although the levels of support by health care professionals varied [13]. Initiation of family participation in patient rounds is often based on the culture and tendencies of the team doing the patient rounds. Focus on the best way of communicating as part of the culture among health care professionals is of great importance in relation to optimizing family participation [35]. In our study, the discussion regarding involving relatives virtually raised diverse topics and attitudes. On one hand, it was experienced as important and useful in treatment and care but, on the other hand, it raised a debate of how much influence the family should have during the patient rounds. Health care professionals stressed that, in future, this would require an agreement on and matching expectations of how to provide family-centered care.

Workload

In this study, both nurses and physicians found the technology easy to use; however, their workload played a significant role and, in some cases, seemed to overshadow the benefits of using the technology. They experienced it as *just one more thing* they had to offer in an already busy schedule. Moreover, they found that this solution was time-consuming. It required them to prioritize time to learn the new technology, attend the patient rounds at a predetermined time, and invest time in involving the relatives. These findings are compatible with the findings in a scoping review, which yielded 74 studies of adopting electronic health [14]. Lack of time and intense workload were significant barriers to the implementation of telehealth solutions. Telehealth was believed to take health care professionals away

XSL•FO
RenderX

from well-known clinical tasks and place a greater workload onto them, which resulted in resistance to adapt [14]. Several studies argue that endorsement and clear visions from management are necessary if the staff is to be willing to invest extra time and effort in implementing new technologies [11,13]. In this study, the use of technology required a change in the existing work organization of the patient rounds for the health care professionals. This was due to the need for a specifically set time for doing patient rounds and also the involvement of relatives in the conversation. This indicates that many factors can have an impact on the success of technological information systems and that experiences of heavy workload and lack of time affect the possibility of adopting new work routines.

Organization

Our findings revealed that health care professionals experienced organizational changes when faced with implementation of the new technology. They were *controlled* by a prearranged time for attending the patient rounds, combined with different professions, depending on each other to be present at that time. Therefore, accommodating the family perspective was experienced as overruling their rights to organize their own work. It left them reluctant to adapt to the implementation of the technology. Mohammadzadeh et al aimed to explain key considerations in telehealth solutions relating to cancer care. They found that the organizational structure and management provide an environment that can have a strong influence on the adoption of new technology. The organizational culture among health care professionals must contain acceptance of the need to implement the technology [36]. A sense of urgency among employees is the first and most crucial step when transforming organizations and making changes happen [21]. If the staff does not feel a sense of urgency, changes are most likely to fail. To prevent that from happening, a group with enough commitment to lead the change through a clear vision is essential for the implementation to succeed [21]. This is also supported by Rising et al, who suggest the presence of clinician champions in the department to take care of concerns raised by late or nonadopters [11].

Double Change

Furthermore, the study discovered that health care professionals experienced the technology as easy to use, but the screen seemed to draw their attention and they would start focusing on the relative instead of the patient. They described an experience of indirectly having to choose between the patient and the relative and allowed the screen to affect them in that way. This was due to the presence of the relative through the screen that highlighted the relative and suppressed the focus on the patient. According to Ihde, technology does nothing in itself, but in the interaction with human beings, the technology can, as in our study, *highlight* the relative as the contact is mediated through technology [37]. Technologies are mediators of human experiences and practices, where a human's behavior and attitude toward the technology will shape its usage patterns [38]. With this understanding, the technology in our study only represents a solution to facilitate a conversation, but the use of the technology was influenced and dependent on the perception of health care professionals who attended the patient rounds. The health care professionals emphasized that it required time to learn to use a new communication tool. This is in line with what Danbjørg et al found in their study of nurses' experiences of using a new app to support parents after early discharge. They found that the nurses needed time to adapt to new ways of communicating when technology is involved. With time, they enhanced possibilities not only in using the technology as a communication tool but also in allowing observations and emotional support [39]. In this study the technology itself seemed to be the smallest part but the health care professionals experienced changes when adopting the new technical skills required for handling the technology. Changes were also observed as health care professionals had to rethink the way they usually interacted with relatives, and heavy workloads seemed to influence the amount of resistance that occurred when adopting the family-orientated telehealth solution. These factors required an organizational change for both nurses and physicians, affecting their previous possibilities to organize their own working day. What was expected as implementation of an *easy-to-use technology* became an extremely complex process, leading us to think that the complexity of the health care organizations cannot be ignored. It tells us that changing one thing will often bring along many changes, which might end up having extensive consequences for health care professionals in clinical practice. Many of these impacts are difficult to predict, which leads us to the introduction of the following figure (Figure 1). The figure illustrates elements regarding the complexity of implementing new technology.

Limitations and Strength

Our study is limited in using IPA as a research approach, because it requires that all the participants involved possess the necessary language as a tool to explain their experiences. Therefore, limitations in using IPA rely upon the validity of language. A small sample size often raises questions concerning representativeness and transferability of findings. Nevertheless, in IPA research, the aim is not to investigate what occurs in *all* settings, but to instead focus on the perceptions and understandings of a specific group within *their* setting [33,40]. Including participants who are colleagues could have affected the discussions by some participants, suppressing their opinions because of the internal dynamics among them. Both focus groups allowed every participant speaking time, and they all contributed in the interactions, which was the intended advantage of using focus groups. All quotes used in the Results section were member-checked by the individual participants, as a technique to improve accuracy and validity [41]. Mixing the methods and focus group interviews, as well as field observational studies, are considered the strength of this study. It allows statements from health care professionals to be supported and validated by observations leading to a broader understanding of the investigated phenomenon. This study also provided knowledge in an area with limited in-depth research—the use of telehealth with relatives in patient rounds.

XSL•FO
RenderX

Figure 1. Illustration of the complexity in health care systems.



Conclusions and Implications for Clinical Practice

In conclusion, this study illustrated that telehealth in relation to relatives is experienced by health care professionals as a possible way to facilitate participation of the patient's relatives in patient rounds. Their virtual presence at patient rounds reduces misunderstandings and improves decisions about treatment and care. It became clear that introducing telehealth with relatives required changes making an impact on health care professionals in more ways than one. We identified a double change, not only a change in using technology but also a change that had consequences for the health care professionals' work routines in connection with workload, culture, and organization.

This study gave insight into the implementation of telehealth involving relatives that relies upon many factors because of the complexity in the health care organization. We learned that changes must be planned carefully into the existing organization structures, as a small change affects changes on many levels across the organization. Moreover, telehealth with relatives improves access to health care and quality in patient rounds because it may empower patients and relatives by providing more individualized information, treatment, and care.

Acknowledgments

This work was supported financially by the AgeCare Academy of Geriatric Cancer Research and the Foundation of Innovation, Odense University Hospital, Denmark.

The authors would like to thank the nurses and physicians who spent their time on participating in the focus group interviews and made this study possible. The authors would also like to convey special thanks to Vickie Svane Kristensen for proofreading.

Conflicts of Interest

None declared.

References

1. Northouse L, Williams AL, Given B, McCorkle R. Psychosocial care for family caregivers of patients with cancer. J Clin Oncol 2012 Apr 10;30(11):1227-1234. [doi: 10.1200/JCO.2011.39.5798] [Medline: 22412124]
2. Heynsbergh N, Heckel L, Botti M, Livingston PM. Feasibility, useability and acceptability of technology-based interventions for informal cancer carers: a systematic review. BMC Cancer 2018 Dec 02;18(1):244 [FREE Full text] [doi: 10.1186/s12885-018-4160-9] [Medline: 29499663]

XSL•FO
RenderX

3. Laidsaar-Powell R, Butow P, Bu S, Charles C, Gafni A, Fisher A, et al. Family involvement in cancer treatment decision-making: a qualitative study of patient, family, and clinician attitudes and experiences. Patient Educ Couns 2016 Dec;99(7):1146-1155. [doi: 10.1016/j.pec.2016.01.014] [Medline: 26873544]
4. Dieperink KB, Coyne E, Creedy DK, Østergaard B. Family functioning and perceived support from nurses during cancer treatment among Danish and Australian patients and their families. J Clin Nurs 2018 Jan;27(1-2):e154-e161. [doi: 10.1111/jocn.13894] [Medline: 28544339]
5. Kahriman F, Zaybak A. Caregiver burden and perceived social support among caregivers of patients with cancer. Asian Pac J Cancer Prev 2015;16(8):3313-3317 [FREE Full text] [Medline: 25921137]
6. Kim Y, Kashy DA, Spillers RL, Evans TV. Needs assessment of family caregivers of cancer survivors: three cohorts comparison. Psychooncology 2010 Jun;19(6):573-582. [doi: 10.1002/pon.1597] [Medline: 19582798]
7. Wright LM, Leahey M. Nurses And Families: A Guide To Family Assessment And Intervention. Philadelphia: FA Davis Company; 2013:1-151.
8. Shin JY, Kang TI, Noll RB, Choi SW. Supporting caregivers of patients with cancer: a summary of technology-mediated interventions and future directions. Am Soc Clin Oncol Educ Book 2018 May 23(38):838-849 [FREE Full text] [doi: 10.1200/EDBK_201397] [Medline: 30231412]
9. Collinge W, Kahn J, Walton T, Kozak L, Bauer-Wu S, Fletcher K, et al. Touch, caring, and cancer: randomized controlled trial of a multimedia caregiver education program. Support Care Cancer 2013 May;21(5):1405-1414 [FREE Full text] [doi: 10.1007/s00520-012-1682-6] [Medline: 23262808]
10. Fuentes C, Hernandez C, Escobedo L, Herskovic V, Tentori M. Promoting self-reflection of social isolation through persuasive mobile technologies: the case of mother caregivers of children with cancer. Int J Hum Comput Interact 2014 Jul 23;30(10):802-814. [doi: 10.1080/10447318.2014.927279]
11. Rising KL, Ricco JC, Printz AD, Hoon Woo S, Hollander JE. Virtual rounds: observational study of a new service connecting family members remotely to inpatient rounds. Gen Int Med Clin Innov 2016;1(3):44-47. [doi: 10.15761/GIMCI.1000113]
12. Jackson JD, Cotton SE, Bruce Wirta S, Proenca CC, Zhang M, Lahoz R, et al. Burden of heart failure on caregivers in China: results from a cross-sectional survey. Drug Des Devel Ther 2018;12:1669-1678 [FREE Full text] [doi: 10.2147/DDDT.S148970] [Medline: 29922041]
13. Stelson EA, Carr BG, Golden KE, Martin N, Richmond TS, Delgado MK, et al. Perceptions of family participation in intensive care unit rounds and telemedicine: a qualitative assessment. Am J Crit Care 2016 Dec;25(5):440-447 [FREE Full text] [doi: 10.4037/ajcc2016465] [Medline: 27587425]
14. de Grood C, Raissi A, Kwon Y, Santana MJ. Adoption of e-health technology by physicians: a scoping review. J Multidiscip Healthc 2016;9:335-344 [FREE Full text] [doi: 10.2147/JMDH.S103881] [Medline: 27536128]
15. Epstein EG, Sherman J, Blackman A, Sinkin RA. Testing the feasibility of Skype and FaceTime updates with parents in the neonatal intensive care unit. Am J Crit Care 2015 Jul;24(4):290-296 [FREE Full text] [doi: 10.4037/ajcc2015828] [Medline: 26134328]
16. Yager PH, Clark M, Cummings BM, Noviski N. Parent participation in pediatric intensive care unit rounds via telemedicine: feasibility and impact. J Pediatr 2017 Jun;185:181-6.e3. [doi: 10.1016/j.jpeds.2017.02.054] [Medline: 28363361]
17. Bateman EH, Keef DM. How can eHealth enhance adherence to cancer therapy and supportive care? Srp Arh Celok Lek 2016;144(1-2):116-121 [FREE Full text] [Medline: 27276871]
18. Kirkovits T, Schinkoethe T, Drewes C, Gehring C, Bauerfeind I, Harbeck N, et al. eHealth in modern patient-caregiver communication: high rate of acceptance among physicians for additional support of breast cancer patients during long-term therapy. JMIR Cancer 2016 Sep 19;2(2):e14 [FREE Full text] [doi: 10.2196/cancer.5132] [Medline: 28410173]
19. Scott KC, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
20. Ross J, Stevenson F, Lau R, Murray E. Exploring the challenges of implementing e-health: a protocol for an update of a systematic review of reviews. BMJ Open 2015;5(4):e006773 [FREE Full text] [doi: 10.1136/bmjopen-2014-006773] [Medline: 25869685]
21. Kotter JP. Leading Change. Boston: Harvard Business Review Press; 2012.
22. Alpert JM, Morris BB, Thomson MD, Matin K, Brown RF. Implications of patient portal transparency in oncology: qualitative interview study on the experiences of patients, oncologists, and medical informaticists. JMIR Cancer 2018 Mar 26;4(1):e5 [FREE Full text] [doi: 10.2196/cancer.8993] [Medline: 29581090]
23. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
24. Polinski JM, Barker T, Gagliano N, Sussman A, Brennan TA, Shrank WH. Patients' satisfaction with and preference for telehealth visits. J Gen Intern Med 2016 Mar;31(3):269-275 [FREE Full text] [doi: 10.1007/s11606-015-3489-x] [Medline: 26269131]
25. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]

XSL•FO
RenderX

26. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
27. Smith JA, Flowers P, Larkin M. Interpretative Phenomenological Analysis Theory, Method and Research. London: SAGE Publications Ltd; 2009:1-203.
28. Green J, Thorogood N. Qualitative Methods For Health Research. London: SAGE Publications Ltd; 2014.
29. Kvale S, Brinkmann S. Interview. København: Hans Reitzels Forlag; 2009.
30. Spradley JP. Participant Observation. New York: Holt, Renehart And Winstone; 1980.
31. Halkier B. Focus Groups. In: Qualitative methods: a Basic Book. København: Hans Reitzels Forlag; 2015:137-151.
32. Northern Nurses' Federation. Northern Nurses' Federation. 2003. Ethical guidelines for nursing research in the Nordic countries URL: https://dsr.dk/sites/default/files/479/ssns_etiske_retningslinjer_0.pdf [accessed 2019-01-31] [WebCite Cache ID 75pk7roN9]
33. Smith JA, Osborn M. Interpretative phenomenological analysis. Great Britain: SAGE Publications Ltd; 2008.
34. Visser M, Deliens L, Houttekier D. Physician-related barriers to communication and patient- and family-centred decision-making towards the end of life in intensive care: a systematic review. Crit Care 2014 Nov 18;18(6):604 [FREE Full text] [doi: 10.1186/s13054-014-0604-z] [Medline: 25403291]
35. Jacobowski NL, Girard TD, Mulder JA, Ely EW. Communication in critical care: family rounds in the intensive care unit. Am J Crit Care 2010 Sep;19(5):421-430 [FREE Full text] [doi: 10.4037/ajcc2010656] [Medline: 20810417]
36. Mohammadzadeh N, Safdari R, Rahimi A. Cancer care management through a mobile phone health approach: key considerations. Asian Pac J Cancer Prev 2013;14(9):4961-4964 [FREE Full text] [Medline: 24175759]
37. Ihde D. Technology and the Lifeworld: From Garden to Earth. USA: Indiana University Press; 1990.
38. Ihde D. Heidegger's Technologies: Postphenomenological Perspectives. USA: Fordham Univ Press; 2010.
39. Danbjørg DB, Wagner L, Kristensen BR, Clemensen J. Nurses' experience of using an application to support new parents after early discharge: an intervention study. Int J Telemed Appl 2015;2015:851803 [FREE Full text] [doi: 10.1155/2015/851803] [Medline: 25699079]
40. Noon EJ. An interpretative phenomenological analysis of the barriers to the use of humour in the teaching of childhood studies. J Perspect Appl Acad Pract 2017 May 10;5(3):-. [doi: 10.14297/jpaap.v5i3.255]
41. Lincoln YS, Guba EG. Naturalistic Inquiry. USA: SAGE Publications; 1985.

Abbreviations

IPA: interpretative phenomenological analysis

Edited by G Eysenbach; submitted 23.10.18; peer-reviewed by J Alpert, N Mohammadzadeh; comments to author 13.12.18; revised version received 21.12.18; accepted 31.12.18; published 25.03.19.

Please cite as:
Østervang C, Vestergaard LV, Dieperink KB, Danbjørg DB
Patient Rounds With Video-Consulted Relatives: Qualitative Study on Possibilities and Barriers From the Perspective of Healthcare Providers
J Med Internet Res 2019;21(3):e12584
URL: http://www.jmir.org/2019/3/e12584/
doi:10.2196/12584
PMID:30907746

©Christina Østervang, Lene Vedel Vestergaard, Karin Brochstedt Dieperink, Dorthe Boe Danbjørg. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessment of the Relationship Between Ambient Temperature and Home Blood Pressure in Patients From a Web-Based Synchronous Telehealth Care Program: Retrospective Study

Ching-Chang Huang[1,2], MD; Ying-Hsien Chen[1,2], MD; Chi-Sheng Hung[1,2], MD, PhD; Jen-Kuang Lee[1,2], MD, PhD; Tse-Pin Hsu[1,3], MSN; Hui-Wen Wu[1,3], MSN; Pao-Yu Chuang[1,3], MSN; Ming-Fong Chen[1,2], MD, PhD; Yi-Lwun Ho[4], MD, PhD

[1]Telehealth Center, National Taiwan University Hospital, Taipei, Taiwan

[2]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[3]Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan

[4]Graduate Institute of Clinical Medicine, Division of Cardiology, Department of Internal Medicine, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan

Corresponding Author:
Yi-Lwun Ho, MD, PhD
Graduate Institute of Clinical Medicine, Division of Cardiology
Department of Internal Medicine
National Taiwan University Hospital and National Taiwan University College of Medicine
7 Chung-Shan South Road
Taipei, 10002
Taiwan
Phone: 886 223123456 ext 66373
Fax: 886 223216938
Email: ylho@ntu.edu.tw

Abstract

Background: Decreased ambient temperature significantly increases office blood pressure, but few studies have evaluated the effect of ambient temperature on home blood pressure.

Objective: We aimed to investigate the relationship between short-term ambient temperature exposure and home blood pressure.

Methods: We recruited patients with chronic cardiovascular diseases from a telehealth care program at a university-affiliated hospital. Blood pressure was measured at home by patients or their caregivers. We obtained hourly meteorological data for Taipei (temperature, relative humidity, and wind speed) for the same time period from the Central Weather Bureau, Taiwan.

Results: From 2009 to 2013, we enrolled a total of 253 patients. Mean patient age was 70.28 (SD 13.79) years, and 66.0% (167/253) of patients were male. We collected a total of 110,715 home blood pressure measurements. Ambient temperature had a negative linear effect on all 3 home blood pressure parameters after adjusting for demographic and clinical factors and antihypertensive agents. A 1°C decrease was associated with a 0.5492-mm Hg increase in mean blood pressure, a 0.6841-mm Hg increase in systolic blood pressure, and a 0.2709-mm Hg increase in diastolic blood pressure. This temperature effect on home blood pressure was less prominent in patients with diabetes or hypertension. Antihypertensive agents modified this negative effect of temperature on home blood pressure to some extent, and angiotensin receptor blockers had the most favorable results.

Conclusions: Short-term exposure to low ambient temperature significantly increased home blood pressure in patients with chronic cardiovascular diseases. Antihypertensive agents may modify this effect.

(J Med Internet Res 2019;21(3):e12369) doi:10.2196/12369

KEYWORDS

ambient temperature; home blood pressure; antihypertensive agents; retrospective studies

XSL•FO
RenderX

Introduction

Background

Exposure to extreme temperatures, either high or low, is known to significantly increase mortality, and most temperature-attributable death is due to low, rather than high, ambient temperatures [1,2]. A previous study also documented the short-term effects of ambient temperature on morbidity from noncommunicable diseases [3]. Since it was first reported by Rose in 1961 [4], seasonal variation in blood pressure has been observed by many investigators, and this phenomenon is evident in a variety of patient populations [5,6]. Later studies confirmed that ambient temperature contributes to this phenomenon, and that decreased ambient temperature significantly increased office blood pressure [7-10]. Hypertension is an important risk factor for many cardiovascular diseases [11-13], and thus ambient temperature–associated fluctuations in blood pressure may be the underlying mechanism driving increased mortality and morbidity.

Recently, a Japanese study showed that home blood pressure (HBP) may be an even more important parameter than office blood pressure for predicting future coronary artery disease (CAD) and stroke events [14]. Few studies have investigated seasonal variations in HBP, and the existing studies have been on a small scale and examined elderly normotensive participants as the patient populations [15,16]. A recent large-scale Japanese population study using a Web-based platform also confirmed seasonal variations in HBP [17].

Objective

We aimed to investigate the relationship between short-term ambient temperature exposure and HBP in patients with chronic cardiovascular diseases using a Web-based platform.

Methods

Study Design

This was a single-center, retrospective study and was approved by the Institutional Review Board of National Taiwan University Hospital, Taipei, Taiwan. We obtained informed consent from all participants.

Patient Population

We conducted this study from January 2009 to December 2013 at the Telehealth Center of National Taiwan University Hospital. Patients older than 20 years with chronic cardiovascular diseases who were receiving telehealth care at our telehealth center were enrolled in the study group. Chronic cardiovascular diseases included (1) CAD with or without percutaneous coronary intervention, (2) prior myocardial infarction (MI), (3) heart failure, (4) peripheral artery disease, (5) prior stroke, and (6) hypertension. Patients who did not reside in Taipei during the study period were excluded. The decision on whether to receive the telehealth program was made by patients, their caregivers, or both.

Telehealth Care Program

According to the classification of 4 generations of telemedicine in heart failure proposed by Anker et al [18], the telehealth care program used in this study was a fourth-generation telehealth program: it was a synchronized and integrated remote management program for chronic medical conditions. The internet-based platform was developed by the Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taiwan. The program details have been reported previously [19-22]. Briefly, the telehealth program consisted of 4 key components: (1) biometric data, including single-lead electrocardiography, blood pressure, heart rate, and oxygen saturation, were transmitted from the patients' devices to our telehealth center daily and on demand; (2) nurse case managers telephoned patients or caregivers daily and on demand for communication and health promotion; (3) full-time nurse case managers and cardiologists were in charge of patient care 24 hours a day; and (4) long-term medications and management were discussed with the patients' primary care physicians after any acute event. The telehealth program bridged acute care and home care and emphasized education, prevention, and early detection of clinical deterioration.

Data Collection

We obtained all demographic and clinical data, including the prescription of antihypertensive agents, from the electronic database of the hospital. Chronic disease diagnosis was based on the electronic database. Blood pressure and other biometric parameters were measured at home by patients or their caregivers, and data were instantaneously transmitted to the hospital server for storage and analysis. In general, patients and caregivers were advised to take measurements twice daily, with the first measurement being in the morning before eating or taking any medications and the second being in the evening. However, each patient chose their own time of day and time interval to measure their blood pressure based on their own habits. The specific model of sphygmomanometer used in this study was the AViTA BPM65ZB (AViTA Corporation, New Taipei City, Taiwan), which is an electronic digital upper arm blood pressure monitor. We obtained hourly meteorological data for Taipei (temperature, relative humidity, and wind speed) for the same time period as the study from the Central Weather Bureau, Taiwan.

Statistical Analysis

We performed statistical analyses using R version 3.4.2 software (R Foundation). For the statistical tests, we considered a 2-sided $P \leq .05$ to be statistically significant. The distributional properties of continuous variables are expressed as mean (standard deviation), and categorical variables are presented as frequency and percentage. In the univariate analysis, we examined the differences in the distributions of continuous variables and categorical variables between male and female participants using Wilcoxon rank sum tests and Fisher exact tests, respectively.

We conducted the multivariate analysis by fitting multiple linear regression models to estimate the adjusted effects of age, sex, comorbidities, heart rate, antihypertensive drugs, seasons,

XSL•FO
RenderX

meteorological factors (see below), and other predictors on 3 continuous outcomes: (1) mean blood pressure (MBP), (2) systolic blood pressure (SBP), and (3) diastolic blood pressure (DBP). Since antihypertensive drug use and meteorological factors varied over time, we defined and included the following 2 groups of time-dependent covariates in our linear regression analyses:

- Antihypertensive drugs: of the 6 frequently used categories of antihypertensive drugs (angiotensin converting enzyme inhibitors [ACEIs], angiotensin receptor blockers [ARBs], calcium channel blockers [CCBs], alpha-blockers [ABs], beta-blockers [BBs], and diuretics), we evaluated the antihypertensive drug categories and the number of categories of antihypertensive drugs used on the day of blood pressure measurement.
- Meteorological factors: we considered hourly averaged temperature, relative humidity, and wind speed within the hour of blood pressure measurement; daily highest and lowest temperatures; highest and lowest temperatures in the 12 hours prior to blood pressure measurement; highest and lowest temperatures in the 24 hours prior to blood pressure measurement; difference between the highest and lowest daily temperature; difference between the highest and lowest temperature in the 12 hours prior to blood pressure measurement; and difference between the highest and lowest temperature in the 24 hours prior to blood pressure measurement.

More statistical details can be found in Multimedia Appendix 1.

Results

Patient Population and Demographics

From January 2009 to December 2013, a total of 253 patients who participated in the Telehealth Care Program of National Taiwan University Hospital who had complete demographic and clinical data were enrolled in the study. We collected a total of 110,715 blood pressure measurements in the database during the study period. The demographic and clinical data are summarized in Table 1 (per patient) and Table 2 (per measurement). Briefly, the mean patient age was 70.28 (SD 13.79) years, and 66.0% (167/253) of patients were male. Hypertension was diagnosed in 61.7% (156/253) of patients, diabetes mellitus in 32.8% (83/253), heart failure in 30.4% (77/253), prior stroke in 14.2% (36/263), and CAD in 57.3% (145/253).

Univariate Analysis

Table 1 details the demographic and clinical characteristics of the 253 patients. The distributional parameters of most variables did not differ statistically between male (n=167) and female (n=86) patients except that male patients had a younger mean age (68.73, SD 13.54 years vs 73.31, SD 13.84 years; *P*=.003), higher proportions of CAD without prior MI (72/167, 43.1% vs 32/86, 37.2%; *P*=.006) and CAD with prior MI (24/167, 20.4% vs 7/86, 8.1%; *P*=.006), and a lower proportion with cancer (17/167, 10.2% vs 18/86, 20.9%; *P*=.02). Table 2 details the demographic and clinical characteristics of 110,715 observations from the 253 patients. On average, each patient provided 437.61 repeated blood pressure measurements (male: 452.56; female: 408.57) during the study period. We recorded and counted daily use of any of the following 6 categories of antihypertensive drugs: ACEIs, ARBs, CCBs, ABs, BBs, and diuretics. In the statistical analysis, we considered the number of categories of antihypertensive drugs and the specific categories of antihypertensive drugs used on the day of blood pressure measurement. The distributional parameters of most of the variables differed statistically between male (n=75,578) and female (n=35,137) patients for 11 of the 12 meteorological factors. Given the large number of observations (110,715), the statistical power was sufficiently high to find relatively small differences between male and female patients on the statistical tests.

Meteorological Data

The mean hourly averaged temperature was 23.43 (SD 5.69) °C, and the mean relative humidity was 72.28 (SD 9.95) %. The mean wind speed was 2.47 (SD 1.60) m/s. The mean temperature differences in the past 12 and 24 hours were 3.71 (SD 2.32) °C and 5.72 (2.60) °C, respectively (Table 2).

Antihypertensive Agents

The most frequently used antihypertensive agent class was diuretics (45,526/110,715, 41.12%), followed by ARBs (41,465/110,715, 37.45%), CCBs (29,923/110,715, 27.03%), and BBs (23,107/110,715, 20.87%; Table 2).

Multivariate Analysis

We conducted multivariate analysis by fitting multiple linear regression models to estimate the adjusted effects of age, sex, hypertension, diabetes mellitus, cancer, atrial fibrillation, congestive heart failure, CAD, prior MI, prior stroke, peripheral artery disease, heart rate, the 6 categories of antihypertensive drugs, seasons, the 12 meteorological factors, and other potential predictors on 3 continuous outcomes: (1) SBP, (2) DBP, and (3) MBP. Since patients' use of antihypertensive drugs and meteorological factors varied over time, we defined and computed both as time-dependent covariates in these 3 regression analyses. Tables 3 and 4, and Multimedia Appendix 2 display the 3 fitted multiple final linear regression models of mean SBP (mm Hg), DBP (mm Hg), and MBP (mm Hg), respectively. As Figures 1-4 show, we determined the cutoff point(s) for discretizing continuous covariates with nonlinear effects on mean SBP (mm Hg) objectively using the corresponding generalized additive model plots during the stepwise variable selection procedure. We applied the same approach to regression analyses of DBP (mm Hg) and MBP (mm Hg), and the generalized additive model plots are shown in Figures 5-8 and Multimedia Appendix 3, respectively.

XSL•FO
RenderX

Table 1. Demographic and clinical characteristics of the 253 patients.

Characteristics	Male	Female	All patients	*P* value[a]
Sample size, n (%)	167 (66.0)	86 (34.0)	253 (100)	N/A[b]
Age (years), mean (SD)	68.73 (13.54)	73.31 (13.84)	70.28 (13.79)	.003
Hypertension, n (%)				.34
No	68 (40.7)	29 (33.7)	97 (38.3)	
Yes	99 (59.3)	57 (66.3)	156 (61.7)	
Diabetes mellitus, n (%)				.89
No	113 (67.7)	57 (66.3)	170 (67.2)	
Yes	54 (32.3)	29 (33.7)	83 (32.8)	
Atrial fibrillation, n (%)				.19
No	138 (82.6)	65 (75.6)	203 (80.2)	
Yes	29 (17.4)	21 (24.4)	50 (19.8)	
Congestive heart failure, n (%)				.47
No	119 (71.3)	57 (66.3)	176 (69.6)	
Yes	48 (28.7)	29 (33.7)	77 (30.4)	
Coronary artery disease (CAD) and myocardial infarction (MI), n (%)				.006
No CAD	61 (36.5)	47 (54.7)	108 (42.7)	
CAD without prior MI	72 (43.1)	32 (37.2)	104 (41.1)	
CAD with prior MI	34 (20.4)	7 (8.1)	41 (16.2)	
Prior stroke, n (%)				.57
No	145 (86.8)	72 (83.7)	217 (85.8)	
Yes	22 (13.2)	14 (16.3)	36 (14.2)	
Peripheral artery disease, n (%)				.47
No	152 (91.0)	81 (94.2)	233 (92.1)	
Yes	15 (9.0)	5 (5.8)	20 (7.9)	
Cancer, n (%)				.02
No	150 (89.8)	68 (79.1)	218 (86.2)	
Yes	17 (10.2)	18 (20.9)	35 (13.8)	

[a]Calculated using the Wilcoxon rank sum test for continuous variables and the Fisher exact test for categorical variables.

[b]N/A: not applicable.

Multimedia Appendix 2 presents the fitted multiple linear regression model of mean MBP (mm Hg) as a linear equation for outcome prediction. Hourly averaged temperature (°C) within the hour of blood pressure measurement had a negative relationship with the mean value of MBP with a slope of −0.55 mm Hg/°C (Multimedia Appendix 3), but this negative slope was modified to some extent by the covariates hypertension, diabetes, ACEIs, ARBs, CCBs, and diuretics, which are the so-called effect modifiers in the epidemiological and statistical literature. In Tables 3 and 4, the fitted multiple linear regression models modeling mean SBP (mm Hg) and DBP (mm Hg), respectively, can be interpreted in the same manner.

Effects of Meteorological Factors on Home Blood Pressure

Hourly averaged temperature had a linear negative effect on all 3 HBP parameters, so lower temperature resulted in higher HBP. The temperature difference between the maximum and minimum in the 12 hours prior to blood pressure measurement had a positive effect on MBP and DBP (ie, a larger temperature difference resulted in higher MBP and DBP), but the effect was not significant for SBP. Relative humidity and wind speed had nonlinear effects on HBP, as shown in the generalized additive model plots (Figures 1-4 for SBP, Figures 5-8 for DBP, and Multimedia Appendix 3 for MBP).

XSL•FO
RenderX

Table 2. Demographic and clinical characteristics of the 110,715 blood pressure observations from the 253 patients.

Characteristics	Male (n=167)	Female (n=86)	All patients (n=253)	*P* value[a]
Observations, n (%)	75,578 (68.26)	35,137 (31.74)	110,715 (100)	N/A[b]
Age (years), mean (SD)	72.19 (12.87)	79.82 (10.23)	74.61 (12.61)	<.001
Mean blood pressure (mm Hg), mean (SD)	106.33 (12.95)	108.06 (13.57)	106.88 (13.17)	<.001
Systolic blood pressure (mm Hg), mean (SD)	125.08 (16.43)	127.73 (16.76)	125.92 (16.58)	<.001
Diastolic blood pressure (mm Hg), mean (SD)	68.83 (11.85)	68.72 (11.65)	68.79 (11.79)	.002
Heart rate (beats/min), mean (SD)	68.73 (12.43)	73.32 (11.92)	70.18 (12.46)	<.001
Antihypertensive drugs used on the day of blood pressure measurement				
Number of categories of used antihypertensive drugs, mean (SD)	1.33 (1.08)	1.50 (1.11)	1.39 (1.09)	<.001
1. Angiotensin converting enzyme inhibitors, n (%)				<.001
No	73,250 (96.92)	33,922 (96.54)	107,172 (96.80)	
Yes	2,328 (3.08)	1,215 (3.46)	3,543 (3.20)	
2. Angiotensin receptor blocker, n (%)				<.001
No	49,161 (65.05)	20,089 (57.17)	69,250 (62.55)	
Yes	26,417 (34.95)	15,048 (42.83)	41,465 (37.45)	
3. Calcium channel blockers, n (%)				<.001
No	56,953 (75.36)	23,839 (67.85)	80,792 (72.97)	
Yes	18,625 (24.64)	11,298 (32.15)	29,923 (27.03)	
4. Alpha-blockers, n (%)				<.001
No	66,043 (87.38)	34,707 (98.78)	100,750 (91.00)	
Yes	9,535 (12.62)	430 (1.22)	9,965 (9.00)	
5. Beta-blockers, n (%)				.005
No	59,982 (79.36)	27,626 (78.62)	87,608 (79.13)	
Yes	15,596 (20.64)	7,511 (21.38)	23,107 (20.87)	
6. Diuretics, n (%)				<.001
No	47,422 (62.75)	17,767 (50.57)	65,189 (58.88)	
Yes	28,156 (37.25)	17,370 (49.44)	45,526 (41.12)	
Season, n (%)				<.001
Spring	16,243 (21.49)	7,870 (22.40)	24,113 (21.78)	
Summer	18,929 (25.05)	8,851 (25.19)	27,780 (25.09)	
Fall	22,143 (29.30)	9,963 (28.36)	32,106 (29.00)	
Winter	18,263 (24.16)	8,453 (24.06)	26,716 (24.13)	
Meteorological factors, mean (SD)				
Hourly averaged temperature[c] (°C)	23.45 (5.69)	23.39 (5.69)	23.43 (5.69)	.21
Hourly averaged relative humidity[c] (%)	72.28 (9.96)	72.28 (9.94)	72.28 (9.95)	.95
Hourly averaged wind speed[c] (m/s)	2.48 (1.60)	2.46 (1.60)	2.47 (1.60)	.09
Daily maximum temperature[d] (°C)	26.60 (6.24)	26.59 (6.28)	26.59 (6.25)	.86
Daily minimum temperature[d] (°C)	20.90 (5.05)	20.89 (5.09)	20.90 (5.06)	.83
Difference between daily maximum and daily minimum temperatures[d] (°C)	5.70 (2.65)	5.70 (2.66)	5.70 (2.65)	.83
Maximum temperature in the 12 hours prior to blood pressure measurement (°C)	25.32 (6.12)	25.31 (6.15)	25.32 (6.13)	.72

XSL•FO
RenderX

Characteristics	Male (n=167)	Female (n=86)	All patients (n=253)	*P* value[a]
Minimum temperature in the 12 hours prior to blood pressure measurement (°C)	21.60 (5.12)	21.62 (5.19)	21.60 (5.14)	.37
Difference between maximum and minimum temperatures in the 12 hours prior to blood pressure measurement (°C)	3.73 (2.32)	3.69 (2.30)	3.71 (2.32)	.04
Maximum temperature in the 24 hours prior to blood pressure measurement (°C)	26.73 (6.18)	26.72 (6.22)	26.73 (6.19)	.99
Minimum temperature in the 24 hours prior to blood pressure measurement (°C)	21.01 (5.05)	21.00 (5.09)	21.01 (5.06)	.88
Difference between maximum and minimum temperatures in the 24 hours prior to blood pressure measurement (°C)	5.72 (2.60)	5.72 (2.61)	5.72 (2.60)	.63

[a]Calculated using the Wilcoxon rank sum test for continuous variables and the Fisher exact test for categorical variables.

[b]N/A: not applicable.

[c]Hourly averaged temperature, hourly averaged relative humidity, and hourly averaged wind speed were the readings within the *hour* of blood pressure measurement.

[d]Daily maximum temperature and daily minimum temperature were the readings within the *day* of blood pressure measurement.

Effect of Temperature Modified by Antihypertensive Agents on Home Blood Pressure

Antihypertensive agents significantly affected HBP measurements and may modify the effect of ambient temperature (Tables 3 and 4, and Multimedia Appendix 2) on blood pressure to some extent. Only ARBs, CCBs, and diuretics significantly modified the effect of temperature on SBP (Table 3). Use of ARBs and CCBs alleviated the negative effect of temperature on SBP (from −0.6841 mm Hg/°C to –0.5523 mm Hg/°C for ARBs and to –0.6426 mm Hg/°C for CCBs), but use of diuretics further increased the negative effect of temperature on SBP (–2.2772 mm Hg/°C). For MBP, use of ACEIs, ARBs, and CCBs attenuated the negative effect of temperature, but use of diuretics potentiated the negative effect of temperature (Multimedia Appendix 2). For DBP, all antihypertensive agent classes increased the negative effect of temperature except for ARBs (Table 4). Therefore, only use of ARBs decreased the effect of temperature on all 3 HBP parameters.

Effect of Temperature Modified by Diabetes Mellitus or Hypertension on Home Blood Pressure

Interestingly, in patients with DM or hypertension, the effect of ambient temperature on HBP was less pronounced. Multimedia Appendix 4 shows a conditional effect plot of temperature on SBP for 65-year-old men with a history of CAD.

XSL•FO
RenderX

Table 3. Multivariate analysis of predictors for systolic blood pressure (SBP) by fitting a multiple linear regression model with the stepwise variable selection method.

Covariate	Parameter estimate	Standard error	*t* value	*P* r>|*t* |
Intercept	118.7080	0.4070	291.6461	<.001
Age (years)	0.2346	0.0043	54.6593	<.001
Male	0.9285	0.1134	8.1917	<.001
CAD[a] without MI[b] vs no CAD	2.6639	0.1071	24.8780	<.001
CAD with MI vs no CAD	−1.4940	0.1634	−9.1437	<.001
Congestive heart failure	−4.2668	0.1154	−36.9753	<.001
Peripheral artery disease	−4.2366	0.1861	−22.7713	<.001
Cancer	3.9610	0.1377	28.7566	<.001
Hourly averaged temperature[c] (°C)	−0.6841	0.0092	−74.0842	<.001
Hypertension × hourly averaged temperature[c] (°C)	0.1488	0.0044	33.5195	<.001
Diabetes × hourly averaged temperature[c] (°C)	0.1776	0.0045	39.8797	<.001
Hourly averaged relative humidity[c] ≤66.508% or hourly averaged relative humidity[c] >85.838%	0.7227	0.0945	7.6479	<.001
Number of categories of used antihypertensive drugs[d]	0.8770	0.1110	7.9004	<.001
ARB[d] × hourly averaged temperature[c] (°C)	0.1318	0.0062	21.2490	<.001
ARB[d] × AB[d]	6.9533	0.4036	17.2274	<.001
ARB[d] × BB[d]	2.1025	0.2311	9.0969	<.001
ARB[d] × CCB[d]	−4.4110	0.2163	−20.3951	<.001
CCB[d] × hourly averaged temperature[c] (°C)	0.0415	0.0074	5.6260	<.001
CCB[d] × AB[d]	−10.2323	0.3818	−26.8019	<.001
CCB[d] × BB[d]	−3.4101	0.2855	−11.9423	<.001
Diuretics[d]	−1.7931	0.1632	−10.9847	<.001
Diuretics[d] × AB[d]	6.2103	0.2903	21.3900	<.001

[a]CAD: coronary artery disease.

[b]MI: myocardial infarction.

[c]Hourly averaged temperature, hourly averaged relative humidity, and hourly averaged wind speed were the readings within the *hour* of blood pressure measurement, and thus varied over time.

[d]Daily use of any of the 6 categories of antihypertensive drugs—angiotensin converting enzyme inhibitors, angiotensin receptor blockers (ARB), calcium channel blockers (CCB), alpha-blockers (AB), beta-blockers (BB), and diuretics—were recorded and counted for all 253 patients. The listed antihypertensive drugs were those taken on the *day* of blood pressure measurement, and thus may have varied over time.

XSL•FO
RenderX

Table 4. Multivariate analysis of predictors for diastolic blood pressure (DBP) by fitting multiple linear regression model with the stepwise variable selection method.

Covariate	Parameter estimate	Standard error	*t* value	*P* r>|*t* |
Intercept	98.7106	0.2923	337.7143	<.001
Age (years)	−0.3451	0.0030	−116.6101	<.001
Male	−0.8468	0.0789	−10.7338	<.001
CAD[a] without MI[b] vs no CAD	−0.5255	0.0736	−7.1373	<.001
CAD with MI vs no CAD	−0.8580	0.1143	−7.5038	<.001
Congestive heart failure	−0.6366	0.0812	−7.8369	<.001
Stroke	3.7983	0.0888	42.7668	<.001
Peripheral artery disease	2.1644	0.1281	16.8949	<.001
Cancer	−0.5488	0.0946	−5.8035	<.001
Hourly averaged temperature[c] (°C)	−0.2709	0.0074	−36.5555	<.001
Hypertension × hourly averaged temperature[c] (°C)	0.1151	0.0030	37.8255	<.001
Diabetes × hourly averaged temperature[c] (°C)	0.0631	0.0030	20.6944	<.001
Hourly averaged relative humidity[c] ≤55.953% or hourly averaged relative humidity[c] >80.318%	0.8000	0.0756	10.5798	<.001
Hourly averaged wind speed[c] (m/s)	0.1549	0.0200	7.7402	<.001
Temperature difference between the maximum and minimum in the 12 hours prior to blood pressure measurement >5.362°C	0.3597	0.0798	4.5053	<.001
AB[d] × hourly averaged temperature[c] (°C)	−0.0709	0.0086	−8.2254	<.001
ACEI[d] × hourly averaged temperature[c] (°C)	−0.0439	0.0097	−4.5262	<.001
ARB[d] × hourly averaged temperature[c] (°C)	0.0740	0.0046	16.2087	<.001
ARB[d] × AB[d]	3.7984	0.3048	12.4603	<.001
ARB[d] × BB[d]	6.5456	0.1632	40.0987	<.001
ARB[d] × CCB[d]	−4.0171	0.1504	−26.7066	<.001
ARB[d] × diuretics[d]	−3.4974	0.1401	−24.9670	<.001
BB[d] × hourly averaged temperature[c] (°C)	−0.0346	0.0044	−7.9548	<.001
CCB[d] × hourly averaged temperature[c] (°C)	−0.0141	0.0053	−2.6683	.008
CCB[d] × AB[d]	2.0324	0.2695	7.5417	<.001
CCB[d] × ACEI[d]	−11.8642	0.8327	−14.2472	<.001
CCB[d] × BB[d]	−0.6426	0.1978	−3.2479	.001
CCB[d] × diuretics[d]	3.0658	0.1502	20.4120	<.001
Diuretics[d] × hourly averaged temperature [c] (°C)	−0.0329	0.0044	−7.4992	<.001
Diuretics[d] × AB [d]	−5.6251	0.2490	−22.5946	<.001
Diuretics[d] × ACEI [d]	2.8185	0.3888	7.2483	<.001

[a]CAD: coronary artery disease.

[b]MI: myocardial infarction.

[c]Hourly averaged temperature, hourly averaged relative humidity, and hourly averaged wind speed were the readings within the *hour* of blood pressure measurement, and thus varied over time.

[d]Daily use of any of the 6 categories of antihypertensive drugs—angiotensin converting enzyme inhibitors (ACEI), angiotensin receptor blockers (ARB), calcium channel blockers (CCB), alpha-blockers (AB), beta-blockers (BB), and diuretics—were recorded and counted for all 253 patients. The listed

XSL•FO
RenderX

antihypertensive drugs were those taken on the *day* of blood pressure measurement, and thus may have varied over time.

Figure 1. Generalized additive model plot showing the relationship between systolic blood pressure (SBP) and the hourly averaged temperature.

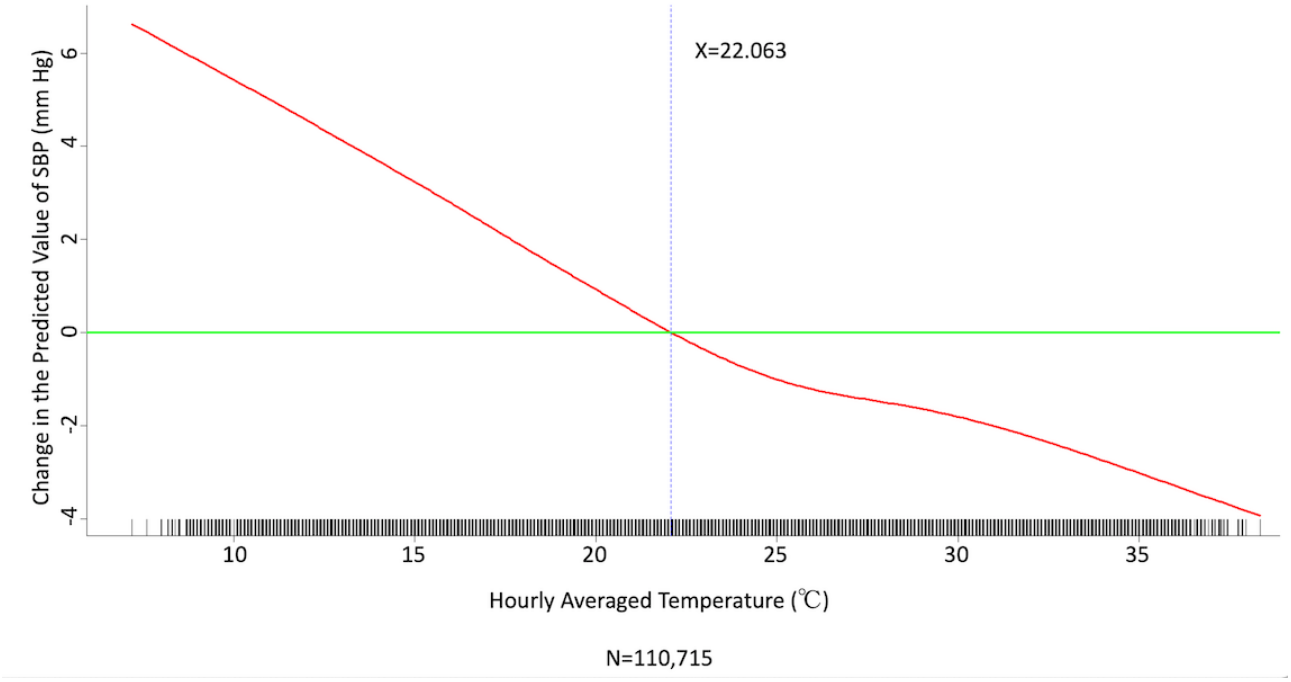
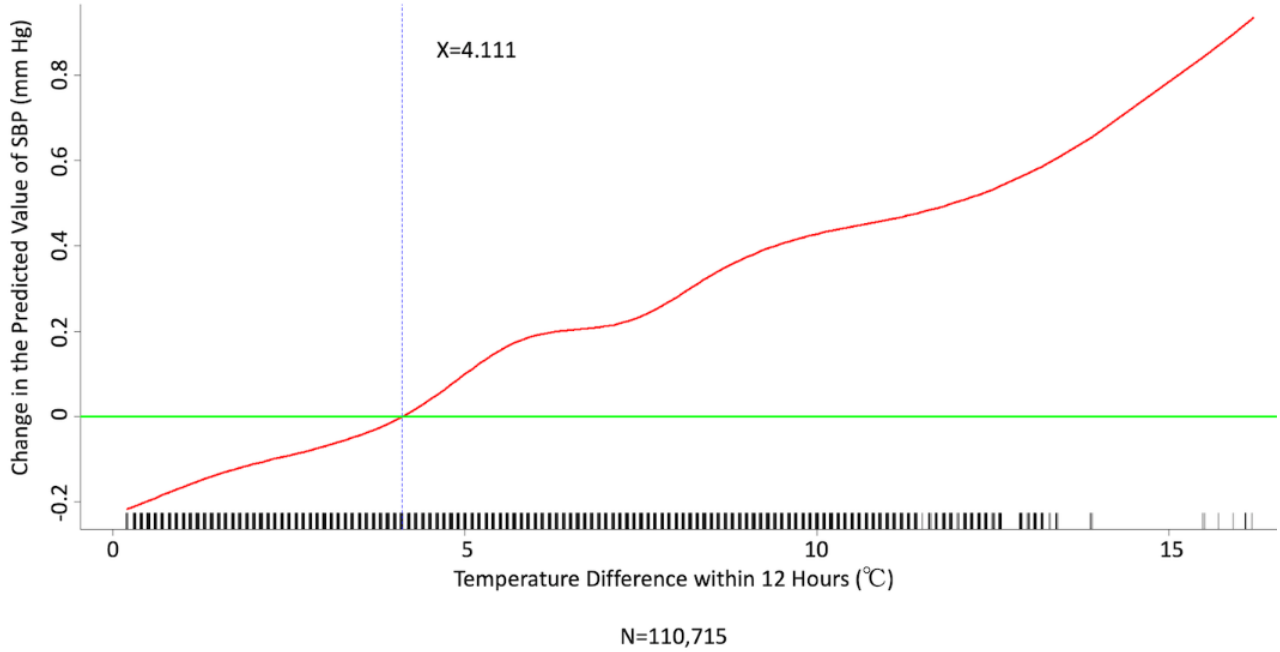


Figure 2. Generalized additive model plot showing the relationship between systolic blood pressure (SBP) and the temperature difference between maximum and minimum in the 12 hours prior to blood pressure measurement.



XSL•FO
RenderX

Figure 3. Generalized additive model plot showing the relationship between systolic blood pressure (SBP) and the hourly averaged relative humidity.

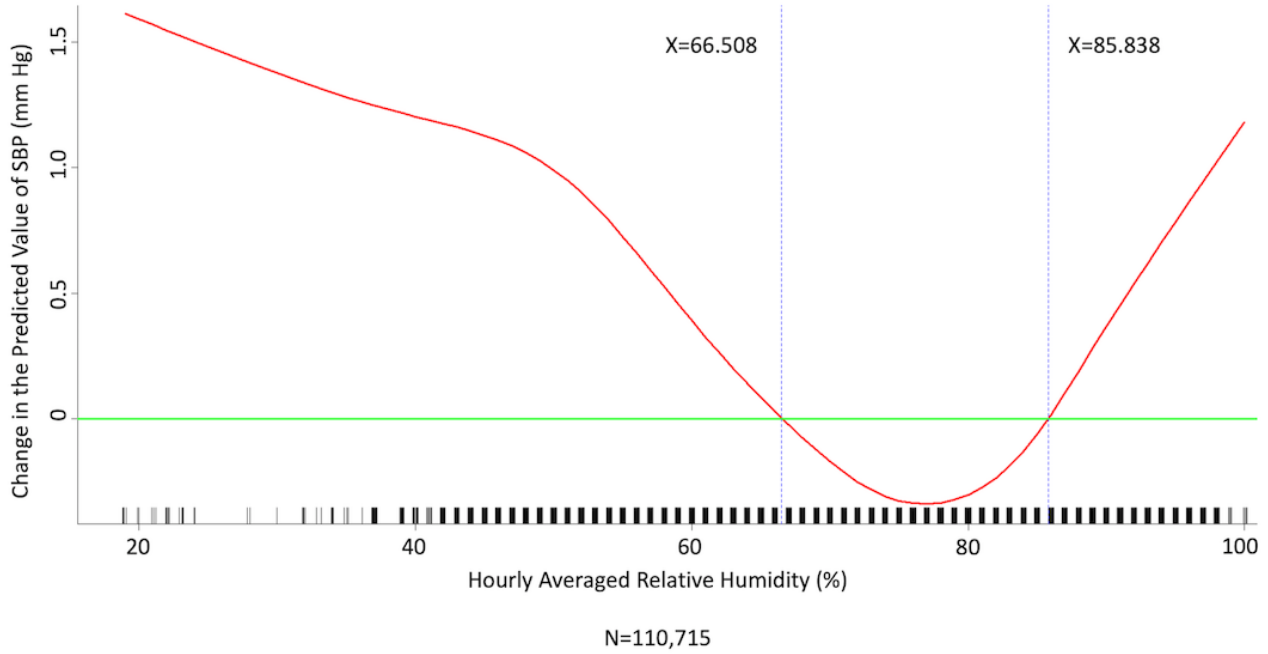
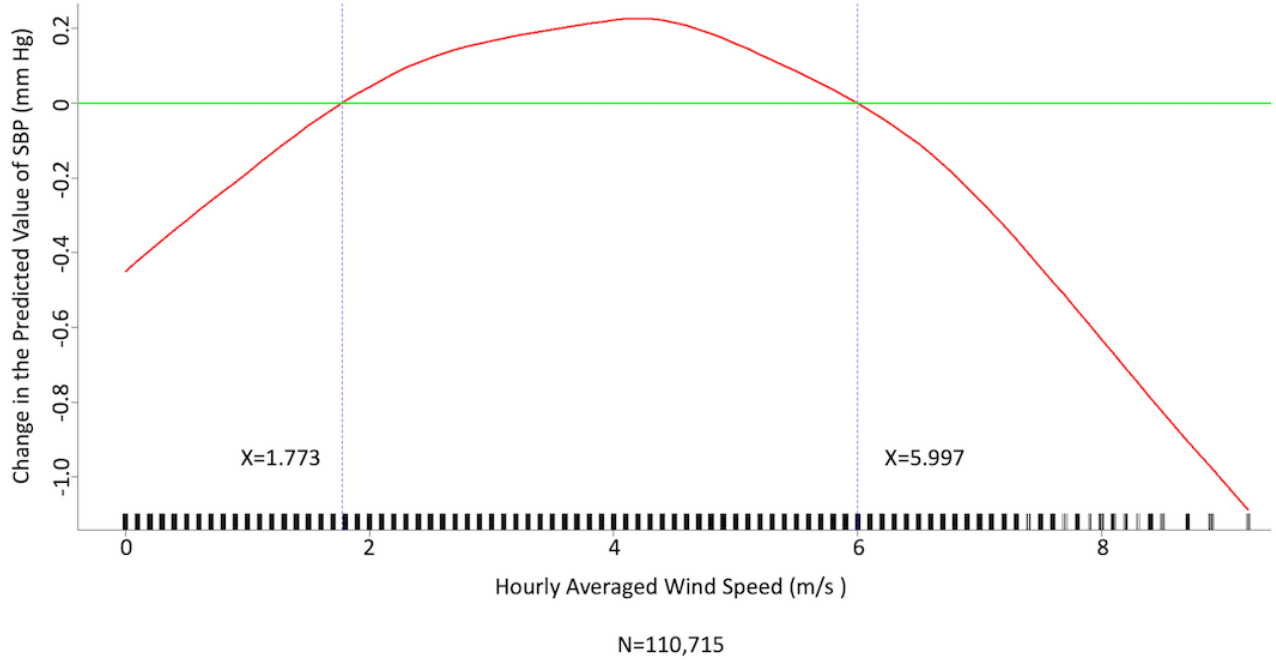


Figure 4. Generalized additive model plot showing the relationship between systolic blood pressure (SBP) and the hourly averaged wind speed.



XSL•FO
RenderX

Figure 5. Generalized additive model plot showing the relationship between diastolic blood pressure (DBP) and the hourly averaged temperature.

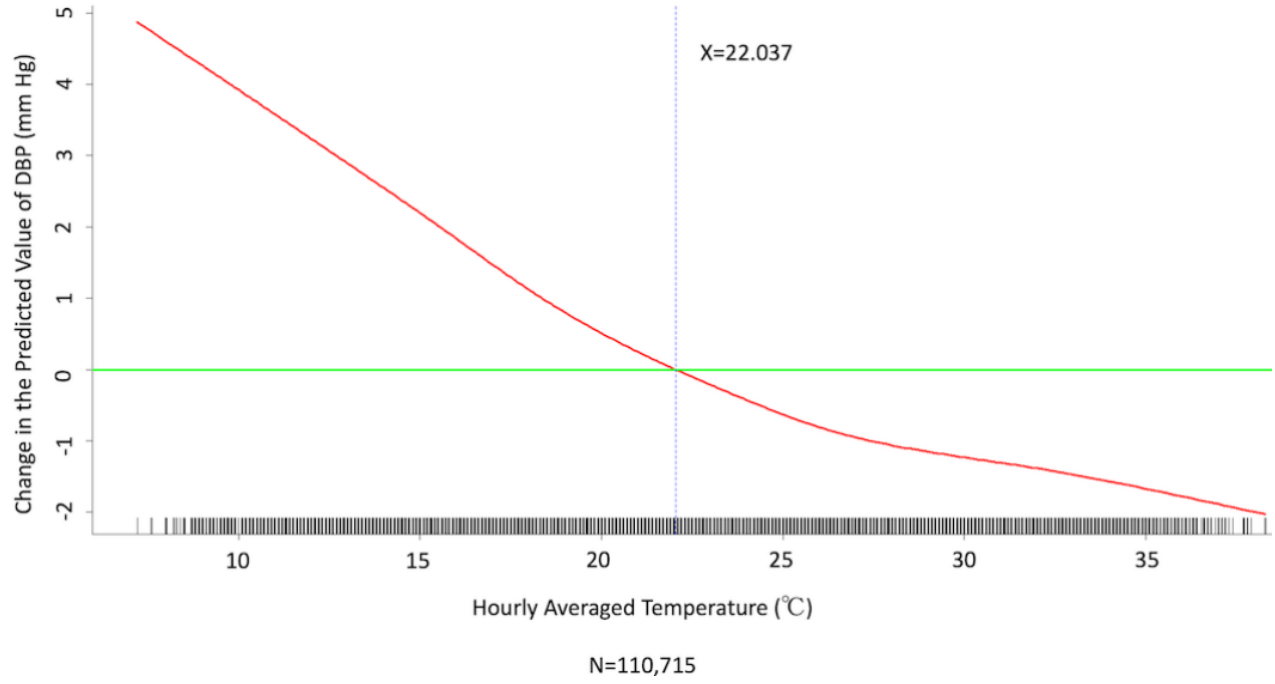
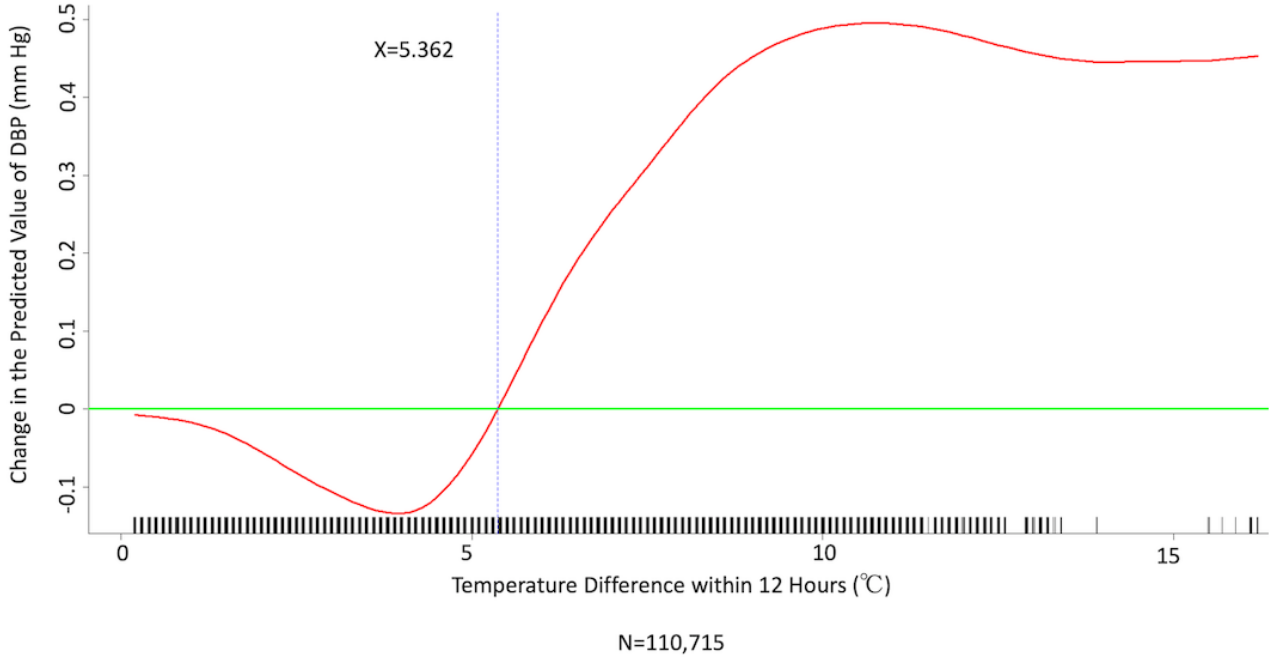


Figure 6. Generalized additive model plot showing the relationship between diastolic blood pressure (DBP) and the temperature difference between maximum and minimum in the 12 hours prior to blood pressure measurement.



XSL•FO
RenderX

Figure 7. Generalized additive model plot showing the relationship between diastolic blood pressure (DBP) and the hourly averaged relative humidity.

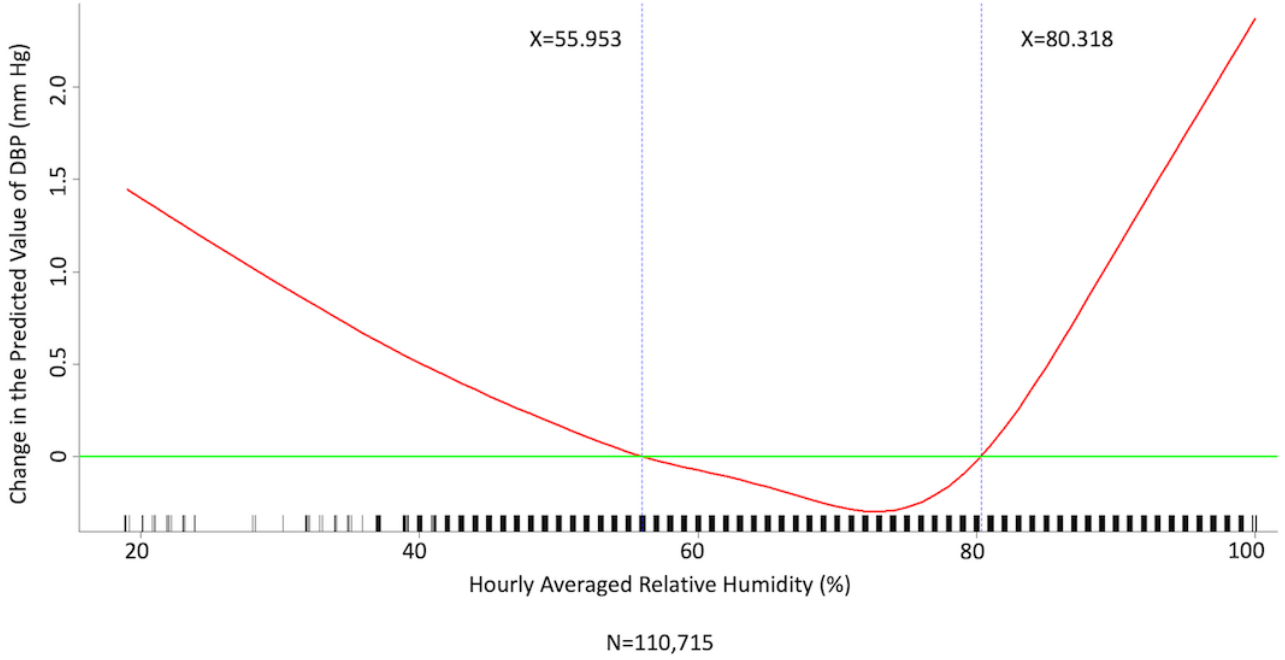
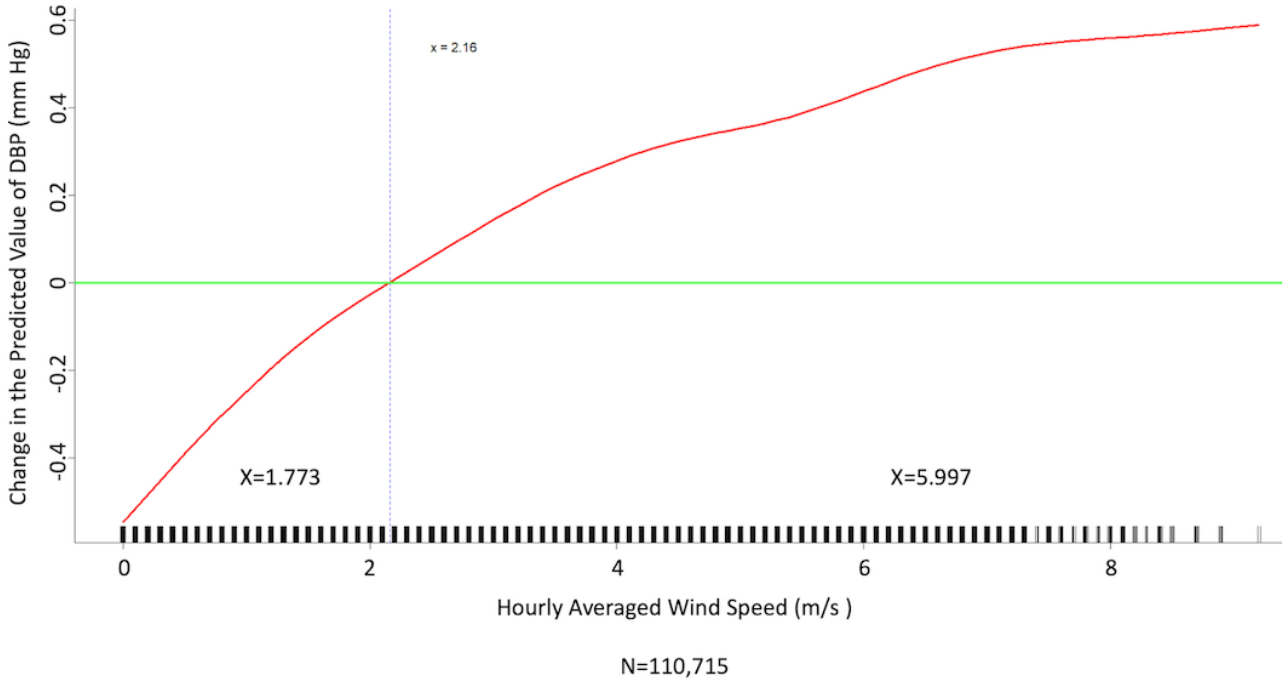


Figure 8. Generalized additive model plot showing the relationship between diastolic blood pressure (DBP) and the hourly averaged wind speed.



Discussion

Effect of Meteorological Factors on Home Blood Pressure

Both extremely high and low temperatures have been shown to increase mortality [2], but the underlying mechanism is not well understood. Seasonal variations in blood pressure have been well reported, and it was assumed to be one of the possible underlying mechanisms contributing to low-temperature–related mortality increases. However, most previous studies have been population or cross-sectional studies.

Hintsala et al reported that even short-term cold exposure increased central aortic blood pressure [23], and the underlying mechanism may be endothelial-dependent vasodilation [24]. HBP was recently found to be a more powerful predictor than office blood pressure of future cardiovascular events [14). To the best of our knowledge, our study is the first to investigate the relationships among demographics, meteorological factors, medications, and HBP at the patient level. The study had a few important findings. HBP was significantly affected by meteorological factors. Of the meteorological factors we investigated, ambient temperature (hourly averaged temperature) was the most important and had a linear inverse relationship with all 3 of the HBP parameters. Kimura et al [15] and Imai

XSL•FO
RenderX

et al [16] reported seasonal HBP variations, but their participants were healthy normotensive elderly individuals and their studies comprised limited patient numbers [15,16]. Iwahori et al [17] also recently confirmed seasonal variations in HBP by using a nationwide Web-based databank in Japan, although effects on individual participants based on repeated measurement could not be investigated, as the study was a population study. Our study confirmed that temperature has an inverse linear effect on HBP in patients with chronic cardiovascular diseases. Moreover, the results showed that both absolute temperature and temperature difference affected HBP. Relative humidity and wind speed also had significant effects on HBP, and the relationships appear to be nonlinear.

Effect of Temperature Modified by Antihypertensive Agents on Home Blood Pressure

In this study, patients with a diagnosis of either diabetes mellitus or hypertension appeared to be more resistant to the effect of temperature on HBP. The true underlying mechanism remains unknown, but medications appear to play a role. Chen el al [10] reported that benazepril attenuated temperature-mediated blood pressure variations. We included the 6 most frequently used antihypertensive agent classes in the multivariate analysis and found that antihypertensive agents significantly modified the effect of temperature on HBP to some extent. However, the effect of each individual drug class was variable. Of the 6 classes of antihypertensive agents, ARBs appeared to have the most favorable outcome, as use of ARBs attenuated the negative effect of temperature on all 3 HBP parameters. On the other hand, use of diuretics appeared to potentiate the negative effect of temperature on HBP, as the slope increased for all HBP parameters. Although diuretics (thiazides) are still among the most frequently used antihypertensive agents in practice and are suggested by the current guidelines as one of the first-line agents [13], this study showed that patients taking diuretics were more sensitive to temperature changes. In this regard, ARBs may be the class of choice for patients whose blood pressure is more sensitive to ambient temperature.

The fact that ACEIs had no significant influence on HBP, in contrast to ARBs, was somewhat surprising. However, the number of patients taking ACEIs was very small, so the analysis in this study may not have had sufficient power to detect the true effect of ACEIs. On the other hand, the proportion of patients taking ARBs and diuretics was the largest of the various classes, and this may have increased the power of the analysis to find an effect.

Impact on Health

As the final multiple linear regression models shown in Tables 3 and 4 suggest, changes to any covariate in the regression model had positive and negative effects on mean SBP and DBP. While the values of the other covariates were held fixed, mean SBP increased by 0.6841 mm Hg (Table 3) and mean DBP increased by 0.2709 mm Hg (Table 4) within an hour due to a 1°C decrease in the hourly averaged environmental temperature. According to a meta-analysis, every 4-mm Hg reduction in SBP and 2-mm Hg reduction in DBP is sufficient to significantly reduce cardiovascular events [25]. The effect of temperature on HBP may thus have a huge impact on health. Additional longer-term outcome studies are warranted to confirm this phenomenon.

Study Limitations

This study was a retrospective registry with a relatively small number of patients. The patients all had chronic cardiovascular diseases with excellent adherence as they participated in a telehealth care program; thus, the study results should be extrapolated to other patient populations with caution. The ambient temperature was outdoor temperature in this study, and thus may have underestimated the true effect of indoor temperature. Previous studies have shown that both indoor and outdoor temperatures are inversely related to blood pressure, and there is a stronger association with indoor temperature than with outdoor temperature [8,26-29]. The true effect of indoor temperature on HBP should thus be more prominent.

We did not include medications other than antihypertensive agents, such as oral antidiabetic drugs and statins, in this analysis. As a result, whether these drugs also modify the effect of temperature on HBP is unknown. We also did not include other environmental factors such as air pollutants in this study. Moreover, although the patients and caregivers were advised to take measurements twice daily, patients and their caregivers chose their own time of day and time intervals to measure blood pressure. There may thus be heterogeneity across patients regarding measurement behavior and potential confounding effects.

Since the repeated measurements of SBP and DBP on the same patient were probably correlated, we could have applied the generalized estimating equations method to obtain robust estimates of standard errors for the estimated regression coefficients of the multiple linear regression models of SBP and DBP. However, we did not do so because the cluster sizes were too big (ie, many repeated measurements on some patients).

Conclusions

Short-term exposure to low ambient temperature significantly increased HBP in patients with chronic cardiovascular diseases, and antihypertensive agents could have modified this effect.

Acknowledgments

We thank Dr Fu-Chang Hu (Institute of Clinical Medicine, National Taiwan University, Taipei, Taiwan) for his support with the statistical analyses. We would like to thank Anthony Abram for editing and proofreading an earlier version of the manuscript.

This study was funded by National Taiwan University Hospital (106-A136).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

More statistical details.

[PDF File (Adobe PDF File), 34KB - jmir_v21i3e12369_app4.pdf]

Multimedia Appendix 2

Multivariate analysis of predictors for mean blood pressure.

[PDF File (Adobe PDF File), 282KB - jmir_v21i3e12369_app1.pdf]

Multimedia Appendix 3

Generalized additive model plots of the relationship between mean blood pressure and meteorological factors.

[PDF File (Adobe PDF File), 514KB - jmir_v21i3e12369_app2.pdf]

Multimedia Appendix 4

Conditional effect plot of temperature on value of systolic blood pressure for 65-year-old men with history of coronary artery disease.

[PDF File (Adobe PDF File), 247KB - jmir_v21i3e12369_app3.pdf]

References

1. The Eurowinter Group. Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. The Eurowinter Group. Lancet 1997 May 10;349(9062):1341-1346. [Medline: 9149695]
2. Gasparrini A, Guo Y, Hashizume M, Lavigne E, Zanobetti A, Schwartz J, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet 2015 Jul 25;386(9991):369-375 [FREE Full text] [doi: 10.1016/S0140-6736(14)62114-0] [Medline: 26003380]
3. Ye X, Wolff R, Yu W, Vaneckova P, Pan X, Tong S. Ambient temperature and morbidity: a review of epidemiological evidence. Environ Health Perspect 2012 Jan;120(1):19-28 [FREE Full text] [doi: 10.1289/ehp.1003198] [Medline: 21824855]
4. Rose G. Seasonal variation in blood pressure in man. Nature 1961 Jan 21;189:235. [Medline: 13743262]
5. Argilés A, Mourad G, Mion C. Seasonal changes in blood pressure in patients with end-stage renal disease treated with hemodialysis. N Engl J Med 1998 Nov 05;339(19):1364-1370. [doi: 10.1056/NEJM199811053391904] [Medline: 9801397]
6. Hopstock LA, Barnett AG, Bønaa KH, Mannsverk J, Njølstad I, Wilsgaard T. Seasonal variation in cardiovascular disease risk factors in a subarctic population: the Tromsø Study 1979-2008. J Epidemiol Community Health 2013 Feb;67(2):113-118. [doi: 10.1136/jech-2012-201547] [Medline: 22859515]
7. Woodhouse PR, Khaw KT, Plummer M. Seasonal variation of blood pressure and its relationship to ambient temperature in an elderly population. J Hypertens 1993 Nov;11(11):1267-1274. [Medline: 8301109]
8. Madsen C, Nafstad P. Associations between environmental exposure and blood pressure among participants in the Oslo Health Study (HUBRO). Eur J Epidemiol 2006;21(7):485-491. [doi: 10.1007/s10654-006-9025-x] [Medline: 16858621]
9. Alpérovitch A, Lacombe J, Hanon O, Dartigues J, Ritchie K, Ducimetière P, et al. Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: the Three-City study. Arch Intern Med 2009 Jan 12;169(1):75-80. [doi: 10.1001/archinternmed.2008.512] [Medline: 19139327]
10. Chen Q, Wang J, Tian J, Tang X, Yu C, Marshall RJ, et al. Association between ambient temperature and blood pressure and blood pressure regulators: 1831 hypertensive patients followed up for three years. PLoS One 2013;8(12):e84522 [FREE Full text] [doi: 10.1371/journal.pone.0084522] [Medline: 24391962]
11. Lawes CMM, Vander Hoorn S, Rodgers A, International Society of Hypertension. Global burden of blood-pressure-related disease, 2001. Lancet 2008 May 03;371(9623):1513-1518. [doi: 10.1016/S0140-6736(08)60655-8] [Medline: 18456100]
12. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1·25 million people. Lancet 2014 May 31;383(9932):1899-1911 [FREE Full text] [doi: 10.1016/S0140-6736(14)60685-1] [Medline: 24881994]
13. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison HC, et al. 2017 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: executive summary: a report of the American College of

XSL•FO
RenderX

Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018 Jun;71(6):1269-1324. [doi: 10.1161/HYP.0000000000000066] [Medline: 29133354]
14. Kario K, Saito I, Kushiro T, Teramukai S, Tomono Y, Okuda Y, et al. Morning home blood pressure is a strong predictor of coronary artery disease: the HONEST Study. J Am Coll Cardiol 2016 Dec 05;67(13):1519-1527 [FREE Full text] [doi: 10.1016/j.jacc.2016.01.037] [Medline: 27150682]
15. Kimura T, Senda S, Masugata H, Yamagami A, Okuyama H, Kohno T, et al. Seasonal blood pressure variation and its relationship to environmental temperature in healthy elderly Japanese studied by home measurements. Clin Exp Hypertens 2010 Jan;32(1):8-12. [doi: 10.3109/10641960902929479] [Medline: 20144067]
16. Imai Y, Munakata M, Tsuji I, Ohkubo T, Satoh H, Yoshino H, et al. Seasonal variation in blood pressure in normotensive women studied by home measurements. Clin Sci (Lond) 1996 Jan;90(1):55-60. [Medline: 8697706]
17. Iwahori T, Miura K, Obayashi K, Ohkubo T, Nakajima H, Shiga T, et al. Seasonal variation in home blood pressure: findings from nationwide web-based monitoring in Japan. BMJ Open 2018 Dec 05;8(1):e017351 [FREE Full text] [doi: 10.1136/bmjopen-2017-017351] [Medline: 29306878]
18. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
19. Chen Y, Lin Y, Hung C, Huang C, Yeih D, Chuang P, et al. Clinical outcome and cost-effectiveness of a synchronous telehealth service for seniors and nonseniors with cardiovascular diseases: quasi-experimental study. J Med Internet Res 2013;15(4):e87 [FREE Full text] [doi: 10.2196/jmir.2091] [Medline: 23615318]
20. Ho Y, Yu J, Lin Y, Chen Y, Huang C, Hsu T, et al. Assessment of the cost-effectiveness and clinical outcomes of a fourth-generation synchronous telehealth program for the management of chronic cardiovascular disease. J Med Internet Res 2014;16(6):e145 [FREE Full text] [doi: 10.2196/jmir.3346] [Medline: 24915187]
21. Hung C, Yu J, Lin Y, Chen Y, Huang C, Lee J, et al. Mortality benefit of a fourth-generation synchronous telehealth program for the management of chronic cardiovascular disease: a longitudinal study. J Med Internet Res 2016 May 13;18(5):e102 [FREE Full text] [doi: 10.2196/jmir.5718] [Medline: 27177497]
22. Hung C, Lee J, Chen Y, Huang C, Wu V, Wu H, et al. Effect of contract compliance rate to a fourth-generation telehealth program on the risk of hospitalization in patients with chronic kidney disease: retrospective cohort study. J Med Internet Res 2018 Jan 24;20(1):e23 [FREE Full text] [doi: 10.2196/jmir.8914] [Medline: 29367185]
23. Hintsala H, Kandelberg A, Herzig K, Rintamäki H, Mäntysaari M, Rantala A, et al. Central aortic blood pressure of hypertensive men during short-term cold exposure. Am J Hypertens 2014 May;27(5):656-664. [doi: 10.1093/ajh/hpt136] [Medline: 23964061]
24. Ejike C, Wang L, Liu M, Wang W, Morishita M, Bard RL, et al. Personal-level exposure to environmental temperature is a superior predictor of endothelial-dependent vasodilatation than outdoor-ambient level. J Am Soc Hypertens 2017 Nov;11(11):746-753.e1. [doi: 10.1016/j.jash.2017.09.006] [Medline: 28989070]
25. Blood Pressure Lowering Treatment Trialists' Collaboration, Sundström J, Arima H, Woodward M, Jackson R, Karmali K, et al. Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data. Lancet 2014 Aug 16;384(9943):591-598. [doi: 10.1016/S0140-6736(14)61212-5] [Medline: 25131978]
26. Barnett AG, Sans S, Salomaa V, Kuulasmaa K, Dobson AJ, WHO MONICA Project. The effect of temperature on systolic blood pressure. Blood Press Monit 2007 Jun;12(3):195-203. [doi: 10.1097/MBP.0b013e3280b083f4] [Medline: 17496471]
27. Kim Y, Kim S, Cheong H, Ahn B, Choi K. Effects of heat wave on body temperature and blood pressure in the poor and elderly. Environ Health Toxicol 2012;27:e2012013 [FREE Full text] [doi: 10.5620/eht.2012.27.e2012013] [Medline: 22888472]
28. Saeki K, Obayashi K, Iwamoto J, Tone N, Okamoto N, Tomioka K, et al. Stronger association of indoor temperature than outdoor temperature with blood pressure in colder months. J Hypertens 2014 Aug;32(8):1582-1589. [doi: 10.1097/HJH.0000000000000232] [Medline: 24937638]
29. Wang Q, Li C, Guo Y, Barnett AG, Tong S, Phung D, et al. Environmental ambient temperature and blood pressure in adults: a systematic review and meta-analysis. Sci Total Environ 2017 Jan 01;575:276-286. [doi: 10.1016/j.scitotenv.2016.10.019] [Medline: 27750133]

Abbreviations

AB: alpha-blocker
ACEI: angiotensin converting enzyme inhibiter
ARB: angiotensin receptor blocker
BB: beta-blocker
CAD: coronary artery disease
CCB: calcium channel blocker
DBP: diastolic blood pressure
HBP: home blood pressure
MBP: mean blood pressure

XSL•FO
RenderX

MI: myocardial infarction
SBP: systolic blood pressure

Edited by G Eysenbach; submitted 30.09.18; peer-reviewed by S Omboni, X Wang, M Singh; comments to author 23.10.18; revised version received 07.02.19; accepted 10.02.19; published 04.03.19.

Please cite as:
Huang CC, Chen YH, Hung CS, Lee JK, Hsu TP, Wu HW, Chuang PY, Chen MF, Ho YL
Assessment of the Relationship Between Ambient Temperature and Home Blood Pressure in Patients From a Web-Based Synchronous Telehealth Care Program: Retrospective Study
J Med Internet Res 2019;21(3):e12369
URL: http://www.jmir.org/2019/3/e12369/
doi:10.2196/12369
PMID:30829574

©Ching-Chang Huang, Ying-Hsien Chen, Chi-Sheng Hung, Jen-Kuang Lee, Tse-Pin Hsu, Hui-Wen Wu, Pao-Yu Chuang, Ming-Fong Chen, Yi-Lwun Ho. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences With Video Consultations in Specialized Palliative Home-Care: Qualitative Study of Patient and Relative Perspectives

Karen Frydenrejn Funderskov[1], RN, MSc; Mette Raunkiær[1], RN, MSc, PhD; Dorthe Boe Danbjørg[2,3], RN, PhD; Ann-Dorthe Zwisler[1], MD, PhD; Lene Munk[4], RN; Mia Jess[1], MSc; Karin Brochstedt Dieperink[1,4], RN, PhD

[1]REHPA - Danish Knowledge Centre for Rehabilitation and Palliative Care, University of Southern Denmark, Nyborg, Denmark

[2]Centre for Innovative Medical Technology, CIMT, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, Odense, Denmark

[3]Quality of Life Research Center, Department of Haematology, Odense University Hospital, Odense, Denmark

[4]Department of Oncology, Odense University Hospital, Odense, Denmark

Corresponding Author:
Karen Frydenrejn Funderskov, RN, MSc
REHPA - Danish Knowledge Centre for Rehabilitation and Palliative Care
University of Southern Denmark
Vestergade 17
Nyborg, 5800
Denmark
Phone: 45 21811011
Email: karen.frydenrejn.funderskov@regionh.dk

Abstract

Background: The work of specialized palliative care (SPC) teams is often challenged by substantial amounts of time spent driving to and from patients' homes and long distances between the patients and the hospitals.

Objective: Video consultations may be a solution for real-time SPC at home. The aim of this study was to explore the use of video consultations, experienced by patients and their relatives, as part of SPC at home.

Methods: This explorative and qualitative study included palliative care patients in different stages and relatives to use video consultations as a part of their SPC between October 2016 and March 2017. Data collection took place in the patients' homes and consisted of participant observations followed by semistructured interviews. Inclusion criteria consisted of patients with the need for SPC, aged more than 18 years, who agreed to participate, and relatives wanting to participate in the video consultations. Data were analyzed with Giorgi's descriptive phenomenological methodology.

Results: A number of patients (n=11) and relatives (n=3) were included and, in total, 86 video consultations were conducted. Patients participating varied in time from 1 month to 6 months, and the number of video consultations per patient varied from 3 to 18. The use of video consultations led to a situation where patients, despite life-threatening illnesses and technical difficulties, took an active role. In addition, relatives were able to join on equal terms, which increased active involvement. The patients had different opinions on when to initiate the use of video consultations in SPC; it was experienced as optional at the initiating stage as well as the final stage of illness. If the video consultations included multiple participants from the SPC team, the use of video consultations could be difficult to complete without interruptions.

Conclusions: Video consultations in SPC for home-based patients are feasible and facilitate a strengthened involvement and communication between patients, relatives, and SPC team members.

(J Med Internet Res 2019;21(3):e10208) doi:10.2196/10208

KEYWORDS

palliative care; relatives; telemedicine; qualitative research

XSL•FO
RenderX

Introduction

Background

Palliative care, according to the World Health Organization (WHO), aims to improve quality of life for patients and their families facing physical, psychosocial, or spiritual problems associated with life-threatening illnesses [1]. Specialized palliative care (SPC) is recommended for people suffering from life-threatening illnesses with complex palliative care needs. In 2015, cancer was the most common cause of death (n=15,953) in Denmark [2]. Internationally and in Denmark, patients with cancer represent the majority of patients receiving SPC [3,4]. In Denmark, from January 2010 to 2015, 9782 cancer patients, who died within this period, were referred to SPC [3].

In Denmark, as in several other European countries, SPC is provided at hospices, special palliative wards at hospitals, or in the patient's own home by multidisciplinary SPC teams in collaboration with community nurses and general practitioners [5]. The work is often made difficult owing to long distances and the amount of time SPC teams spend driving to the patients' homes [6,7]. Many consultations are conducted via telephone; however, the lack of nonverbal communication is a disadvantage as health care professionals overlook signs of treatable conditions or deteriorations [7]. Furthermore, relatives or community nurses usually take over telephone consultations when the patient's illness progresses, which makes it difficult to maintain the relation between the SPC team and patient [7].

Objectives

The increasing use of telehealth in palliative care is generally accepted, given that increased home care is seen as the preferred option by both patients and health care professionals [8-10]. The first randomized controlled trial measuring the effectiveness of weekly SPC teleconsultations was conducted in Holland based on patient-experienced symptom burden [11]. The study concluded that adding weekly teleconsultations compared with usual palliative care led to higher reported symptom scores among patients with advanced cancer but also a high degree of satisfaction with telemedicine. Hoeck et al advice future research to focus on the potential impact of technology as on the patient's sense of well-being and appropriate timing for teleconsultations [11].

A high level of user satisfaction was also reported in a systematic literature review from Chile, and the number of hospital admissions decreased by 66% and number of bed days by 77% after introducing video phone devices and text messaging in palliative care [12,13]. However, Denmark has yet to initiate the use of video consultations in the palliative care setting as only 1 Danish qualitative study has been conducted, where 17 health care professionals discussed the opportunities found in using telemedicine without trying it but expressed a preference for face-to-face contact for optimal communication with the patients [7].

Ahead of this study, a literature review of qualitative studies was conducted to explore how video consultations work between patients and their relatives, and health care professionals in SPC. In total, 8 studies were found. There is an increasing use of telemedicine in palliative care, and it seems to be effective in the care of patients with palliative needs [6,14]. An American study showed that pain management, coordination of care, and technical issues using technology were main topics being discussed between 4 interdisciplinary hospice teams and 12 caregivers for hospice patients during videoconferencing [15]. An early introduction leads to the most benefit; however, the technologies must be reliable, mobile, and easy to use [9,16]. The use of video favors the nonverbal communication of emotions, body language, and facial expressions of suffering and happiness [17]. In total, 18 patients receiving palliative care emphasized that being able to see one another's facial expression and detect situational contexts allows them to be absorbed in a digital connectedness [18]. In the care of complex patients, teleconsultations allowed the health care professionals in primary and in specialized care teams to support each other when a concentrated direct interaction was performed [8].

However, the above studies did not explore the opportunities for patients and their relatives in the direct involvement in SPC with the health care professionals. Therefore, the purpose of this study was to clarify if, when, and how the use of video consultations is feasible in an SPC setting and to explore how patients in their own homes and with complex palliative needs and their relatives experience the use of video consultations.

Methods

Overview

This study was carried out as an explorative, qualitative study with the purpose of exploring the experiences connected to video consultations. A descriptive phenomenological approach and participant observations followed by semistructured interviews were used to investigate patient and relative experiences in the use of video consultations in SPC at home [19].

Video Consultation Intervention

This study was conducted from October 2016 to February 2017 and initiated by an SPC team at the Department of Oncology at a university hospital.

The included patients were consulted in their homes during initial home visits by an SPC team, who asked for consent to participate. The patients and their families were primarily in telephone contact with the SPC team nurse, who then consulted the SPC team physician, when necessary. Approximately once a week, a video consultation was set by the SPC team nurse and the patient. The technical application was a tablet used for one-way calls between patients, relatives, and the SPC team nurse. Each patient kept the tablet throughout the study period or until the terminal stage of illness occurred. Video consultations were initiated by 1 SPC nurse. Video calls were encrypted to secure confidential information. An app was developed for relatives to participate in group consultations. Supervision was available from the company supplying the tablets. A total of 7 tablets were available during the study (Figure 1).

XSL•FO
RenderX

Figure 1. Tablet device used in video consultations in specialized palliative care.



Participants

In this study, 1 SPC team nurse recruited the patients.

The inclusion criteria include the following: patients with need for SPC, aged above 18 years, who agreed to participate and were willing to employ the tablet. Relatives were asked for participation when they occurred in the video consultations.

The exclusion criteria were patients who were too cognitively impaired to enter on their own and relatives who did not participate in video consultations.

Location, age, sex, diagnosis, relatives, and reason(s) for contacting the SPC team professionals were described to illuminate which patients and relatives were suitable for video consultations and how these took place.

Ethical Approval

Informed consent was obtained from all individual participants included in the study. The study is registered at the Danish Data Protection Agency (file number: 2012-52-0018). The date of approval was September 29, 2016. This study was executed according to the Declaration of Helsinki [20].

Data Collection

Participant observations were performed with regard to setting, behavior, and physical aspects, with inspiration from Spradley [21].

Interviews were carried out to gain firsthand information from patients and their relatives when participating in video consultations [22,23].

Participant observations were performed by one of the authors (KFF) during video consultations, directly followed by interviews, based on a semistructured interview guide (Multimedia Appendix 1).

The immediate interview made it possible to question and elaborate on the observed consultation. Image-based data were used to document the setting where participants used video consultations [24]. Method triangulation was used to validate and improve the understanding of participants' experience in using video consultations [25].

Data Analysis

Data analysis comprised 2 datasets: field notes from participant observations and interview transcripts. The observation guide and the interview guide were based on the research questions to create consistency. Interviews were audio recorded,

XSL•FO
RenderX

transcribed ad verbatim, and all data were saved in a secure database, SharePoint.

The 2 datasets were thematically coded and categorized using NVivo-11 (QSR International), and a checklist for consolidated criteria for reporting qualitative research was used for reporting the study [26]. Interview transcripts were analyzed using the phenomenological descriptive approach described in Giorgi's 4 steps [27-29]. The steps have been illustrated in Textbox 1. Examples are provided in Textbox 2.

Field notes from participant observations were thematically coded and compared with the interview transcripts. Main themes arising from the data analysis are described in the following section.

Textbox 1. The 4 steps in Giorgi's descriptive analysis in a qualitative study about video consultations in specialized palliative care in Denmark.

Step 1: To achieve an overall understanding, all transcribed material was read to reach the sense of the whole with maximum openness and by setting preunderstandings aside.

Step 2: When knowledge in the material was reached, a slower rereading was made to conduct meaning units. Text unrelated to the experiences in use of video consultations was deleted.

Step 3: A transformation of meaning units took place to create categories and concepts, which can express the meaning of interests from the researcher's point of view.

Step 4: To investigate the meaning units based on the research questions, to reveal the essential structure, and to synthesize the transformed meaning units into descriptive statements.

Textbox 2. Examples of steps 2, 3, and 4 in the process of analysis as part of the main theme: video consultation strengthens communication despite technical difficulties.

Interview of Patient 3

Step 2: Meaning units described by patients

- When you speak with them on the phone, you have this feeling that they turn up their nose at you, like you are some hypochondriac [...] but when you sit in front of the screen you see their facial expression, they listen to what you have to say.

Step 3: Essential aspects described by researcher

- The patient felt misunderstood during a telephone consultation as the facial expressions were left out.

Step 4: Patients' experiences

- The use of the video consultation made the patient feel safe as the visualization of the specialized palliative care team nurse gave a trustworthy relation

Results

Overview

From September 2016 to March 2017, 14 participants were included. The study period was extended for 1 month as we wished to generate more descriptive data according to multiple video consultations in the study period. Furthermore, the included patients, who wished to continue to use video consultations in their SPC, were given the option to keep the tablet.

In this study, 11 patients with complex palliative care needs (79%) 7 men and 4 women, aged 30 to 68 years and 3 relatives (21%) were included. The total number of video consultations (n=86) varied from 3 to 18 per patient during the study period as the patients were included throughout the period and had different needs for video consultations with the SPC team nurse (Table 1).

A total of 3 main themes emerged from the descriptive data analysis:

1. Becoming an active patient in own care.
2. Video consultation strengthens communication, despite technical difficulties.
3. Gaining access for relatives.

XSL•FO
RenderX

Table 1. Patient (n=11) and relative (n=3) characteristics of Danish participants in an explorative study about video consultations in specialized palliative care.

Patient number and gender	Age, mean (n=59)	Diagnosis	Relative: Yes or No	Relatives included (n=3)	Video consultations (n=86) in the study period per patient (months)	Main topics in video consultation
Patient 1, male	64	Head and neck cancer	Yes	1 spouse, 1 daughter	13 October 2016 to March 2017 (6 months)	Pain relief
Patient 2, female	65	Bile duct cancer	No	—[a]	11 October 2016 to March 2017 (6 months)	Pain relief and nausea
Patient 3, male	63	Rectal cancer	No	—	18 November 2016 to March 2017 (5 months)	Pain relief
Patient 4, male	68	Prostate cancer	Yes	—	7 November 2016 to February 2017 (4 months)	Pain relief and dizziness
Patient 5, male	60	Prostate cancer	Yes	—	4 November 2016 to December 2016 (1 month)	Pain relief
Patient 6, female	64	Ovarian cancer	Yes	—	4 November 2016 to December 2016 (1 month)	Insomnia
Patient 7, male	30	Cystic fibrosis	Yes	—	9 January 2017 to March 2017 (3 months)	Nausea
Patient 8, female	67	Head and neck cancer	No	—	6 February 2017 to March 2017 (2 months)	Increased saliva
Patient 9, female	66	Lung cancer	No	—	6 January 2017 to March 2017 (2 months)	Pain relief and psychosocial
Patient 10, male	59	Lung cancer	Yes	1 spouse	3 February 2017 to March 2017 (2 months)	Dyspnea
Patient 11, male	47	Thymus cancer	No	—	5 February 2017 to March 2017 (2 months)	Pain relief

[a]No relatives participated.

Table 2. Cause of exclusion to participate in video consultations in specialized palliative care in Denmark.

Cause of exclusion	Number of patients (n=18), n (%)
Cognitive impairment due to illness progression	3 (17)
High risk of death before startup	3 (17)
No use of video consultation due to hospital admission	1 (6)
No tablets available	4 (22)
Lack of electronic skills	2 (11)
Suicidal patient	1 (6)
Hearing impairment	1 (6)
No desire for video consultations	3 (17)

Becoming an Active Patient in Own Care

A majority of the patients were familiar with using a tablet from their everyday lives.

In the study, the patients found the tablet easy to use, following an introduction:

> *It´s so easy to use, it´s like she´s sitting right in front of you[...]I really think this is the future.* [Patient 4]

They were pleased that they could see the SPC team nurse on the screen during the video consultations when discussing different issues (Figure 2).

During the participant observations, it was notable that patients were prepared for upcoming video consultations as questions were written down or discussed with the community nurses or relatives:

> *The tablet is placed on the dining table, so the patient's wife and the community can see the screen. Notes were written for questions.* [Patient 5]

Although patients found the use of video consultations feasible in their contact to the SPC team nurse, they did not agree on when to start using the video consultation in relation to their physical conditions. Moreover, 1 patient found the use of video consultation more intense than using the telephone for consultations and did not find its use feasible later in her illness progression:

> *I don´t think the tablet will help me later on as I´m getting worse. At that time, it will be the community nurses who take over because I will have more physical needs. Some things cannot be taken care of on a tablet.* [Patient 6]

XSL•FO
RenderX

Figure 2. Patient using video consultations in specialized palliative home care.



Another patient found telephone consultations more convenient in relation to how her condition was at the time but figured that visual contact would be more beneficial for the SPC team nurse later on:

> *Right now, I could just as easily talk to her (the SPC team nurse) on the phone as I could talk to her using video[...]Perhaps if I felt worse it[the video consultations] would be beneficial[for the SPC team nurse] as she would be able to consider my situation.* [Patient 2]

Video Consultation Strengthens Communication, Despite Technical Difficulties

A total of 3 patients were observed experiencing minor problems finding the correct button for activating the tablet or controlling the volume button:

> *The patient seems insecure in how to start up the tablet but comfortable enough to press all the buttons...* [Patient 8]

They kept pressing different buttons until the right screen appeared or the volume was suitable; however, they did not express any difficulties when using the tablet during interviews:

> *This is great, I have to say[...] there´s no problem with this.* [Patient 8]

A total of 2 patients with limited ability to speak up owing to tracheotomy experienced great support using the visual communication:

> *She sees the colors on my face or when I´m tired or feeling bad.* [Patient 1]

These patients were used to writing everything down when communicating.

The tablet gave patients an opportunity to see the person they spoke to, which made them feel safe. The use of telephone consultations was not unwelcoming, contrary to previous telephone consultations with other health care professionals, which had led to misunderstandings:

> *When you speak with them on the phone, you have this feeling that they turn up their nose at you, like you are some hypochondriac[...]but when you sit in front of the screen, you see their facial expression, and they listen to what you have to say.* [Patient 3]

Video consultation gave patients an option of a clearer communication with the SPC team nurse because it generated

XSL•FO
RenderX

the possibility to ask questions if anything was unclear or misunderstood as facial expressions showed confusion, etc.

Moreover, 1 patient mentioned the number of participants could affect the video consultation if too many joined the consultation:

> *When so many have to join in [the video consultation], you need to know when to be quiet when someone wants to say something[...]which would be more obvious if we sat in front of each other.* [Patient 10]

If multiple participants from the SPC team are attending the video consultation, it must be clear who is speaking so that misunderstandings and interruptions can be avoided.

In 4 cases, out of the 11 participant observations, an unclear image or reduced volume occurred; however, it did not seem to influence the conversations between patients and the SPC team nurse. The tablet showed the image of the SPC team nurse and then they had the consultation by telephone. The patients stated that it was satisfying to have scheduled appointments because it made it easier to plan other things and for relatives who lived in other parts of the country to participate:

> *We have set a date and that´s fine by me* [Patient 4]

Gaining Access for Relatives

Relatives could join video consultations on equal terms with patients. They attended the consultation with the patients at home or from a distance, using an app, which gave them the ability to ask the SPC team nurse about treatments, appropriate times for visiting, etc.

Visualization had great value among relatives who played an active role in medical administration or scheduling appointments with the doctor:

> *I think it´s all right, if something is missing because then I can elaborate.* [Relative 3]

This opportunity to share the video consultation with a relative was found comforting to patients as no one had to make follow-up phone calls if anyone missed out:

> *Well, the names of the pills, forget it! I can´t remember any of them. She[relative]knows what I take and don't take. I feel good about that and I appreciate that she´s there.* [Patient 11]

Furthermore, one relative joining a video consultation through the app on a cell phone was excited about the opportunity it provided. Due to living and working in another part of the country, the app made it possible to join in, which would not be possible over telephone:

> *Especially when he[patient 1] has the type of cancer that makes it difficult to speak. It has a great impact that you can see each other.* [Relative 1]

Discussion

Principal Findings

The study describes experiences of using video consultations among patients and their relatives in SPC at home. Both patients and relatives found visualization of the SPC team nurse comforting; however, if too many participants joined in, it could negatively affect consultations. Different perspectives were described on when to initiate the video consultations. Finally, relatives could participate more actively, which gave patients a sense of security as relatives helped with medication or elaborating on medical conditions.

Embracing the Technology

During this study, patients referred to the use of tablets for video consultations as easy, convenient, and recognizable. This embodied relation, according to a post phenomenological approach as described by Don Ihde, occurs when a device, such as the tablet in use for video consultations, becomes incorporated in people's everyday lives [30].

A transformation of roles was identified as video consultations enabled visual communication between participants, which gave patients a sense of security as the SPC team nurse appeared on the tablet screen. The technology-mediated contact made it possible for all participants to be active in a dialogue instead of being listeners or observers.

A total of 3 patients were observed experiencing minor difficulties activating the tablet; however, they did not express any issues with the item during the interview. This might be caused by awareness of being part of a study carried out by the organization providing their palliative care and due to the researchers' overt roles in their homes [31,32].

The need to use video consultations differed. One patient found it useful at the time whereas another predicted it becoming more convenient as the illness progressed. Stern et al found the use of telecare-like video consultations viable, recommending such intervention early in palliative care. Due to illness progression and the feeling of being overwhelmed by the technology, early introduction was found beneficial for the patients and their relatives [9]. According to van Gurp et al, long-term interaction during teleconsultations results in trusting relationships and a feeling of relief and intimacy [18]. According to Huniche and Olesen, the constitution of health care changes when technology is adopted. The use of tablets constitutes a new method in SPC, and a technologized mediation invites particular ways of taking action [33]. This was shown in this study; the patients' roles were pulled in the direction of cooperation and away from traditional care because they were taking an active role when preparing for video consultations. The patient-relative cooperation was strengthened as relatives could join in, strengthening a supporting role as they could elaborate when patients consulted the SPC team nurse.

Will the Video Consultations Embrace the Palliative Needs?

According to WHO, palliative care should implicate psychosocial and spiritual problems as well as physical problems to relieve suffering [1]. The use of technology can improve health care access and create cost savings. In addition, the chronically ill and elderly may receive care while remaining at home [34]. However, the question is whether WHO's recommendations for palliative care can be met using video consultation [1,5,34]. According to the European Association of Palliative Care, good communication is essential to palliative care, and sometimes painful concerns have to be considered,

XSL•FO
RenderX

which requires time and commitment [35]. In this study, the issues being discussed were mainly medical adjustments due to pain relief (as shown in Table 1) rather than considerations of sensitive subjects widely, which might occur in the palliative care [35]. In addition to this, van Gurp et al showed that SPC team clinicians found the image of the patient a valuable and supportive addition, but they avoided talking about sensitive topics with vulnerable patients as they did not feel sufficiently close to be able to comfort them [18]. The patients in van Gurp et al's study reported that physical distance provided exactly the freedom they needed to open up, after which normal life could be resumed [18].

This study showed that the simultaneous involvement of relatives in video consultations with the patient, or from a distance using an app, was both feasible and satisfying. This is in accordance with recommendations for palliative care, where relative involvement is recommended as the relatives suffer equally with the patient and after the patient has died [5].

Strength and Limitations

The study was conducted as an explorative study, the first study in a Danish context conducted with patients' and their relatives' experiences during 82 video consultations in SPC at home. The SPC team nurse carried out the patient recruitment to avoid resistance to changes. When she was unavailable, patients were not included by other SPC team nurses. Patient inclusion was ongoing during the study period because of the focus on the increasing vulnerability of patients and their relatives. Some video consultations were rescheduled owing to hospice admission for optimizing stay.

The low participation from relatives (n=3) may be due to not all relatives being offered participation at the initial home visit.

Of the 18 excluded patients (Table 2), 7 were excluded owing to terminal stage of illness, cognitive impairment, or hospital admission caused by illness progression when first referred to the SPC. This could be prevented, according to Dalgaard, by early integration of palliative care, which may result in better symptom management, prolonged survival, improved patient perception of prognosis, and less aggressive care at the end of life [36].

The combination of participant observations and interviews helped vary the study and express discrepancies during fieldwork [37]. The phenomenological approach was helpful in maintaining openness as well as setting preconceptions aside in the process when exploring the participants' experiences in using video consultations in SPC at home [29].

The findings in this study will be accompanied by a study focusing on health care professionals, the feasibility of video consultations in their work, and the organizational changes that may occur owing to implementing video consultations in the SPC team care.

Clinical Implications and Future Research

As video consultations seem feasible and the study was initiated from 1 SPC team, it could be relevant to explore the feasibility in more SPC teams in an international context [38]. The analysis of qualitative data could be hypothesis-generating, with the aim of conducting a quantitative study. In this study, the qualitative data could generate the initial stages in identifying issues for respondents in questionnaire development [39]. Caring for the relatives in SPC is recommended [5], and by involving a family focusing approach in the video consultations, the patients and relatives can contribute with knowledge and experiences from their everyday lives [40].

Conclusions

The use of video consultations strengthens communication. Despite their serious illness, patients took an active role, and despite technical issues, the tablet allowed access for the relatives, both present and via app, to attend in equal terms as the patient in the video consultation. The video consultations require introduction and reliable internet access before the video consultation can take place. Furthermore, agreement must be sought in SPC at home before video consultations are initiated and for as long as they continue.

Acknowledgments

The authors would like to thank all participants. The project was funded by the Danish Knowledge Centre for Rehabilitation and Palliative Care REHPA, EWII Telecare, Patient@home, Co-Lab, CIMT Center for Innovative Medical Technology, the University of Southern Denmark, and Department of Oncology, Odense University Hospital.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Observation guide and semistructured interview guide in an explorative study about video consultations in specialized palliative home care in Denmark.

[PDF File (Adobe PDF File), 35KB - jmir_v21i3e10208_app1.pdf]

References

1. World Health Organization. 2018 Feb. Palliative Care URL: https://www.who.int/news-room/fact-sheets/detail/palliative-care [accessed 2019-01-20] [WebCite Cache ID 75ZQZbygg]

XSL•FO
RenderX

2. Danish Health Data Authority. 2016. Course of Death Register-Numbers and Analysis-Danish Health Data Authority (DHDA) Internet URL: https://sundhedsdatastyrelsen.dk/da/tal-og-analyser/analyser-og-rapporter/andre-analyser-og-rapporter/doedsaarsagsregisteret [accessed 2019-01-20] [WebCite Cache ID 75ZQr1yLP]
3. Danish Multidisciplinary Cancer Group for palliative effort. [Danish Multidisciplinary Cancer Group for palliative efforts]. [Annual Report from the Danish Palliative Care Database (DPCD) 2015] URL: http://www.dmcgpal.dk/573/%C3%85rsrapport-fra-dansk-palliativ-database [accessed 2019-01-20] [WebCite Cache ID 75ZR0yvk6]
4. Danish Health Authority. [Danish Health Authority]. [Palliative action] URL: https://www.sst.dk/da/sygdom-og-behandling/palliativ-indsats [accessed 2019-01-20] [WebCite Cache ID 75ZR4Tdyf]
5. Danish Health Authority. [Danish Health Authority]. Copenhagen; 2011. [Recommendations for Palliative Care] URL: http://www.sst.dk/~/media/3B57BB8B65014D73B47A7023546B4A62.ashx [accessed 2019-01-23] [WebCite Cache ID 75dOmm3WR]
6. Collier A, Morgan DD, Swetenham K, To TH, Currow DC, Tieman JJ. Implementation of a pilot telehealth programme in community palliative care: a qualitative study of clinicians' perspectives. Palliat Med 2016 Apr;30(4):409-417. [doi: 10.1177/0269216315600113] [Medline: 26290500]
7. Neergaard MA, Warfvinge JE, Jespersen TW, Olesen F, Ejskjaer N, Jensen AB. The impact of "modern telecommunication" in palliative care--the views of the professionals. Telemed J E Health 2014 Jan;20(1):24-31. [doi: 10.1089/tmj.2013.0047] [Medline: 24160901]
8. van Gurp J, van Selm M, van Leeuwen E, Vissers K, Hasselaar J. Teleconsultation for integrated palliative care at home: a qualitative study. Palliative Medicine 2016 Mar;30(3):257-269. [doi: 10.1177/0269216315598068] [Medline: 26269323]
9. Stern A, Valaitis R, Weir R, Jadad AR. Use of home telehealth in palliative cancer care: a case study. J Telemed Telecare 2012 Jul;18(5):297-300. [doi: 10.1258/jtt.2012.111201] [Medline: 22790013]
10. Johnston B, Kidd L, Wengstrom Y, Kearney N. An evaluation of the use of telehealth within palliative care settings across Scotland. Palliat Med 2012 Mar;26(2):152-161. [doi: 10.1177/0269216311398698] [Medline: 21378067]
11. Hoek PD, Schers HJ, Bronkhorst EM, Vissers KC, Hasselaar JG. The effect of weekly specialist palliative care teleconsultations in patients with advanced cancer-a randomized clinical trial. BMC Med 2017 Dec 19;15(1):119 [FREE Full text] [doi: 10.1186/s12916-017-0866-9] [Medline: 28625164]
12. Capurro D, Ganzinger M, Perez-Lu J, Knaup P. Effectiveness of eHealth interventions and information needs in palliative care: a systematic literature review. J Med Internet Res 2014 Mar 7;16(3):e72 [FREE Full text] [doi: 10.2196/jmir.2812] [Medline: 24610324]
13. Maudlin J, Keene J, Kobb R. A road map for the last journey: home telehealth for holistic end-of-life care. Am J Hosp Palliat Care 2006;23(5):399-403. [doi: 10.1177/1049909106290807] [Medline: 17060308]
14. Kidd L, Cayless S, Johnston B, Wengstrom Y. Telehealth in palliative care in the UK: a review of the evidence. J Telemed Telecare 2010;16(7):394-402. [doi: 10.1258/jtt.2010.091108] [Medline: 20813893]
15. Wittenberg-Lyles E, Oliver DP, Demiris G, Washington KT, Regehr K, Wilder HM. Question asking by family caregivers in hospice interdisciplinary team meetings. Res Gerontol Nurs 2010 Apr;3(2):82-88 [FREE Full text] [doi: 10.3928/19404921-20090731-05] [Medline: 20415357]
16. Taylor A, Morris G, Pech J, Rechter S, Carati C, Kidd MR. Home telehealth video conferencing: perceptions and performance. JMIR Mhealth Uhealth 2015 Sep 17;3(3):82-88 [FREE Full text] [doi: 10.2196/mhealth.4666] [Medline: 26381104]
17. Hennemann-Krause L, Lopes AJ, Araújo JA, Petersen EM, Nunes RA. The assessment of telemedicine to support outpatient palliative care in advanced cancer. Palliat Support Care 2015 Aug;13(4):1025-1030. [doi: 10.1017/S147895151400100X] [Medline: 25159308]
18. van Gurp J, van Selm M, Vissers K, van Leeuwen E, Hasselaar J. How outpatient palliative care teleconsultation facilitates empathic patient-professional relationships: a qualitative study. PLoS One 2015;10(4):e0124387 [FREE Full text] [doi: 10.1371/journal.pone.0124387] [Medline: 25902263]
19. Green J, Thorogood N. Qualitative methodology and health research. In: Qualitative Methods for Health Research Third Edition. Los Angeles: Sage Publications; 2014:3-34.
20. World Medical Association. World Medical Association. 2018 Jul 9. WMA Declaration of Helsinki - Ethical Principles For Medical Research Involving Human Subjects URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2019-01-20] [WebCite Cache ID 75ZT8dc5e]
21. Spradley JP. Participant Observation. New York: Holt, Rinehart and Winston; 1980.
22. Holstein JA, Gubrium JF. The Active Interview. Thousand Oaks, California: Sage Publications; 1995:38-51.
23. Green J, Thorogood N. In-depth Interviews. In: Qualitative Methods for Health Research Third Edition. Los Angeles: Sage Publications; 2014:95-125.
24. Azzarito L. Photography as a pedagogical tool for shedding light on 'bodies-at-risk' in physical culture. Vis Stud 2012 Oct 26;27(3):295-309. [doi: 10.1080/1472586X.2012.717746]
25. Green J, Thorogood N. Mixing methods and designs. In: Qualitative Methods for Health Research Third Edition. Los Angeles: Sage Publications; 2014:279-301.

XSL•FO
RenderX

26. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
27. Giorgi A. An application of phenomenological method in psychology. In: Giorgi A, Fischer WF, von Eckartsberg R, editors. Duquesne Studies in Phenomenological Psychology. Pittsburgh: Duquesne University Press; 1975:82-103.
28. Brinkmann S, Tanggaard L. [Phenomenology]. In: Qualitative methods: A Basic Book. Copenhagen: Hans Reitzel; 2015:217-239.
29. Giorgi A. The Descriptive Phenomenological Method in Psychology: A Modified Huessrlian Approach. Pittsburgh: Duquesne University Press; 2009.
30. Ihde D. Technology and the Lifeworld: From Garden to Earth. Bloomington: Indiana University Press; 1990.
31. Green J, Thorogood N. Observational methods. In: Qualitative Methods for Health Research Third Edition. Los Angeles: Sage Publications; 2014:151-178.
32. Repstad P. [Between Proximity and Distance: Validative Methods in Social Sciences]. Oslo: Universitetsforlaget; 2009.
33. Huniche L, Olesen F. [Technology understanding and health practices]. In: [Technology in Health Practice]. Copenhagen: Munksgaard; 2014:35-59.
34. World Health Organization. Telehealth URL: https://www.who.int/sustainable-development/health-sector/strategies/telehealth/en/ [accessed 2019-01-20] [WebCite Cache ID 75ZWCtEga]
35. Radbruch L, Payne SA. White paper on standards and norms for hospice and palliative care in Europe: part 1. Eur J Palliat Care 2009;16(6):278-289 [FREE Full text]
36. Dalgaard KM, Bergenholtz H, Nielsen ME, Timm H. Early integration of palliative care in hospitals: a systematic review on methods, barriers, and outcome. Palliat Support Care 2014 Dec;12(6):495-513. [doi: 10.1017/S1478951513001338] [Medline: 24621947]
37. Malterud K. [Qualitative Methods in Medical Research an Introduction]. Oslo: Universitetsforlaget; 2011.
38. REHPA. [REHPA-Knowledge Center for Rehabilitation and Palliation]. URL: http://www.rehpa.dk/professionelle/palliation/fakta/ [accessed 2019-01-20] [WebCite Cache ID 75ZWWVueN]
39. Green J, Thorogood N. Developing qualitative research designs. In: Qualitative Research Methods for Health Research Third Edition. Los Angeles: Sage Publications; 2014:35-63.
40. Wright LM, Leahey M. Nurses and Families: A Guide to Family Assessment and Intervention Sixth Edition. Philadelphia: FA Davis Company; 2013.

Abbreviations

SPC: specialized palliative care
WHO: World Health Organization

Edited by G Eysenbach; submitted 24.02.18; peer-reviewed by V Klaas, H Yu; comments to author 03.08.18; revised version received 17.09.18; accepted 03.10.18; published 21.03.19.

Please cite as:
Funderskov KF, Raunkiær M, Danbjørg DB, Zwisler AD, Munk L, Jess M, Dieperink KB
Experiences With Video Consultations in Specialized Palliative Home-Care: Qualitative Study of Patient and Relative Perspectives
J Med Internet Res 2019;21(3):e10208
URL: https://www.jmir.org/2019/3/e10208/
doi:10.2196/10208
PMID:30896436

©Karen Frydenrejn Funderskov, Mette Raunkiær, Dorthe Boe Danbjørg, Ann-Dorthe Zwisler, Lene Munk, Mia Jess, Karin Brochstedt Dieperink. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Use of Online or Paper Surveys by Australian Women: Longitudinal Study of Users, Devices, and Cohort Retention

David Fitzgerald[1], MSc; Richard Hockey[1], BSc; Mark Jones[1], PhD; Gita Mishra[1], PhD; Michael Waller[1], PhD; Annette Dobson[1], PhD

Centre for Longitudinal and Life Course Research, School of Public Health, University of Queensland, Brisbane, Australia

Corresponding Author:
Annette Dobson, PhD
Centre for Longitudinal and Life Course Research
School of Public Health
University of Queensland
Herston
Brisbane,
Queensland 4006
Australia
Phone: 61 7 3365 5346
Fax: 61 7 3365 5442
Email: a.dobson@sph.uq.edu.au

Abstract

Background: There is increasing use of online surveys to improve data quality and timeliness and reduce costs. While there have been numerous cross-sectional studies comparing responses to online or paper surveys, there is little research from a longitudinal perspective.

Objective: In the context of the well-established Australian Longitudinal Study on Women's Health, we examined the patterns of responses to online or paper surveys across the first two waves of the study in which both modes were offered. We compared the following: differences between women born between 1946 and 1951 and between 1973 and 1978; types of device used for online completion; sociodemographic, behavioral, and health characteristics of women who responded online or using mailed paper surveys; and associations between mode of completion in the first survey and participation and mode of completion in the second survey.

Methods: Participants in this study, who had responded to regular mailed surveys since 1996, were offered a choice of completing surveys using paper questionnaires or Web-based electronic questionnaires starting in 2012. Two groups of women were involved: an older cohort born between 1946 and 1951 aged in their 60s and a younger cohort born between 1973 and 1978 aged in their 30s when the online surveys were first introduced. We compared women who responded online on both occasions, women who responded online at the first survey and used the paper version of the second survey, women who changed from paper to online, and those who used paper for both surveys.

Results: Of the 9663 women in their 60s who responded to one or both surveys, more than 50% preferred paper surveys (5290/9663, 54.74%, on the first survey and 5373/8621, 62.32%, on the second survey). If they chose the online version, most used computers. In contrast, of the 8628 women in their 30s, 56.04% (4835/8628) chose the online version at the first survey. While most favored computers to phones or tablets, many did try these alternatives on the subsequent survey. Many women who completed the survey online the first time preferred the paper version on the subsequent survey. In fact, for women in their 60s, the number who went from online to paper (1151/3851, 29.89%) exceeded the number who went from paper to online (734/5290, 13.88%). The online option was more likely to be chosen by better educated and healthier women. In both cohorts, women who completed paper surveys were more likely than online completers to become nonrespondents on the next survey. Due to the large sample size, almost all differences were statistically significant, with $P<.001$.

Conclusions: Despite the cost-saving advantages of online compared to paper surveys, paper surveys are likely to appeal to a different population of potential respondents with different sociodemographic, behavioral, and health characteristics and greater likelihood of attrition from the study. Not offering a paper version is therefore likely to induce bias in the distribution of responses unless weighting for respondent characteristics (relative to the target population) is employed. Therefore, if mixed mode (paper

XSL•FO
RenderX

or online) options are feasible, they are highly likely to produce more representative results than if only the less costly online option is offered.

(J Med Internet Res 2019;21(3):e10672) doi:10.2196/10672

KEYWORDS

online survey; paper survey; longitudinal study; participant retention

Introduction

There is increasing use of online surveys for research to increase participation; improve data completeness, quality, and availability; and reduce costs [1-3]. There have been randomized trials of paper-based surveys compared with online surveys [4-8], nonrandomized studies [9-11], comparisons of different devices for online completion [12], and mixed-methods approaches in which participants are offered both modes simultaneously or sequentially [1,5,9-12].

Researchers usually find that younger and more educated participants favor internet-based modes of completion [13-15], although many participants prefer postal surveys completed with paper and pen [7,14]. Most of the studies have been cross-sectional, and much less has been reported about associations between mode of survey delivery and device use, participant characteristics, and cohort retention in the context of longitudinal studies.

While a mixed mode of data collection has advantages, allowing participants to switch the mode of questionnaire completion has the potential to change the overall and partial response rates, alter representativeness of the sample being assessed, and change the way in which participants respond to particular questions [12,16]. Data from the United Kingdom Household Longitudinal Study (UKHLS) have been used to assess the effect of switching from a face-to-face mode of collection to include telephone or Web-based methods. In some cases, the mixed modes of data collection had a detrimental effect on response rates [16,17]. It has also been hypothesized that the mode of data collection may contribute to differences in the ability of participants to recall events accurately, and the different reliabilities that result from a mixed mode approach may hamper the longitudinal comparisons within a study [18]. To date, results from the UKHLS have shown broad agreement in the accuracy [19] and reliability [18] of responses collected using different modes; however, less research has focused directly on the effects of changing from paper-based surveys to a mixed mode choice of paper and Web-based surveys in a longitudinal study setting [12,20].

By 2012-2013, most households in Australia had internet access, ranging from 85% in the capital cities to 79% in rural areas [21]. Therefore, in order to reduce costs and improve data quality, participants in the Australian Longitudinal Study on Women's Health, who had responded to regular surveys since 1996, were offered a choice of completing surveys using mailed questionnaires or Web-based electronic questionnaires starting in 2012. Two groups of women were involved: an older cohort born between 1946 and 1951 and a younger cohort born between 1973 and 1978. For this report, our objectives were to examine longitudinal patterns of paper and online responses over the first two waves in which both options were offered:

- Between women born between 1946 and 1951 and between 1973 and 1978
- By type of device used for online completion
- By the sociodemographic, behavioral, and health characteristics of women who responded online or using mailed paper surveys
- Between mode of completion of the first survey and participation and mode of completion in the second survey
- By survey performance criteria, including response rates, noncompletion (break off), missing item responses, and measurement equivalence.

Methods

Australian Longitudinal Study on Women's Health

In 1996, women in three age cohorts—born between 1921 and 1926, born between 1946 and 1951, and born between 1973 and 1978—were recruited to the Australian Longitudinal Study on Women's Health, with more than 12,000 participants in each cohort responding to the baseline surveys. The sampling frame was the database of the universal national health insurance scheme, now called Medicare, which covers all citizens and permanent residents. Women living in rural and remote areas were recruited at twice the population proportion to ensure sufficient power to detect important differences in health and use of services. Details of the study recruitment, national representativeness, attrition, and measures have been published elsewhere [22,23]. This paper uses data from the women born between 1946 and 1951 and between 1973 and 1978.

Survey Design

Since the baseline survey in 1996, the women have been resurveyed approximately every three years. Initially paper questionnaires with reply-paid envelopes were mailed to participants. The surveys have many common items but are tailored to address health and health service issues relevant to each cohort and time.

From survey 6 in 2012, women in the 1973-1978 cohort were offered the option of completing the questionnaires online using versions designed for use on mobile phones, tablets, and computers. From survey 7 in 2013, those in the 1946-1951 cohort had the same options. Since then both cohorts have completed another survey enabling us to look at changes in mode of completion and associations between mode of completion and response to the subsequent survey.

At each survey, women who have not previously withdrawn from the study or died and for whom we have current contact details are invited to participate. Since 2012, women for whom

XSL•FO
RenderX

we have email addresses have been emailed an invitation and a link to the online survey. If they have not responded after 4 to 6 weeks, during which they are sent email, postal, and text-message reminders, they are mailed paper surveys (followed by email, text-message, and telephone reminders to nonresponders). Women for whom we do not have current email addresses are initially mailed an invitation and the link to the online survey (followed by a mailed or text-message reminder). If they have not completed the online survey within 4 to 6 weeks, they are mailed paper surveys (and followed up using the same reminder protocol as the other group).

When the online surveys were introduced, the paper versions were reformatted to improve questionnaire equivalence. For example, long lists of items that had fit on a page but might not fit on a single screen were broken into shorter lists, each with the same instructions. In addition, several steps were taken to adapt the online surveys to mobile devices. Maximum pixel width for a survey question was limited. Lists of question items with numerous response options that took up a large width were transformed into single items with reduced width that used more scrolling on a mobile device screen. The size of question response controls (radio buttons, checkboxes, buttons, etc) were increased for more accurate touch screen interaction.

Measurements

The device used by the participant to complete the survey online was determined from the text of the browser's user agent string. The user agent records which browser and operating system are being used. This information was provided by the online platform used for the surveys. If the user agent string contained the term iPhone, iPad, or Android, the device was categorized as tablet/phone. If the terms Windows, Macintosh, or Linux appeared, the device was categorized as a computer. The device type was considered missing if the user agent string did not contain any of the identified words.

Area of residence was based on a measure of distance from where the woman lived to population centers with varying levels of services (for health, education, shops, etc) [24]. For this paper we used the categories major city, inner regional area, outer regional area, and remote or very remote area. The highest level of education the woman had completed was categorized as school certificate or less, higher school certificate, trade certificate or diploma, or university degree. Marital status was categorized as partnered (married/de facto relationship) or not partnered (separated, divorced, widowed, or never married). A woman's work pattern was categorized according to her responses to questions about time use. Full-time work was defined as working a total of 35 or more hours per week, and part time work was between 1 and 35 hours; women were classified as not working if they used the response option "I don't do this activity" to questions about time spent working.

Based on questions about cigarette use, women were categorized as never smoker, former smoker, or current smoker. Alcohol consumption was calculated from questions about frequency and quantity of alcohol consumed and classified according to national guidelines as nondrinker, low-risk drinker or rarely drinks, or risky drinker [25]. At each survey, women reported their height and weight, and these data were used to calculate body mass index (BMI) (kg/m^2). Each survey included the 36-Item Short Form Health Survey (SF-36) [26]. The item "In general, would you say your health is" with response options excellent, very good, good, fair, or poor was used as a categorical measure of general health. In addition, the physical function and mental health subscales from the SF-36 were calculated.

Statistical Analysis

To compare modes of completion using the two surveys, matched cross-tabulations and tests for marginal homogeneity (Stuart-Maxwell test) and symmetry were used. To examine the characteristics of women who maintained the same mode of completion versus those who changed mode, four groups were defined: online both times, online followed by paper, paper followed by online, and paper both times. To assess the effect of changing mode of completion on trend estimates, we calculated changes in responses by the same individuals between the two successive surveys before the introduction of the online option and then between the survey immediately before and the survey at which the online option was introduced. Differences between the four groups of women were compared using chi-square tests for categorical variables and one-way analysis of variance (ANOVA) and Tukey tests for continuous variables to identify which groups were statistically significantly different.

Results

A total of 9663 women in the 1946-1951 cohort responded to surveys 7 or 8 or both. We were unable to identify the device used by 26 of those who completed the survey online at survey 7; for the remaining 9637, the modes of completion and nonresponse are shown in Table 1. Most women used paper surveys on both occasions, with computers the second most frequent method. The use of tablets or phones increased from survey 7 to survey 8, but this remained the less favored mode of completion. While 13.88% (734/5290) of women who used paper at survey 7 changed to an online mode (computer, tablet, or phone) at survey 8, 29.89% (1151/3851) of those who completed survey 7 online chose to use the paper version at survey 8. All these differences were highly statistically significant ($P<.001$). The overall response rates (for all modes of completion) for surveys 6, 7, and 8 were 82.99% (10,011/12,063), 81.05% (8622/10,719), and 80.44% (8622/10,719), respectively. While break off percentages at survey 7 were low (14/3575, 0.39%) for online participants who also responded to survey 8, the percentage was higher (14/286, 4.90%) for women who did not respond at survey 8. Similarly, the percentage of women with missing responses for more than 10% of items at survey 7 was low (116/8109, 1.43%) among respondents in all four mode of completion groups who also responded to survey 8, but it was higher for those who did not respond to survey 8 (6.6% [19/286] for those who responded online to survey 7 and 6.5% [49/756] for those who responded on paper).

XSL•FO
RenderX

Table 1. Comparison of mode of completion of survey 7 in 2013 and survey 8 in 2016 women in the 1946-1951 cohort of the Australian Longitudinal Study on Women's Health.

	Survey 8				
Survey 7	Computer	Tablet/phone	Paper	Nonresponse	Total
Computer	1772 (51.80)	369 (10.79)	1022 (29.87)	258 (7.54)	3421 (100)
Tablet/phone	79 (19.08)	184 (44.44)	123 (29.71)	28 (6.76)	414 (100)
Paper	533 (10.08)	201 (3.80)	3800 (71.83)	756 (14.29)	5290 (100)
Nonresponse	60 (12.72)	24 (4.69)	428 (83.59)	—	512 (100)
Total	2444 (25.36)	778 (8.07)	5373 (55.75)	1042 (10.81)	9637 (100)

Characteristics of women in the 1946-1951 cohort in the four main groups are compared in Table 2. While women who responded online in one or both surveys were very similar, women who used the paper survey both times were more likely to live in regional (but not remote) areas, less educated, less likely to be working, more likely to be smokers, and less likely to drink alcohol. These women were also more likely to be overweight or obese and had poorer health on all measures. Like the women who changed from paper to online completion, the paper-only group were less likely to have partners than the groups of women who completed survey 7 online. All differences were statistically significant at $P<.001$.

Table 3 shows that these differences between the groups were (retrospectively) evident even at the baseline survey in 1996. The Tukey tests showed that the differences were consistently greater between the women in the paper-only group and the other three groups. They were slightly older (even though all women were within a 5-year age range), had higher BMIs, and had poorer levels of physical function and mental health. In contrast, the changes over time were of similar magnitude in all four groups. That is, although the levels of these variables were different at baseline, they tracked approximately in parallel over time. While the changes between surveys 5 and 6 appear somewhat different from the changes between surveys 6 and 7, possibly due to the mode change, the effects were similar across all four groups of women (except for a greater decline in physical function in the paper-only group).

Among the women born between 1973 and 1978, 8628 completed one or both of surveys 6 and 7. For 10 of the online responders we could not identify the device they used, but for the remainder the modes of completion and nonresponse are shown in Table 4. As expected, this younger group of women were more likely than the 1946-1951 cohort to complete the surveys online and to use tablets or phones. Also, unlike the older group, more changed from paper to online surveys (1006/3164, 31.80%) than went from online to paper (963/4835, 19.92%, of those who completed survey 6 online).

XSL•FO
RenderX

Table 2. Comparison of women in the 1946-1951 cohort of the Australian Longitudinal Study on Women's Health who responded to survey 7 in 2013 and survey 8 in 2016 (N=8109), according to their modes of survey completion.

Characteristics	Online/online (n=2424[a,b]), n (%)	Online/paper (n=1151[c,b]), n (%)	Paper/online (n=734[b]), n (%)	Paper/paper (n=3800[b]), n (%)
Area of residence				
Major city	1055 (44.03)	456 (39.93)	317 (43.37)	1314 (34.84)
Inner regional	898 (37.48)	461 (40.37)	283 (38.71)	1572 (41.69)
Outer regional	396 (16.53)	195 (17.08)	118 (16.14)	794 (21.06)
Remote/very remote	47 (1.96)	30 (2.63)	13 (1.78)	91 (2.41)
Educational qualifications				
School certificate or less	572 (25.05)	336 (31.26)	193 (28.34)	1620 (47.07)
Higher school certificate	426 (18.66)	208 (19.35)	121 (17.77)	695 (20.19)
Trade certificate/diploma	545 (23.87)	265 (24.65)	182 (26.73)	664 (19.29)
University degree	740 (32.41)	266 (24.74)	185 (27.17)	463 (13.45)
Marital status				
Partnered	1825 (75.63)	925 (81.07)	534 (73.25)	2774 (73.48)
Not partnered	588 (24.37)	216 (18.93)	195 (26.75)	1001 (26.52)
Work				
Full-time	370 (15.33)	200 (17.53)	143 (19.51)	504 (13.37)
Part-time	791 (32.77)	354 (31.03)	248 (33.83)	1045 (27.72)
Not working	1253 (51.91)	587 (51.45)	342 (46.66)	2221 (58.91)
Smoking				
Never smoked	1496 (61.95)	728 (63.80)	494 (67.49)	2401 (63.60)
Former smoker	824 (34.12)	372 (32.60)	213 (29.10)	1053 (27.89)
Current smoker	95 (3.93)	41 (3.59)	25 (3.42)	321 (8.50)
Alcohol consumption				
Nondrinker	244 (10.10)	123 (10.76)	90 (12.31)	797 (21.32)
Low-risk/rarely drinks	1997 (82.66)	962 (84.16)	600 (82.08)	2732 (73.09)
Risky drinker	175 (7.24)	58 (5.07)	41 (5.61)	209 (5.59)
General health				
Excellent/very good	1441 (59.50)	619 (53.87)	363 (49.52)	1512 (39.89)
Good	782 (32.29)	415 (36.12)	297 (40.52)	1659 (43.77)
Fair/poor	199 (8.22)	115 (10.01)	73 (9.96)	619 (16.33)

[a]There were 20 women who responded online to both surveys who were not included in Table 1 because it was not possible to know if they had used a computer, tablet, or phone in survey 7.

[b]Not all women responded to every item.

[c]There were 6 women who responded online for survey 7 and used paper for survey 8 who were not included in Table 1 because it was not possible to know if they had used a computer, tablet, or phone on survey 7.

XSL•FO
RenderX

Table 3. Comparison of measures at surveys 1 (1996), 6 (2010), 7 (2013), and 8 (2016) from women in the 1946-1951 cohort of the Australian Longitudinal Study on Women's Health who responded to surveys 7 and 8, according to their mode of survey completion.

Characteristics	Online/online (n=2424)	Online/paper (n=1151)	Paper/online (n=73)	Paper/paper (n=3800)	*P* value
Age at survey 1, mean	47.44	47.45	47.45	47.68	<.001
Body mass index					
Survey 1	25.40	25.45	25.51	25.92	<.001
Change from survey 6 to survey 7	0.09	0.23	0.23	0.17	.20
Change from survey 7 to survey 8	–0.01	0.01	–0.01	0.09	.23
SF-36[a] subscale for physical function					
Survey 1	88.94	88.00	88.91	86.40	<.001
Change from survey 6 to survey 7	–1.75	–2.07	–1.07	–1.65	.61
Change from survey 7 to survey 8	–1.07	–1.97	–1.40	–2.34	.02
SF-36 subscale for mental health					
Survey 1	75.89	75.09	75.52	74.24	<.001
Change from survey 6 to survey 7	0.72	1.00	0.99	0.56	.77
Change from survey 7 to survey 8	–0.39	–0.68	–0.18	0.07	.35

[a]SF-36: 36-Item Short Form Health Survey.

Table 4. Comparison of modes of completion of survey 6 in 2012 and survey 7 in 2015 women in the 1973-1978 cohort of the Australian Longitudinal Study on Women's Health.

	Survey 7				
Survey 6	Computer	Tablet/Phone	Paper	Nonresponse	Total
Computer	1981 (46.52)	833 (19.56)	846 (19.87)	598 (14.04)	4258 (100)
Tablet/Phone	137 (23.74)	229 (39.69)	117 (20.28)	94 (16.29)	577 (100)
Paper	595 (18.81)	411 (12.99)	1408 (44.50)	750 (23.70)	3164 (100)
Nonresponse	203 (32.79)	126 (20.36)	290 (46.85)	—	619 (100)
Total	2916 (33.84)	1599 (18.55)	2661 (30.88)	1442 (16.73)	8618 (100)

The overall response rates (for all modes of completion) for surveys 5, 6, and 7 were 62.01% (8200/13,223), 61.63% (8010/12,997), and 56.61% (7186/12,693), respectively, of all eligible women. While break off percentages at survey 6 were low (36/4152, 0.87%) for online participants who also responded to survey 7, the percentage was higher (43/692, 6.21%) for women who did not respond at survey 7. Similarly, the percentage of women with missing responses for more than 10% of items at survey 7 was low (91/6567, 1.39%) among respondents in all four mode of completion groups who also responded to survey 8, but it was higher for those who did not respond to survey 7 (7.4% [51/692] for those who responded online at survey 6 and 3.2% [24/750] for those who responded on paper).

The women born between 1973 and 1978, who were aged 34 to 39 years when they responded to survey 6 in 2012, were more urban, much better educated, more likely to work full-time or part-time, and generally in better health (see Table 5) than the women in the 1946-1951 cohort, who were aged 62 to 67 years when they responded to their survey 7 in 2013. The younger women were more homogeneous than the older women, but differences between the four response groups followed broadly the same patterns. Most of the differences were still highly statistically significant (*P*<.001), although the differences were less pronounced for some characteristics (eg, for marital status, *P*=.006) and the differences between the groups on general health were not statistically significant (*P*=.15).

As for the older cohort, these differences between the groups were (retrospectively) evident even at the baseline survey in 1996 (see Table 6). The differences were mainly that the women in the online-only group had better physical function and mental health than women in the other three groups. While the changes over time differed between surveys 5 and 6 and surveys 6 and 7 (suggestive of a possible mode of completion effect), they were of similar magnitude in all four groups of women.

XSL•FO
RenderX

Table 5. Comparison of women in the 1973-1978 cohort of the Australian Longitudinal Study on Women's Health who responded to survey 6 in 2012 and survey 7 in 2015 (N=6567), according to their modes of survey completion.

Characteristics	Online/online (n=3189[a,b]), n (%)	Online/paper (n=963[b]), n (%)	Paper/online (n=1007[c,b]), n (%)	Paper/paper (n=1408[b]), n (%)
Area of residence				
Major city	1917 (64.68)	539 (57.16)	543 (54.96)	749 (53.92)
Inner regional	701 (23.65)	261 (27.68)	272 (27.53)	397 (28.58)
Outer regional	289 (9.75)	132 (14.00)	147 (14.88)	217 (15.62)
Remote/very remote	57 (1.92)	11 (1.17)	26 (2.63)	26 (1.87)
Educational qualifications				
School certificate or less	108 (3.41)	44 (4.64)	50 (5.04)	127 (9.20)
Higher school certificate	297 (9.39)	110 (11.59)	118 (11.88)	200 (14.49)
Trade certificate/diploma	786 (24.84)	257 (27.08)	263 (26.49)	418 (30.29)
University degree	1973 (62.36)	538 (56.69)	562 (56.60)	635 (46.01)
Marital status				
Partnered	2493 (78.79)	785 (82.72)	808 (80.24)	1078 (76.95)
Not partnered	671 (21.21)	164 (17.28)	199 (19.76)	323 (23.05)
Work				
Full-time	1429 (45.15)	368 (38.78)	389 (38.71)	493 (35.09)
Part-time	1189 (37.57)	414 (43.62)	432 (42.99)	626 (44.56)
Not working	547 (17.28)	167 (17.60)	184 (18.31)	286 (20.36)
Smoking				
Never smoked	2008 (63.32)	601 (62.93)	635 (63.18)	855 (60.85)
Former smoker	875 (27.59)	272 (28.48)	265 (26.37)	353 (25.12)
Current smoker	288 (9.08)	82 (8.59)	105 (10.45)	197 (14.02)
Alcohol consumption				
Nondrinker	330 (10.41)	102 (10.68)	115 (11.44)	209 (14.98)
Low-risk/rarely drinks	2697 (85.08)	804 (84.19)	859 (85.47)	1127 (80.79)
Risky drinker	143 (4.51)	49 (5.13)	31 (3.08)	59 (4.23)
General health				
Excellent/very good	1979 (62.21)	600 (62.50)	593 (58.95)	818 (58.18)
Good	943 (29.64)	284 (29.58)	323 (32.11)	465 (33.07)
Fair/poor	259 (8.14)	76 (7.92)	90 (8.95)	123 (8.75)

[a]There were 9 women who responded online to both surveys who were not included in Table 5 because it was not possible to know if they had used a computer, tablet, or phone on survey 7.

[b]Not all women responded to survey.

[c]There was 1 woman who used paper on survey 6 and responded online on survey 7 who was not included in Table 5 because it was not possible to know if she had used a computer, tablet, or phone in survey 7.

XSL•FO
RenderX

Table 6. Comparison of measures at surveys 1 (1996), 5 (2009), 6 (2012) and 7 (2015) from women in the 1973-1978 cohort of the Australian Longitudinal Study on Women's Health who responded to surveys 6 and 7, according to their mode of survey completion.

Characteristics	Online/online (n=3189)	Online/paper (n=963)	Paper/online (n=1007)	Paper/paper (n=1408)	*P* value
Age at survey 1, mean	20.75	20.79	20.81	20.87	.08
Body mass index					
Survey 1	22.81	22.24	22.73	22.77	.003
Change from survey 5 to survey 6	0.36	0.44	0.42	0.50	.37
Change from survey 6 to survey 7	0.60	0.60	0.61	0.55	.92
SF-36[a] subscale for physical function					
Survey 1	92.61	92.45	91.67	91.40	<.001
Change from survey 5 to survey 6	0.14	0.25	–0.22	0.57	.73
Change from survey 6 to survey 7	–0.45	–0.14	0.58	–0.28	.35
SF-36 subscale for mental health					
Survey 1	70.65	69.93	68.13	68.22	.008
Change from survey 5 to survey 6	0.03	0.14	0.25	0.53	.82
Change from survey 6 to survey 7	–1.07	–1.06	–1.70	–1.68	.50

[a]SF-36: 36-Item Short Form Health Survey.

Discussion

Principal Findings

Introducing the choice of online or paper completion of successive surveys in a well-established longitudinal study enabled us to examine how participants would respond over time. As expected, we found that the online option was more likely to be chosen by younger, better educated women. Surprisingly, we found that a large number of women who completed the survey online the first time this was offered preferred the paper version on the subsequent survey. In fact, for women in their 60s, the number who went from online to paper exceeded the number who went from paper to online.

Overall the women in their 60s tended to prefer to continue with paper surveys, and if they chose the online version most used computers. In contrast, the younger women, in their 30s, were more likely to adopt the online version and, while most favored computers to phones or tablets, many did try these alternatives at the subsequent survey.

In this study, it is not uncommon for participants to miss one (or more) of the triennial surveys but complete subsequent surveys. We found that women who completed paper surveys were more likely than online completers to become nonrespondents at the next survey.

Although there was some evidence of an effect of introducing the online option on the temporal changes in key variables, the effects were similar for women who made different choices about mode of completion. That is, there was little evidence of confounding between self-selection of mode of completion and measurement in these health-related variables. Changing the format of the paper surveys to be more like the format of the online surveys may have achieved measurement equivalence in the same survey wave but introduced discontinuity over time.

Limitations

In this paper, we report survey response behavior of women in their 30s and 60s in Australia. The results may not be generalizable to men, people in other age groups, or people with less Web access. Nevertheless, this is a large study with more than 18,000 participants who were originally randomly selected from almost all women in Australia in their age groups at baseline in 1996. Over time, national migration and study attrition have changed the representativeness of the study population, favoring more educated and Australian-born women, but the resultant bias is unlikely to affect the direction of the main results [27].

Although internet access in Australia was more than 80% when the online surveys were first introduced, it may not have been sufficiently reliable for completion of a long survey. At the end of each survey (paper or online), women are encouraged to provide additional comments and suggestions in free text format. While many women do provide comments on a variety of topics, fewer than five mentioned the mode of completion, mainly reporting they had experienced difficulties accessing the online version.

Comparison With Prior Work

Most previous studies comparing online and paper survey respondents have been cross-sectional. They have found similar results to ours: younger and more educated people are more likely to respond online [13-15]. This effect was strong and consistent in our data for women born between 1946 and 1951 but less pronounced in the younger, more homogeneous cohort. Preference for paper questionnaires and postal surveys has also been found by others [7,13,14]. What our study adds is the longitudinal aspect, showing that respondents may choose to swap between online and paper modes of completion at successive surveys. While this swapping behavior might be due

XSL•FO
RenderX

in part to the time when participants received the invitations, it also suggests that many women preferred the paper format, which may enable slower, more deliberative response.

As others have reported [12], we found that online and paper responses produced equivalent scores on key health-related variables reported at the same time. While formatting the paper survey to more closely mimic the online version may have enhanced measurement equivalence across different groups of women, it may also have biased estimates of change over time.

Conclusions

Despite the cost-saving advantages of online compared to paper surveys, we found the latter appeal to a different population of potential respondents. This difference is likely to induce bias in the distribution of responses unless weighting for respondent characteristics (relative to the target population) is employed. From a longitudinal perspective, we also found that women who completed paper surveys were less likely to respond to the subsequent survey, adding to the potential bias. In conclusion, if mixed mode options are feasible, they are highly likely to produce more representative results than if the only response option is online.

Acknowledgments

The research on which this paper is based was conducted as part of the Australian Longitudinal Study on Women's Health by the University of Queensland and the University of Newcastle. We are grateful to the Australian Government Department of Health for funding and to the women who provided the survey data.

Conflicts of Interest

None declared.

References

1. Couper MP. Web surveys: a review of issues and approaches. Public Opin Q 2000;64(4):464-494. [doi: 10.1086/318641] [Medline: 11171027]
2. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
3. Uhlig CE, Seitz B, Eter N, Promesberger J, Busse H. Efficiencies of Internet-based digital and paper-based scientific surveys and the estimated costs and time for different-sized cohorts. PLoS One 2014;9(10):e108441 [FREE Full text] [doi: 10.1371/journal.pone.0108441] [Medline: 25313672]
4. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. J Med Internet Res 2007;9(3):e25 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]
5. Hohwü L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective. J Med Internet Res 2013;15(8):e173 [FREE Full text] [doi: 10.2196/jmir.2595] [Medline: 23978658]
6. Akl EA, Maroun N, Klocke RA, Montori V, Schünemann HJ. Electronic mail was not better than postal mail for surveying residents and faculty. J Clin Epidemiol 2005 Apr;58(4):425-429. [doi: 10.1016/j.jclinepi.2004.10.006] [Medline: 15862729]
7. Loban A, Mandefield L, Hind D, Bradburn M. A randomized trial found online questionnaires supplemented by postal reminders generated a cost-effective and generalizable sample but don't forget the reminders. J Clin Epidemiol 2017 Dec;92:116-125. [doi: 10.1016/j.jclinepi.2017.08.003] [Medline: 28844786]
8. Ebert JF, Huibers L, Christensen B, Christensen MB. Paper- or Web-based questionnaire invitations as a method for data collection: cross-sectional comparative study of differences in response rate, completeness of data, and financial cost. J Med Internet Res 2018 Jan 23;20(1):e24 [FREE Full text] [doi: 10.2196/jmir.8353] [Medline: 29362206]
9. Horevoorts NJ, Vissers PA, Mols F, Thong MS, van de Poll-Franse LV. Response rates for patient-reported outcomes using web-based versus paper questionnaires: comparison of two invitational methods in older colorectal cancer patients. J Med Internet Res 2015;17(5):e111 [FREE Full text] [doi: 10.2196/jmir.3741] [Medline: 25953059]
10. Rübsamen N, Akmatov MK, Castell S, Karch A, Mikolajczyk RT. Comparison of response patterns in different survey designs: a longitudinal panel with mixed-mode and online-only design. Emerg Themes Epidemiol 2017 Mar;14:4 [FREE Full text] [doi: 10.1186/s12982-017-0058-2] [Medline: 28344629]
11. Callas PW, Solomon LJ, Hughes JR, Livingston AE. The influence of response mode on study results: offering cigarette smokers a choice of postal or online completion of a survey. J Med Internet Res 2010;12(4):e46 [FREE Full text] [doi: 10.2196/jmir.1414] [Medline: 20965873]
12. de Leeuw ED, Hox JJ. Internet surveys as part of a mixed-mode design. In: Das M, Ester P, Kaczmirek L, editors. Social and Behavioral Research and the Internet. New York: Routledge; 2011:45-76.

XSL•FO
RenderX

13. Barentsz MW, Wessels H, van Diest PJ, Pijnappel RM, Haaring C, van der Pol CC, et al. Tablet, web-based, or paper questionnaires for measuring anxiety in patients suspected of breast cancer: patients' preferences and quality of collected data. J Med Internet Res 2014;16(10):e239 [FREE Full text] [doi: 10.2196/jmir.3578] [Medline: 25364951]
14. Mlikotic R, Parker B, Rajapakshe R. Assessing the effects of participant preference and demographics in the usage of Web-based survey questionnaires by women attending screening mammography in British Columbia. J Med Internet Res 2016 Mar 22;18(3):e70. [doi: 10.2196/jmir.5068] [Medline: 27005707]
15. Smith B, Smith TC, Gray GC, Ryan MAK, Millennium Cohort Study Team. When epidemiology meets the Internet: Web-based surveys in the Millennium Cohort Study. Am J Epidemiol 2007 Dec 1;166(11):1345-1354 [FREE Full text] [doi: 10.1093/aje/kwm212] [Medline: 17728269]
16. Jackle A, Lynn P, Burton J. Going online with a face-to-face household panel: effects of a mixed mode design on item and unit non-response. Surv Res Methods 2015;9(1):57-70. [doi: 10.18148/srm/2015.v9i1.5475]
17. Lynn P. Alternative sequential mixed-mode designs: effects on attrition rates, attrition bias, and costs. J Surv Stat Methodol 2013 Sep 18;1(2):183-205. [doi: 10.1093/jssam/smt015]
18. Cernat A. The impact of mixing modes on reliability in longitudinal studies. Soc Meth Res 2014 Oct 31;44(3):427-457. [doi: 10.1177/0049124114553802]
19. Allum N, Conrad F, Wenz A. Consequences of mid-stream mode-switching in a panel survey. Surv Res Meth 2018;12(1):43-58. [doi: 10.18148/srm/2018.v12i1.6779]
20. Dex S, Gumy J. On the experience and evidence about mixing models of data collection in large-scale surveys where the web is used as one of the modes in data collection.: National Centre for Research Methods Review URL: http://eprints.ncrm.ac.uk/2041/1/mixing_modes_of_data_collection_in_large_surveys.pdf [accessed 2019-02-10] [WebCite Cache ID 765aZXUf4]
21. 8146.0 - Household Use of Information Technology, Australia, 2014-15.: Australia Bureau of Statistics URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/8146.0Main+Features12014-15?OpenDocument [accessed 2019-02-10] [WebCite Cache ID 765af9fzz]
22. Lee C, Dobson AJ, Brown WJ, Bryson L, Byles J, Warner-Smith P, et al. Cohort profile: the Australian Longitudinal Study on Women's Health. Int J Epidemiol 2005 Oct;34(5):987-991 [FREE Full text] [doi: 10.1093/ije/dyi098] [Medline: 15894591]
23. Dobson AJ, Hockey R, Brown WJ, Byles JE, Loxton DJ, McLaughlin D, et al. Cohort profile update: Australian Longitudinal Study on Women's Health. Int J Epidemiol 2015 Oct;44(5):1547. [doi: 10.1093/ije/dyv110] [Medline: 26130741]
24. Pink B. Australian Standard Geographical Classification (ASGS).: Australian Bureau of Statistics; 2011 Jul. URL: http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/32FBEDE1EA4C5800CA25791F000F2E1C/$File/att98dqt.pdf [accessed 2019-02-24]
25. Australian guidelines to reduce health risks from drinking alcohol. Canberra: National Health and Medical Research Council, Australian Government; 2009. URL: https://nhmrc.gov.au/file/1641/download?token=DT3b2cTT [accessed 2019-02-10] [WebCite Cache ID 765aSyO9I]
26. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
27. Powers J, Loxton D. The impact of attrition in an 11-year prospective longitudinal study of younger women. Ann Epidemiol 2010 Apr;20(4):318-321. [doi: 10.1016/j.annepidem.2010.01.002] [Medline: 20149968]

Abbreviations

ANOVA: analysis of variance
BMI: body mass index
SF-36: 36-Item Short Form Health Survey
UKHLS: United Kingdom Household Longitudinal Study

Edited by G Eysenbach; submitted 02.04.18; peer-reviewed by M Callegaro, A Cernat; comments to author 14.10.18; revised version received 10.12.18; accepted 30.12.18; published 14.03.19.

Please cite as:
Fitzgerald D, Hockey R, Jones M, Mishra G, Waller M, Dobson A
Use of Online or Paper Surveys by Australian Women: Longitudinal Study of Users, Devices, and Cohort Retention
J Med Internet Res 2019;21(3):e10672
URL: http://www.jmir.org/2019/3/e10672/
doi:10.2196/10672
PMID:30869647

XSL•FO
RenderX

©David Fitzgerald, Richard Hockey, Mark Jones, Gita Mishra, Michael Waller, Annette Dobson. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based Form With Interactive Charts Used to Collect and Analyze Data on Home Births in Italy

Michele Zanetti[1], BEng; Rita Campi[1], MAT; Paola Olivieri[2], MID; Marta Campiotti[2], MID; Alice Faggianelli[2], MID; Maurizio Bonati[1], MD

[1]Laboratory for Mother and Child Health, Department of Public Health, Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy

[2]Associazione Nazionale Culturale Ostetriche Parto a Domicilio e Casa Maternità, Varese, Italy

Corresponding Author:
Michele Zanetti, BEng
Laboratory for Mother and Child Health
Department of Public Health
Istituto di Ricerche Farmacologiche Mario Negri
19 Via G La Masa
Milan, 20156
Italy
Phone: 39 0239014249
Fax: 39 023550924
Email: michele.zanetti@marionegri.it

Abstract

Background: The use of Web-based forms and data analysis can improve the collection and visualization of data in clinical research. In Italy, no register exists that collects clinical data concerning home births.

Objective: The purpose of this study was (1) to develop a Web portal to collect, through a Web-based form, data on home births in Italy and (2) to provide those interested with a graphic visualization of the analyses and data collected.

Methods: Following the World Health Organization's guidelines, and adding questions based on scientific evidence, the case report form (CRF) on the online form was drafted by midwives of the National Association of Out-of-Hospital Birth Midwives. During an initial phase, a group of midwives (n=10) tested the CRF, leading to improvements and adding the necessary questions to achieve a CRF that would allow a more complete collection of data. After the test phase, the entire group of midwives (n=166) registered themselves on the system and began filling out birth questionnaires. In a subsequent phase, the administrators of the portal were able to view the completed forms in a graphic format through the use of interactive maps and graphs.

Results: From 2014 to 2016, 58 midwives included 599 birth questionnaires via the Web portal; of these, 443 were home-based, 76% (321/424) of which were performed at home and 24% (103/424) at a midwifery unit. Most of the births assisted (79%, 335/424) were in northern Italy, and the average ages of the mother and father were 33.6 (SD 4.7) years and 37.0 (SD 5.6) years, respectively.

Conclusions: We developed an innovative Web-based form that allows, for the first time in Italy, the collection of data on home births and births in the midwifery unit. Furthermore, the data collected are viewable online by the midwives through interactive maps and graphs that allow them to have a general and continuously updated view of the situation of out-of-hospital births performed by the National Association of Out-of-Hospital Birth Midwives. The future goal is to be able to expand this data collection to all out-of-hospital births throughout the national territory. With an increase in the number of enrolled midwives, it would be possible to use the portal as a Web-based form and also as a portal for sharing resources that would help midwives in their clinical practice.

(J Med Internet Res 2019;21(3):e10335) doi:10.2196/10335

KEYWORDS

Web-based form; home birth; interactive charts; internet; survey methods

XSL•FO
RenderX

Introduction

From the mid-20th century, giving birth moved from the home into the hospital in most well-resourced countries. This change was related to the decreased rates of perinatal, neonatal, and maternal mortality observed in the same period. In-hospital birthing was one of the possible explanatory causes, perhaps the most important, as even living conditions, in general, began to change at that time. The rapid improvement of medical technologies and the availability of more knowledge and resources led to the over-medicalization of procedures once considered natural and physiological events, such as childbirth, which needs few obstetrical interventions when of low risk [1]. Thus, although birthing in a hospital is the cultural norm, in recent decades in a few countries, the number of out-of-hospital births is increased [2-4].

Rates of planned out-of-hospital births (ie, births intended to occur at home or at a freestanding birth unit) are low in high-income countries, although they vary widely. The highest rate is in the Netherlands; in the past, almost 30% of all babies born here were born at home, and home birth had always been an integrated part of the maternity system. Women in the Netherlands also currently have the option of giving birth in a birth center (a home-like setting); 11.4% of births occur in this setting and 16.3% at home [5,6]. In Wales, England, Scotland, Iceland, and Switzerland, out-of-hospital birth rates are 1% to 3%, whereas in other European countries rates are lower at 1% [5]. In the United States, although it recently increased, the rate is around 1.5% [7].

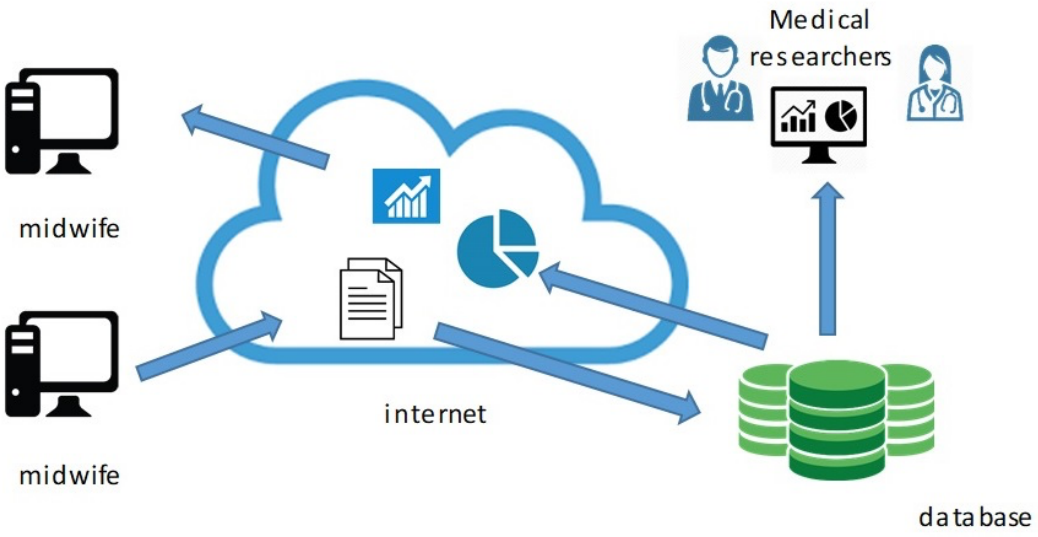
The Certificate of Delivery Assistance (CeDAP) is the nationwide mandatory questionnaire, completed by the midwife or physician attending the delivery, which collects information about the parents, the pregnancy, and the newborn baby. The CeDAP constitutes the richest source of health, epidemiological, and sociodemographic information concerning the birth event. This certificate is not always complete for out-of-hospital births and therefore does not contain all the information necessary to better assess the quality of births performed at home or in a midwifery unit.

To obtain more information on and to evaluate out-of-hospital births, it was necessary to provide the birth assistance certificate with additional questions; for this reason, a few years ago we began to collect paper data sheets and PDF forms. However, problems were found during collection. In fact, to be processed, the data collected had to be sent by mail to the processing site, which involved a waste of resources. In the last few years, the paper data collection questionnaires have been joined by increasingly advanced digital data collection systems: Web pages allow users to enter data collected in their studies and share it immediately with other users or simply analyze it for their own purposes (Figure 1).

The continuous development of information technology has allowed us to improve the computerized data collection questionnaires further, making them increasingly user-friendly for those entering the data collected during clinical practice.

With this information and with the help of the National Association of Out-of-Hospital Birth Midwives (called “the association” from here on), we decided to develop a portal [8] that would permit the collection of data on home births in Italy by the midwives who belong to the association.

Figure 1. Data flow overview.



XSL•FO
RenderX

Methods

Recruitment

The association is a national network of qualified midwives providing care and support to pregnant women and their babies before, during, and after childbirth. The association, founded in 1981, covers almost all midwives who attend out-of-hospital births. All qualified midwives who adhere to the association assist women who meet the eligibility criteria for out-of-hospital birth defined by national and international guidelines [9]. For the data acquisition, all 166 midwives who are part of the association participated. The midwives work privately, outside the national health system. They assist births at home or in private units run by midwives (a midwifery unit).

Portal Development

The portal was developed using ASP.NET WebForms, an open source Web framework for building modern Web apps and services with .Net. For data management, it used Microsoft SQL SERVER 2008 R2, a relational database management system with which all information was saved, loaded, and managed. In the database, the data from the forms were saved in JSON (JavaScript Object Notation) format.

Form Development

Case report form (CRF) variables were defined in an eXtensible Markup Language (XML), then automatically transformed into an HTML page using XSLT transformation; the HTML page contains tags and attributes managed by AngularJS, a framework JavaScript client-side compatible with both desktop and mobile browsers [10]. All data services (loading and saving) were handled via ANGULARJS and JQUERY as seen in Figure 2.

The XML, which contains the CRF variables, can be changed at any time and according to need (eg, additional CRF versions, new variables, bug detection) without having to rewrite other files or change the database structure. All the architecture described is downloadable from GitLab [11].

Data Collection

Following the World Health Organization's guidelines, and adding questions based on scientific evidence, the CRF was created by the midwives of the association. The CRF records every moment of pregnancy from the first visit to the end of each birth, including the weeks immediately following the birth (Figure 3). The main parts of the online questionnaires are described subsequently.

- *Information person completing form*: information regarding the midwife filling out the form. It is not necessarily the same person who followed the birth, but the general information is required.
- *First part*: to be completed at the beginning of the visit. Some information on the mother is collected (eg, age, diet, blood pressure).
- *Last third-trimester exams*: a report of the tests performed (eg, blood tests, fetal presentation).
- *Second part*: subsequent checks (only in the event of hospitalization). This part collects data on whether the exclusion from home care took place before labor. If the answer is affirmative, a series of questions must be filled out regarding the outcome of the childbirth and the motivations that led to exclusion.
- *Third part*: essential data concerning the childbirth (eg, place, date, time).
- *Fourth part*: information regarding the weeks following childbirth, type of breastfeeding, and whether the child has had any problems.
- *Lotus*: asks if lotus birth has been carried out. Lotus birth is the practice of leaving the umbilical cord uncut so that the baby remains attached to the placenta until the cord naturally separates at the navel as a cut cord does 3 to 10 days after birth. If a lotus birth was carried out, information on type of lotus and related information is collected, including the parent's experience.

Interactive Charts

The compilation of the questionnaires by midwives has made it possible to obtain a source of useful information not only for the group of clinicians involved in the processing of data but also for the midwives of the association. To facilitate the reading of the data, we created an area of the portal dedicated to the graphic visualization of the collected data. This area was created using Highcharts, a scalable vector graphics-based, multiplatform charting library, which makes it easy to add interactive charts [12].

On one page, for example, a map of Italy is divided into provinces, which are colored differently based on the number of questionnaires. By selecting a single province, information about the births in that area can be visualized by month and by place of birth as seen in Figure 4 the bar graph in the top right. The bottom graph shows births across Italy by month and place of birth.

Data Processing

Statistical Methods

The results presented compare women who gave birth at home to those who gave birth in a freestanding midwifery unit.

Categorical variables were summarized using proportions and associations tested using chi square or Fisher exact tests, where applicable. Continuous variables were summarized using means and standard deviations for normally distributed data, whereas skewed data were summarized using medians. A two-tailed independent *t* test was used to test differences of means for normally distributed continuous variables; the Mann-Whitney *U* test was used for skewed continuous variables.

To identify risk factors, we computed relative risks (RRs) using a multivariate log-binomial regression model, considering the significance of the confidence intervals. Statistical significance was evaluated using 95% confidence intervals and a two-tailed *P* value of <.05.

All data management and analyses were performed using SAS software.

XSL•FO
RenderX

Figure 2. The case report form system developed.

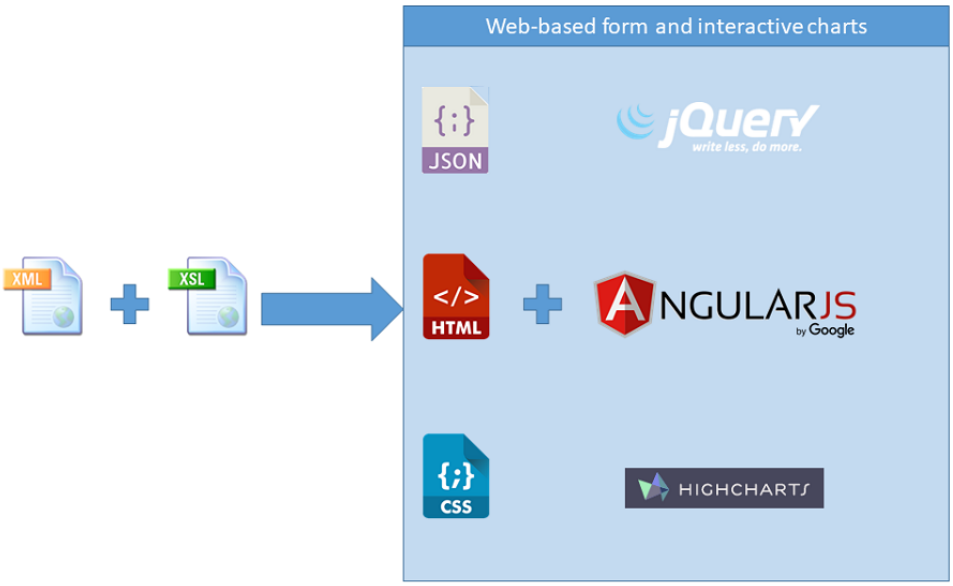
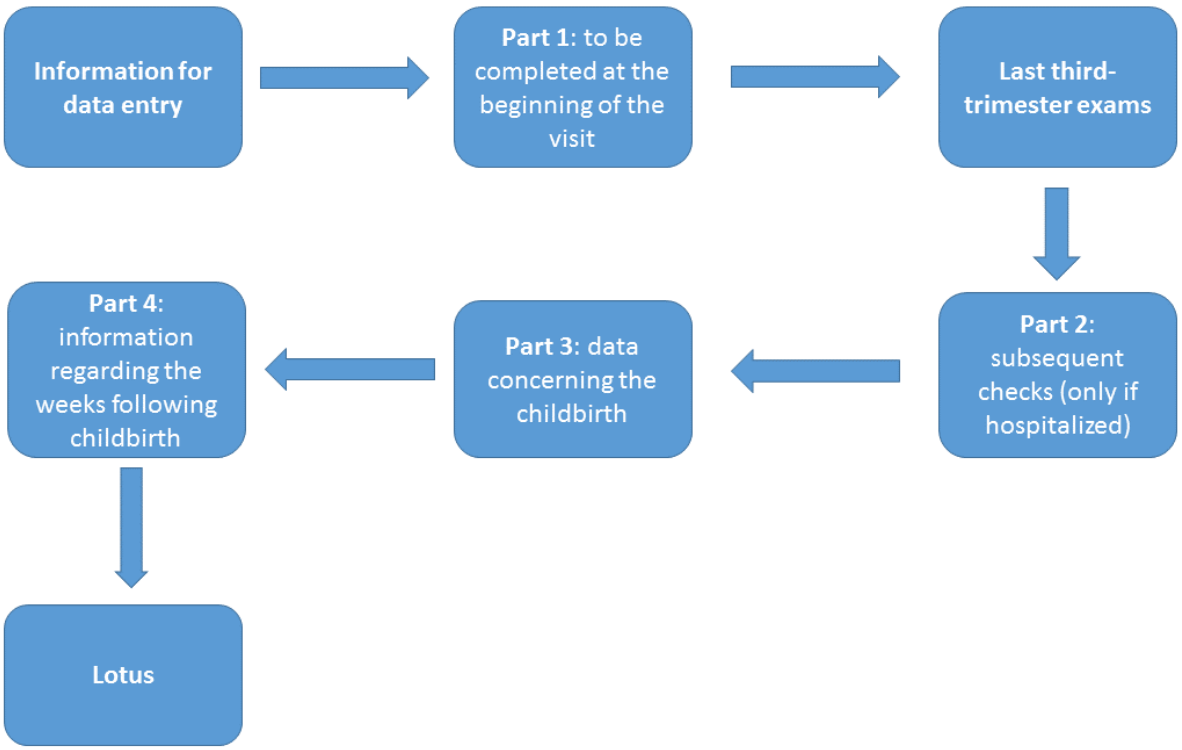
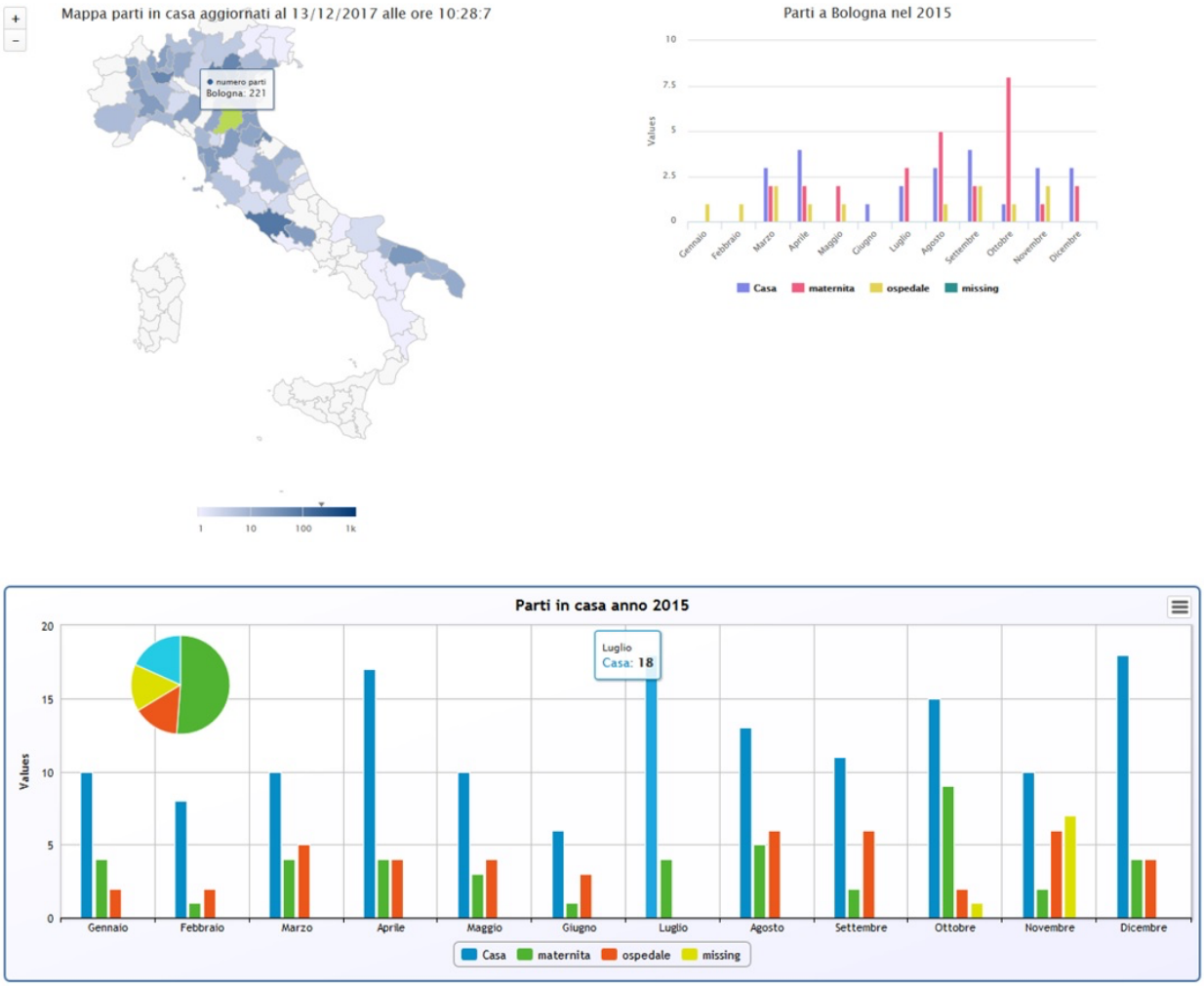


Figure 3. Case report form block diagram.



XSL•FO
RenderX

Figure 4. Interactive charts of births. The map of Italy is divided into provinces, colored differently based on the number of questionnaires. Selecting a single province provides information about the births in that area by month in a bar graph (top right). The bottom graph shows births across Italy by month and place of birth.



Results

Data were collected on 424 Italian women who delivered out-of-hospital, 321 (75.7%) of whom delivered at home and 103 (24.3%) at a freestanding midwifery unit. The majority of recruited deliveries (79.0%, 335/424) took place in northern Italy. The mean ages of the mothers and fathers were 33.6 (SD 4.7) years and mean 37.0 (SD 5.6) years, respectively (Table 1). The mothers had a higher level of education than the fathers, and the fathers worked more often (99%, 413/424) than mothers (78%, 331/424). The distributions of mothers for parity, age at first delivery, and educational level were different between mothers who delivered at home and those who delivered in a freestanding midwifery unit. In the multivariate regression model, being a primipara and having the first child at age 35 years or older was associated with a slightly higher risk of delivering in a freestanding midwifery unit compared to at home (age <35 years: RR 1.9, 95% CI 1.1-3.2; age ≥35 years: RR 2.3, 95% CI 1.1-4.8).

Of the 247 multiparas, the majority (63.2%, 156/247) had previously given birth in a hospital, and most had a normal vaginal birth (91.9%, 227/247). The majority of multiparas who delivered at home (76.1%, 188/247) had previously given birth in a hospital, whereas about half of the multiparas who delivered in a freestanding midwifery unit (51.4%, 127/247) had previously given birth at home.

A quarter of the mothers reported desire for intimacy as a reason for delivering out-of-hospital. Positive previous experience at home and a desire for a new experience were the other frequent reasons for justifying the choice of delivering at home, whereas naturalness and a trusting relationship with midwives were the main reasons for choosing to deliver in a freestanding midwifery unit.

XSL•FO
RenderX

Table 1. Characteristics of women having planned births at home or in a freestanding midwifery unit, and of their partners.

Characteristic	Delivery location			Relative risk (95% CI)	*F* (*df*1,*df*2)	*P* value
	At home (n=321)	In a freestanding midwifery unit (n=103)	Overall (N=424)			
Maternal characteristics						
Age (years), mean (SD)	33.5 (4.6)	33.9 (4.9)	33.6 (4.7)		1.13 (320,102)	.41
Age group (years), n (%)						
18-24	9 (2.8)	3 (2.9)	12 (2.8)	0.98 (0.70-1.37)		
25-34	171 (53.3)	52 (50.5)	223 (52.6)	Reference		
≥35	141 (43.9)	48 (46.6)	189 (46.6)	0.97 (0.87-1.09)		
Residential area, n (%)						
Large city	117 (42.1)	41 (50.0)	158 (43.9)	Reference		
Medium-size city	111 (39.9)	21 (25.6)	132 (36.7)	1.14 (1.01-1.28)		
Small town	50 (18.0)	20 (24.4)	70 (19.4)	0.96 (0.81-1.15)		
Missing	43	21	64			
Marital status, n (%)						
Married and/or cohabiting	229 (71.3)	74 (71.8)	303 (71.4)	Reference		
Other	92 (28.8)	29 (28.2)	121 (28.6)	1.25 (1.10-1.42)		
Number of children, n (%)						
First	100 (32.9)	51 (54.3)	151 (37.9)	Reference		
Second or more	204 (67.1)	43 (45.7)	247 (62.1)	1.25 (1.10-1.42)		
Missing	17	9	26			
First delivery >35 years, n (%)						
Yes	22 (6.9)	20 (19.4)	42 (9.9)	0.67 (0.50-0.90)		
No	299 (93.1)	83 (80.6)	382 (90.1)	Reference		
Level of education, n (%)						
Primary	5 (1.6)	3 (2.9)	8 (1.9)	Reference		
Secondary	102 (31.9)	29 (28.2)	131 (31.0)	1.25 (0.72-2.15)		
Postsecondary	214 (66.6)	71 (68.6)	285 (67.1)	1.20 (0.70-2.06)		
Occupational status before index birth, n (%)						
Working	246 (76.6)	85 (82.4)	331 (78.0)	Reference		
Not working	75 (23.4)	18 (17.6)	93 (22.0)	1.08 (0.96-1.22)		
Annual income (€), n (%)						
<20,000	41 (15.1)	10 (10.5)	51 (13.9)	Reference		
20,000-29,000	109 (40.1)	33 (34.7)	142 (38.7)	0.95 (0.81-1.12)		
≥30,000	122 (44.9)	52 (54.7)	174 (47.4)	0.57 (0.74-1.03)		
Missing	49	8	57			
Diet, n (%)						
Omnivorous	238 (74.1)	80 (77.7)	318 (75.0)	Reference		
Other	83 (25.9)	23 (22.3)	106 (25.0)	1.05 (0.93-1.18)		
Smoker, n (%)						
Yes	10 (3.1)	7 (6.8)	17 (4.0)	0.77 (0.52-1.15)		
No	311 (96.9)	96 (93.2)	407 (96.0)	Reference		
Partner characteristics						

XSL•FO
RenderX

Characteristic	Delivery location			Relative risk (95% CI)	*F* (*df*1,*df*2)	*P* value
	At home (n=321)	In a freestanding midwifery unit (n=103)	Overall (N=424)			
Age (years), mean (SD)	36.7 (5.5)	37.7 (6.2)	37.0 (5.6)		1.27 (306,96)	.12
Level of education, n (%)						
Primary	25 (7.9)	8 (7.9)	33 (7.9)	Reference		
Secondary	134 (42.5)	47 (46.5)	181 (43.5)	0.98 (0.79-1.21)		
Postsecondary	156 (49.5)	46 (45.5)	202 (48.6)	1.02 (0.83-1.25)		
Missing	6	2	8			
Occupational status, n (%)						
Working	312 (98.7)	101 (100.0)	413 (99.0)	Reference		
Not working	4 (1.3)	0	4 (1.0)	1.32 (1.25-1.40)		
Missing	5	2	7			

The position most frequently used by women for delivering at home was on all fours; for delivering in a freestanding midwifery unit, it was squatting (Table 2). Delivering with two or more midwives was slightly more frequent in a freestanding midwifery unit, as was the use of a uterotonic agent (mainly oxytocin) at birth. None of the other monitored obstetric and neonatal parameters differed between the two delivery settings. No third- or fourth-degree perineal tears were observed in the studied population, and only two episiotomies were performed. Within one week of delivery, one mother and eight newborns were hospitalized, all after delivering at home, and all were discharged from the hospital after a few days.

XSL•FO
RenderX

Table 2. Birth-related characteristics and birth outcomes of women having planned births at home or in a freestanding midwifery unit.

Characteristics	Delivery location			Relative risk (95% CI)	*Z*	*F* (*df*1,*df*2)	*P* value
	At home (n=321)	In a freestanding midwifery unit (n=103)	Overall (N=424)				
Birth-related							
Gestational age (weeks), median (IQR[a])	40 (1)	40 (1)	40 (1)		0.75		.45
Birthweight (g), mean (SD)	3419.6 (451.4)	3397.1 (392.8)	3414.1 (437.3)			1.32 (312,102)	.22
Small for gestational age, n (%)							
Yes	23 (7.2)	14 (13.6)	37 (8.7)	0.81 (0.62-1.04)			
No	298 (92.8)	89 (86.4)	387 (91.3)	Reference			
Position in delivering, n (%)							
Lying down	46 (14.3)	30 (29.1)	76 (17.9)	0.94 (0.75-1.17)			
Squatting	75 (23.4)	41 (39.8)	116 (27.4)	Reference			
Kneeling	18 (5.6)	1 (1.0)	19 (4.5)	1.47 (1.23-1.74)			
On all fours	130 (40.5)	16 (15.5)	146 (34.4)	1.38 (1.19-1.59)			
On the side	32 (10.0)	15 (14.6)	47 (11.1)	1.05 (0.83-1.34)			
Others	20 (6.2)	0	20 (4.7)	—[b]			—
Number of midwives at delivery, n (%)							
1	23 (7.2)	1 (1.0)	24 (5.7)	1.29 (1.16-1.42)			
≥2	298 (92.8)	102 (99.0)	400 (94.3)	Reference			
Cord clamping (min), mean (SD)	84.2 (106.7)	88.8 (44.8)	85.3 (95.4)			1.25 (224,71)	.18
Uterotonic agent use, n (%)							
Yes	98 (30.5)	45 (43.7)	143 (24.3)	1.43 (1.09-1.88)			
No	223 (69.5)	58 (56.3)	281 (75.7)	Reference			
Lotus, n (%)							
Yes	76 (23.7)	27 (26.2)	103 (24.3)	0.97 (0.85-1.10)			
No	245 (76.3)	76 (73.8)	321 (75.7)	Reference			
Exclusive breastfeeding at 10 days, n (%)							
Yes	314 (93.7)	100 (97.1)	414 (94.6)	Reference			
No	7 (2.2)	3 (2.9)	10 (2.4)	0.92 (0.61-1.39)			
Birth outcomes							
Postpartum hemorrhage, n (%)							
≤500 mL	292 (91.0)	97 (94.2)	389 (91.7)	Reference			
>500 mL	29 (9.0)	6 (5.8)	35 (8.3)	1.61 (0.65-3.98)			
Perineal tear (degree), n (%)							
None	171 (53.3)	57 (55.3)	228 (53.8)	Reference			
First	103 (32.1)	33 (32.0)	136 (32.1)	1.01 (0.89-1.74)			
Second	45 (14.0)	12 (11.7)	57 (13.4)	1.05 (0.90-1.23)			
Third	0	0	0	—			—
Fourth	0	0	0	—			—

XSL•FO
RenderX

Characteristics	Delivery location			Relative risk (95% CI)	*Z*	*F* (*df*1,*df*2)	*P* value
	At home (n=321)	In a freestanding midwifery unit (n=103)	Overall (N=424)				
Mother's postpartum hospitalization (within 1 week of delivery), n	1	0	1	—			—
Newborn's hospitalization (within 1 week of birth), n	8	0	8	—			—

[a]IQR: interquartile range.

[b]Cannot be calculated because the frequency is 0.

Discussion

Principal Results

The population enrolled in this first Italian study corresponded to 47% of the expected national deliveries over the study period and for all causes of out-of-hospital births. Considering that the investigated population was selected and only low-risk, planned out-of-hospital births were the target population, the findings are representative and support the choice of out-of-hospital planned births in Italy.

Being older than 35 years and being a primipara increased the probability of the delivery occurring in a freestanding midwifery unit compared to at home. Findings are in line with previous studies performed in different European countries, as well as in the United States and Canada [13-18]. Among the reasons that influence women in planning the place of birth are cultural attitudes, religion, and peer and family views, but another determinant guiding the choice to deliver out-of-hospital is also previous birth experience [19-21]. These reasons are valid for comparable settings in countries with middle and high availability of resources. In developing countries, where out-of-hospital birth rates are high, factors such as poverty, access to hospitals, and lack of transportation determine the choice [22,23].

Strengths and Limitations

This study had limited power to detect small differences in variables with low incidence. As in similar countries [24], the rate of out-of-hospital births in Italy is low, so it is difficult to obtain large study groups. Therefore, the results of this study can be indicative for settings with similar services and societal structures. The strength of the study was its ability to obtain for the first time, using a rigorous data collection process, detailed information on out-of-hospital births in Italy for a large population that was based on a formal, updated, and evidence-based assistance protocol. The first national dataset was created and provided reasonable detail in terms of women's characteristics, pregnancy monitoring, labor, birth, and neonatal outcomes of interest. These data have allowed the body of knowledge to be expanded by providing evidence on the results of comparisons performed between delivery settings.

Conclusions

This study made it possible, for the first time, to obtain a large amount of information about out-of-hospital births in Italy. The tools used for data collection and visualization have allowed us to optimize the acquisition and monitor the information. The future goal is to be able to expand this data collection to all out-of-hospital births in the national territory. This would also allow midwives to have a more complete and detailed view of the work they perform and help monitor and improve the clinical practice of out-of-hospital births. Moreover, with an increase in the number of enrolled midwives, it would be possible to use the portal not only as a Web-based form but also as a portal for sharing resources that would help midwives in their clinical practice. The interactive graphs and maps [25] used for visualizing and processing data can be valuable instruments for sharing results.

Acknowledgments

The authors would like to thank Matteo Mondini, Daniele Crespi, and Claudio Previtali (Istituto di Ricerche Farmacologiche Mario Negri) for their contributions to the development of the key aspect of this research project, and Chiara Pandolfini and Maria Grazia Calati for manuscript editing.

No specific funding was received to conduct the study. The costs were covered by resources from both the Laboratory for Mother and Child Health and the National Association of Out-of-Hospital Birth Midwives.

Authors' Contributions

MB and MC conceptualized the study. PO and AF coordinated and supervised data collection. MZ guaranteed informatics forms and resources for the collection and management of data. RC carried out the initial analyses. MB drafted the initial manuscript. All authors interpreted data and critically reviewed and revised the final manuscript as submitted. All authors agree to be accountable for aspects of their contribution to the work.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Improving Birth Coalition. The mother-friendly childbirth initiative URL: http://www.motherfriendly.org/MFCI/steps [accessed 2018-10-29] [WebCite Cache ID 73X4OJJF4]
2. American Academy of Pediatrics. Planned Home Birth. Pediatrics 2013 Apr 29;131(5):1016-1020. [doi: 10.1542/peds.2013-0575]
3. College of Midwives of British Columbia. Registrant's handbook: home birth standards URL: http://cmbc.bc.ca/standards-policies-forms/standards-policies-and-forms/ [accessed 2018-03-02] [WebCite Cache ID 6xcALUM3G]
4. National Institute for Health and Care Excellence. 2017 Feb. Intrapartum care for healthy women and their babies during childbirth URL: http://www.nice.org.uk/guidance/cg190 [accessed 2018-02-27] [WebCite Cache ID 6xcAScLDA]
5. Euro-Peristat. The European Perinatal Health Report: health and care of pregnant women and babies in Europe in 2010 URL: http://www.europeristat.com/ [accessed 2018-03-02] [WebCite Cache ID 6xcAbxKNp]
6. Zielinski R, Ackerson K, Kane LL. Planned home birth: benefits, risks, and opportunities. Int J Womens Health 2015;7:361-377 [FREE Full text] [doi: 10.2147/IJWH.S55561] [Medline: 25914559]
7. MacDorman MF, Declercq E, Cabral H, Morton C. Recent increases in the US maternal mortality rate: disentangling trends from measurement issues. Obstet Gynecol 2016 Sep;128(3):447-455 [FREE Full text] [doi: 10.1097/AOG.0000000000001556] [Medline: 27500333]
8. Associazione Nazionale Culturale Ostetriche Parto a Domicilio e Casa Maternità, Istituto Mario Negri di Milano. Nascere in Casa URL: https://nascereacasa.marionegri.it/ [accessed 2018-02-23] [WebCite Cache ID 6xS2NCZSc]
9. Campiotti M, Dalle Pezze M, Forasacco L, Iop P, Montani M, Olivieri P, Associazione Nascere in Casa. Linee guida di assistenza al travaglio e parto fisiologico a domicilio e casa maternità. 2017. URL: http://www.nascereacasa.it/lineeguida_revisione2017/ [accessed 2018-02-23] [WebCite Cache ID 6xS3BGoEH]
10. AngularJS. URL: https://angularjs.org/ [accessed 2018-10-29] [WebCite Cache ID 73X7lMCiI]
11. GitLab. Gatherings Crafter URL: https://gitlab.com/dellaltromondo/gatherings-crafter [accessed 2018-10-29] [WebCite Cache ID 73X7Tnjur]
12. Highcharts. URL: https://www.highcharts.com/products/highcharts/ [accessed 2018-10-29] [WebCite Cache ID 73X4zV1xL]
13. Hildingsson IM, Lindgren HE, Haglund B, Rådestad IJ. Characteristics of women giving birth at home in Sweden: a national register study. Am J Obstet Gynecol 2006 Nov;195(5):1366-1372. [doi: 10.1016/j.ajog.2006.03.050] [Medline: 16707081]
14. Birthplace in England Collaborative Group, Brocklehurst P, Hardy P, Hollowell J, Linsell L, Macfarlane A, et al. Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study. BMJ 2011 Nov 23;343:d7400 [FREE Full text] [Medline: 22117057]
15. de Jonge A, Geerts CC, van der Goes BY, Mol BW, Buitendijk SE, Nijhuis JG. Perinatal mortality and morbidity up to 28 days after birth among 743 070 low-risk planned home and hospital births: a cohort study based on three merged national perinatal databases. BJOG 2015 Apr;122(5):720-728 [FREE Full text] [doi: 10.1111/1471-0528.13084] [Medline: 25204886]
16. Hutton EK, Cappelletti A, Reitsma AH, Simioni J, Horne J, McGregor C, et al. Outcomes associated with planned place of birth among women with low-risk pregnancies. CMAJ 2016 Mar 15;188(5):E80-E90 [FREE Full text] [doi: 10.1503/cmaj.150564] [Medline: 26696622]
17. Hutton EK. The safety of home birth. J Obstet Gynaecol Can 2016 Apr;38(4):331-333. [doi: 10.1016/j.jogc.2016.02.005] [Medline: 27208600]
18. Snowden JM, Tilden EL, Snyder J, Quigley B, Caughey AB, Cheng YW. Planned out-of-hospital birth and birth outcomes. N Engl J Med 2015 Dec 31;373(27):2642-2653 [FREE Full text] [doi: 10.1056/NEJMsa1501738] [Medline: 26716916]
19. Coxon K, Sandall J, Fulop NJ. How do pregnancy and birth experiences influence planned place of birth in future pregnancies? Findings from a longitudinal, narrative study. Birth 2015 Jun;42(2):141-148 [FREE Full text] [doi: 10.1111/birt.12149] [Medline: 25676885]
20. Halfdansdottir B, Olafsdottir OA, Hildingsson I, Smarason AK, Sveinsdottir H. Maternal attitudes towards home birth and their effect on birth outcomes in Iceland: a prospective cohort study. Midwifery 2016 Mar;34:95-104. [doi: 10.1016/j.midw.2015.12.010] [Medline: 26809368]
21. Happel-Parkins A, Azim KA. At pains to consent: a narrative inquiry into women's attempts of natural childbirth. Women Birth 2016 Aug;29(4):310-320. [doi: 10.1016/j.wombi.2015.11.004] [Medline: 26639037]
22. Moindi RO, Ngari MM, Nyambati VCS, Mbakaya C. Why mothers still deliver at home: understanding factors associated with home deliveries and cultural practices in rural coastal Kenya, a cross-section study. BMC Public Health 2016 Feb 03;16:114 [FREE Full text] [doi: 10.1186/s12889-016-2780-z] [Medline: 26842657]
23. Sarker BK, Rahman M, Rahman T, Hossain J, Reichenbach L, Mitra DK. Reasons for preference of home delivery with traditional birth attendants (TBAs) in rural Bangladesh: a qualitative exploration. PLoS One 2016;11(1):e0146161 [FREE Full text] [doi: 10.1371/journal.pone.0146161] [Medline: 26731276]

XSL•FO
RenderX

24. Halfdansdottir B, Smarason AK, Olafsdottir OA, Hildingsson I, Sveinsdottir H. Outcome of planned home and hospital births among low-risk women in Iceland in 2005-2009: a retrospective cohort study. Birth 2015 Mar;42(1):16-26. [doi: 10.1111/birt.12150] [Medline: 25613161]
25. Highsoft. URL: https://www.highcharts.com/ [accessed 2018-02-23] [WebCite Cache ID 6xS5PzMfT]

Abbreviations

CeDAP: birth assistance certificate [Certificato Di Assistenza al Parto]
CRF: case report form
IQR: interquartile range
RR: relative risk

Edited by G Eysenbach; submitted 08.03.18; peer-reviewed by S Tao, R Hauser; comments to author 05.09.18; revised version received 30.10.18; accepted 06.11.18; published 22.03.19.

Please cite as:
Zanetti M, Campi R, Olivieri P, Campiotti M, Faggianelli A, Bonati M
A Web-Based Form With Interactive Charts Used to Collect and Analyze Data on Home Births in Italy
J Med Internet Res 2019;21(3):e10335
URL: https://www.jmir.org/2019/3/e10335/
doi:10.2196/10335
PMID:30900993

©Michele Zanetti, Rita Campi, Paola Olivieri, Marta Campiotti, Alice Faggianelli, Maurizio Bonati. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factors Influencing Web-Based Survey Response for a Longitudinal Cohort of Young Women Born Between 1989 and 1995

Deborah Loxton[1*], PhD; Melissa L Harris[1*], BAPsyc (Hons), PhD; Peta Forder[1], MBiostat; Jennifer Powers[1], MMedStat; Natalie Townsend[1], BPsych (Hons); Julie Byles[1], PhD; Gita Mishra[2], PhD

[1]Research Centre for Generational Health and Ageing, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia
[2]Institute for Social Science Research, Faculty of Humanities and Social Sciences, The University of Queensland, Herston, Australia
[*]these authors contributed equally

Corresponding Author:
Melissa L Harris, BAPsyc (Hons), PhD
Research Centre for Generational Health and Ageing
Faculty of Health and Medicine
University of Newcastle
University Drive
Callaghan, 2308
Australia
Phone: 61 240420621
Fax: 61 240420044
Email: melissa.harris@newcastle.edu.au

Abstract

Background: With health research practices shifting toward rapid recruitment of samples through the use of online approaches, little is known about the impact of these recruitment methods on continued participation in cohort studies.

Objective: This study aimed to report on the retention of a cohort of young women who were recruited using an open recruitment strategy.

Methods: Women from the 1989-95 cohort of the Australian Longitudinal Study on Women's Health, recruited in 2012 and 2013 were followed up annually via Web-based surveys in 2014, 2015, and 2016. Prevalence ratios for survey response were calculated using log-binomial models with generalized estimating equations including demographic, health-related, and recruitment method characteristics examined as explanatory factors.

Results: Of the 17,012 women who completed the baseline survey (Survey 1) in 2012 to 2013, approximately two-thirds completed Survey 2 (2014), and just over half completed Surveys 3 (2015) and 4 (2016). Women demonstrated transient patterns of responding with 38.21% (6501/17,012) of women completing all 4 surveys. Although retention of young women was associated with older age, higher education, higher self-rated health status, and low engagement with adverse health behaviors, the method of recruitment was a key determinant of study participation in the multivariate model. Although women were more likely to be recruited into the cohort via social media (eg, Facebook), retention over time was higher for women recruited through traditional media and referral approaches.

Conclusions: A balance must be obtained between achieving representativeness, achieving rapid cohort recruitment, and mitigating the pitfalls of attrition based on recruitment method in the new era of cohort studies, where traditional recruitment methods are no longer exclusively viable options.

(J Med Internet Res 2019;21(3):e11286) doi:10.2196/11286

KEYWORDS

retention; attrition; young adult; Web-based survey; social media; women's health

XSL•FO
RenderX

Introduction

Participation in epidemiologic studies has been declining over the past 30 years. This has been attributed to declining volunteerism, the proliferation of market research studies, perceived irrelevance, and increased demands on participant involvement [1]. Together with the changing technological landscape and shift toward digital communication practices, this has led to innovative approaches to participant recruitment that take advantage of online recruitment through social media advertising and blog-seeding [2-4]. Although these techniques might be successful in recruiting participants to a single study [2,5], little is known about the impact of these recruitment techniques on participant retention in cohort studies.

A review of the current literature revealed consistent associations between sociodemographic and health factors and attrition. In conventional cohort studies using paper-based surveys, attrition (permanent or temporary loss to follow-up) is more likely to occur if participants are younger [6,7], less educated [8,9], have poorer health [6,7,10], or are smokers [6,11]. Retention (active participation at a survey wave) within conventional longitudinal studies among younger people has been credited to ongoing contact with participants, via such avenues as regular postcards and phone calls [12-14]. As internet-based research increases, it is important to understand what leads participants to return to a Web-based survey and what biases might be introduced by the disparities between those who return and those who do not, particularly for young people who are notoriously hard to engage in longitudinal health research [1,2]. It is also important to understand the influence of sociodemographic factors on recruitment method and the joint impact on retention. For example, we found that recruitment method differed by age [3] and, on the basis of past research, we would expect age to also influence retention [6,7].

The aim of this study was to examine factors that influence retention in a longitudinal cohort study of young women who were openly recruited through the internet and social media, as well as traditional media recruitment methods [3]. Specifically, we wished to identify the characteristics associated with retention, with a particular focus on whether retention was influenced by different recruitment methods. On the basis of past research, we hypothesized that women who were younger [6,7], had lower levels of education [8,9], had relatively poorer health [6,7,10], and who used tobacco [6,11] would be less likely to continue completing surveys compared with other women. We further hypothesized that the recruitment method may influence retention and that some of this influence may be due to age [3].

Methods

Overview of the Study Design

This study included data from the 1989-95 cohort of the Australian Longitudinal Study on Women's Health (ALSWH), a national population-based health study of Australian women. Details regarding the recruitment strategy for this cohort are described in detail elsewhere [3]. Briefly, however, between 2012 and 2013, women born between 1989 and 1995 (aged 18 to 23 years) were recruited using an open recruitment method that relied on both social media as well as traditional recruitment strategies. Additional eligibility for study inclusion included living in Australia, possessing a Medicare (ie, Australia's national health insurer) number (Australian and New Zealand citizens and permanent residents living in Australia are eligible for a Medicare number), and consenting to have survey data linked with administrative data (eg, records of health service use). Verification for data linkage was conducted by the Australian Department of Human Services matching participants on name, address, date of birth, and Medicare number.

Unlike the previous ALSWH cohorts who were recruited through the Medicare database, recent recruitment of young women demonstrated that this approach was no longer a viable option for the recruitment of young women for the purpose of conducting health surveys [15]. Recruitment strategies for the new cohort of women aged 18 to 23 years included paid Facebook advertising, promotion using social and other internet-based media (eg, forum posts, Gumtree advertisements, and Twitter), paid and unpaid promotion through traditional media (eg, face-to-face events, posters, print, and radio media promotion), incentives (eg, chance to win a AU $50 gift voucher or fashion promotion which entailed a chance to win an exclusive pair of Black milk tights), and peer referral. Owing to the slow rate of recruitment at the beginning, the campaign comprised 2 promotions: the first delivered by ALSWH (October 2012 to December 2013) and the second by a marketing company under the branding of Women's Health of Australia! (October 2013 to December 2013). The slogans, branding, and incentives offered under each of the promotions are described in detail elsewhere [3].

The 17,012 women who responded to the open recruitment invitation and completed the Web-based survey were found to be broadly representative of similarly aged Australian women in terms of demographics, with some over-representation of more educated women [5]. Unlike the original 3 ALSWH cohorts who completed surveys on a 3-year rolling schedule, the newest cohort has completed yearly surveys.

To meet the aims of the study, surveys were focused on demographic, economic and social factors, health behaviors, self-reported anthropometric measures, physical and mental health, and health service use. To reduce participant burden, not all questions are asked at every survey, with a core set of items included and themed items included in every second or third survey (eg, complementary and alternative therapies are included in Survey 2 (measured in 2014), whereas the additional theme for Survey 3 (measured in 2015) was on parental socioeconomic position and adverse childhood experiences) [16].

Participants

Data for this study were obtained from the ALSWH 1989-95 cohort who completed surveys in 2012 to 2013 (Survey 1), 2014 (Survey 2), and 2015 (Survey 3).

Measures

Response status: The women were resurveyed annually via Web-based surveys in 2014, 2015, and 2016. Women were

XSL•FO
RenderX

classified as respondents to a particular survey if they provided answers to at least 5 items from that survey. The number of questions varied between 56 and 113 across the 3 follow-up surveys.

Demographic variables at Survey 1 included age and area of residence. Area of residence was based on the Accessibility/Remoteness Index of Australia Plus index of distance from the nearest urban center (major cities; inner regional; and outer regional, remote, or very remote) [17]. The highest level of education completed was classified as (1) less than 12 years of schooling, (2) Grade 12 or equivalent, (3) certificate or diploma, and (4) university. Student and employment status were each classified as not studying or not employed and part-time or full-time. The ability to manage with available income was reported as easy, not too bad, difficult some of the time, difficult all the time, or impossible. Relationship status was classified as married; living in a de facto relationship; or not married or partnered.

Health-related characteristics at Survey 1 included self-rated health, psychological distress, health behaviors, body mass index (BMI), physical activity, and experience of partner violence. Self‐rated health was measured by the general health item, "In general would you say your health is:" with response options of excellent; very good; good; fair; or poor [18]. Psychological distress was measured by the Kessler Psychological Distress Scale (K10) [19], which comprises 10 items that measure recent depression and anxiety. The summed responses to these items were categorized to indicate low, moderate, high, and very high levels of psychological distress [20]. Smoking was classified into never smoker, ex-smoker, and current smoker. Alcohol consumption was based on the level of risk identified according to the 2009 guidelines [21] with episodic risk defined as drinking 5 or more standard drinks on 1 occasion and long-term risk as drinking more than 2 drinks per day on average. Using these definitions, 4 categories of alcohol consumption were used: (1) no risk; (2) low long-term risk, low episodic risk; (3) low long-term risk, high episodic risk; and (4) high long-term risk, irrespective of episodic risk. Marijuana use and other illicit drug use were categorized according to most recent use: never; recent use (in the last 12 months); past use (more than 12 months ago); recent and past use. BMI (kg/m^2) was categorized according to World Health Organization guidelines as underweight (BMI less than 18.5), healthy weight (BMI 18.5 to 24.9), overweight (BMI 25 to 29.9), and obese (BMI 30 or more) [22]. Physical activity level was based upon the frequency and duration of leisure-time activity in the last week and classified as sedentary, low, moderate, or high [23]. Partner violence was an affirmative response to the item: "Have you ever been in a violent relationship with a partner or spouse?"

Recruitment method: When women were first recruited into the 1989-95 cohort, an item in the baseline survey asked them what had led them to take part in the study. Responses were grouped into 5 broad recruitment methods: (1) Facebook, including posts on the ALSWH pages and paid advertising; (2) other online media (eg, Twitter, Tumblr, and blogs); (3) referral (eg, direct contact from ALSWH staff, referral from professional and personal networks, and existing ALSWH participants); (4) traditional media (eg, television, radio, and magazine advertising); and (5) a fashion promotion, which involved an incentive of the chance to win a pair of exclusive tights [3].

Analyses

Baseline characteristics at Survey 1, including mode of recruitment, are presented according to participants who responded to follow-up surveys (respondents) versus participants who did not (nonrespondents). Response rates across the first 3 follow-up surveys are presented in a multimedia appendix according to pattern of response, demographic, and health-related characteristics as well as recruitment method.

Prevalence ratios (PR) for survey response were calculated using a log-binomial model with generalized estimating equations to account for the correlation of repeated observations across surveys. A total of 3 models were investigated incorporating explanatory variables measured at baseline: (1) survey; (2) survey and recruitment method; and (3) survey, recruitment method, selected demographic, and health-related variables. An interaction between age and recruitment method was also investigated. Results were deemed to be statistically significant if $P<.01$, and analyses were conducted using SAS Software 9.4 (TS1M3) for Windows (SAS Institute Inc, Cary, NC, USA).

Results

In 2012 to 2013, 17,012 women completed Survey 1 (baseline). Surveys were conducted annually thereafter, with 66.68% (11,344/17,012) of women completing the second survey in 2014, 52.67% (8961/17,012) of women completing the third survey in 2015, and 52.94% (9007/17,012) of women completing the fourth survey in 2016. All 3 follow-up surveys were completed by 38.21% (6501/17,012) of women, whereas 19.54% (3324/17,012) of women completed 2 follow-up surveys, 18.58% (3161/17,012) of women completed 1 follow-up survey only, and 23.67% (4026/17,012) women did not respond to any of the follow-up surveys. The frequency of response patterns across 3 follow-up surveys is presented in Table 1.

Women who responded at subsequent surveys (2, 3, or 4) were more likely to be older at baseline, reported better health, and were more likely to manage on their available income at baseline, whereas women who did not respond again after Survey 1 were more likely to report smoking, drinking, or other drug use at baseline, be sedentary, or report experiencing domestic violence (see Multimedia Appendix 1). Responders were more likely to have been recruited via referral or traditional media methods, whereas nonrespondents were more likely to report being recruited through Facebook and other social media strategies.

Response rates across the first 3 follow-up surveys are presented in Multimedia Appendix 2.

XSL•FO
RenderX

Table 1. Frequency of response patterns across 3 follow-up surveys (Survey 2, Survey 3, and Survey 4) for the women who completed Survey 1 (N=17,012).

Pattern of response to follow-up surveys	Statistics, n (%)
0 follow-up surveys	4026 (23.67)
1 follow-up survey only	3161 (18.58)
Survey 2 only	2089 (12.28)
Survey 3 only	474 (2.79)
Survey 4 only	598 (3.52)
2 follow-up surveys	33324 (19.54)
Survey 2 and Survey 3	1416 (8.32)
Survey 2 and Survey 4	1338 (7.87)
Survey 3 and Survey 4	570 (3.35)
3 follow-up surveys	6501 (38.21)
Survey 2, Survey 3, and Survey 4	6501 (38.21)

Participants were 20% less likely to respond after the first follow-up survey (95% CI 0.79-0.80, see Table 2). Compared with women who were recruited via Facebook, women who were recruited via referral or traditional media were more likely to respond at subsequent surveys (PR=1.11; 95% CI 1.07-1.15 and PR=1.18; 95% CI 1.14-1.22, respectively), whereas some recruited via the fashion promotion strategy were slightly less likely to respond (PR=0.94; 95% CI 0.91-0.97). Women were less likely to respond if they had not completed Grade 12 at baseline (PR=0.80; 95% CI 0.76-0.84), had extreme difficulties managing with their available income (PR=0.95; 95% CI 0.93-0.97), were smokers (PR=0.82; 95% CI 0.79-0.85), or had very high levels of psychological distress (PR=0.96; 95% CI 0.93-0.98). There was no evidence of an interaction between age and recruitment method (*P*>.01, data not shown).

XSL•FO
RenderX

Table 2. Factors associated with survey response over 3 follow-up surveys using baseline (Survey 1) characteristics as explanatory variables (N=17,012; 187 participants were excluded due to missing data). The included explanatory factors for predicting response at follow-up surveys were model 1 (survey), model 2 (survey + recruitment method), and model 3 (survey + recruitment method + other characteristics; eg, selected demographic/health factors).

Baseline characteristics	Model 1		Model 2		Model 3	
	PR (95% CI)[a]	*P* value	PR (95% CI)	*P* value	PR (95% CI)	*P* value
Survey						
2 (2014)	1	—[b]	1	—	1	—
3 (2015)	0.79 (0.78-0.80)	<.001	0.79, (0.78-0.80)	<.001	0.80 (0.79-0.81)	<.001
4 (2016)	0.79 (0.78-0.80)	<.001	0.80 (0.79-0.81)	<.001	0.80 (0.79-0.81)	<.001
Recruitment method						
Facebook	—	—	1	—	1	—
Other social media	—	—	1.06 (1.01-1.11)	.0114	1.04 (0.99-1.08)	.10
Referral	—	—	1.14 (1.11-1.18)	<.001	1.11 (1.07-1.15)	<.001
Traditional media	—	—	1.22 (1.18-1.26)	<.001	1.18 (1.14-1.22)	<.001
Fashion promotion	—	—	0.97 (0.94-1.00)	.0321	0.94 (0.91-0.97)	<.001
Baseline age						
18-20 years	—	—	—	—	1	—
21-23 years	—	—	—	—	1.08 (1.06-1.10)	<.001
Area of residence						
Major cities	—	—	—	—	1	—
Inner regional	—	—	—	—	1.01 (0.98-1.03)	.59
Outer regional and remote areas	—	—	—	—	0.97 (0.93-1.01)	.11
Education						
Grade 12 or more	—	—	—	—	1	—
Less than Grade 12	—	—	—	—	0.80 (0.76-0.84)	<.001
Income management						
Easy/not too bad/difficult sometimes	—	—	—	—	1	—
Impossible/difficult all the time	—	—	—	—	0.95 (0.93-0.97)	<.001
Obesity						
Not obese (BMI[c]<30)	—	—	—	—	1	—
Obese (BMI ≥ 30 or more)	—	—	—	—	0.95 (0.92-0.98)	.0035
Smoking status						
Nonsmoker/exsmoker	—	—	—	—	1	—
Current smoker	—	—	—	—	0.82 (0.79-0.85)	<.001
Self-rated health						
Excellent/very good/good	—	—	—	—	1	—
Fair/poor	—	—	—	—	0.99 (0.96-1.02)	.61
Psychological distress						
Not very high (K10 score<30)	—	—	—	—	1	—
Very high distress (K10 score ≥30)	—	—	—	—	0.96 (0.93-0.98)	.0016

[a]PR (95% CI): prevalence ratio with 95% CI.

[b]Not applicable.

[c]BMI: body mass index.

XSL•FO
RenderX

Discussion

Principal Findings

This study aimed to provide an understanding of what factors influence longitudinal study retention rates for young people recruited through an open recruitment strategy. Importantly, similar to the ALSWH 1973-78 cohort, this new cohort of young women demonstrated generational patterns in responding, with younger women more likely to be transient in their participation over time. On the basis of previous research utilizing traditional recruitment methods [7], we also hypothesized that women who had lower levels of education were obese and those who smoked would be less likely to be retained at follow-up surveys compared with other women. Although the findings support these hypotheses, we found that the method of recruitment was associated with ongoing survey response, even after controlling for sociodemographic and health-related characteristics. This is a novel finding in the new era of online recruitment.

Comparison With Prior Work

Although relatively new as a recruitment tool in health research, social media has previously been lauded for its cost-effectiveness, increased reach of participants (including those that are hard to reach and minority populations), flexibility, and ability to provide largely representative samples [2,5,24]. Here, we were able to extend our knowledge regarding the impact of recruitment method on cohort retention. The majority of women were successfully recruited into the cohort via Facebook or other online avenues, with 65% to 69% of these women returning to complete the first follow-up survey. This is consistent with cohort retention rates using traditional random sampling recruitment strategies [25]. Particularly, retention among similarly aged women recruited into the ALSWH in 1996 demonstrated a similar pattern, with 68% of participants retained at the second survey (conducted 4 years later), with participant attrition stabilizing over time. Importantly, for the newest 1989-95 ALSWH cohort of young women, a part of the study design included shorter surveys to be conducted annually in an effort to mitigate the attrition seen in the previous cohort of younger women who were administered surveys on a 3-year rolling schedule. However, this strategy did not appear to be effective in mitigating the attrition between Surveys 1 and 2, with an overall response rate of 67% just 12 months later. The main reason for nonresponse was that they could not be contacted, which was the same reason as that found for the 1973-78 cohort [25]. It is important to note, however, that although attrition was the highest between Surveys 1 and 2, this effect had plateaued by Survey 4.

Although a smaller number of participants were recruited through traditional media and referral avenues, the retention rate for this group of women remained higher than those recruited through social media. This finding has important implications for online recruitment particularly for young women. It is posited that recruitment through Facebook and other social media provided an immediate opportunity to click directly through to the first survey, potentially reducing the time to consider the ongoing commitment for participation and also involving less contact with the research team. This is supported by Frandsen et al [26] who argued that the use of social media does not lead to better informed and potentially self-screened participants compared with traditional media. Participants recruited through traditional media and referral have to consciously decide to do the survey and then access it on the web, possibly indicating more thought about completing the first survey and a better understanding that there would be additional questionnaires to follow. In addition, traditional media and referral both provide contact with the research team, either through interviews in traditional media or personal contact through referral. As a result, researchers using social media, such as Facebook, for the purposes of recruitment into longitudinal cohort studies need to consider the amount of participant information that is collected to enable accurate tracking of participants to minimize participant attrition.

Although labor intensive, previous research has identified that implementing multiple methods of retention increased rates by 70% [27]. With reduced initial contact with participants recruited through social media, more personalized methods of follow-up (eg, mail, email, short message service, and telephone) may be even more important to reduce attrition. Therefore, when recruiting a cohort for the purposes of longitudinal research, a balance between achieving representativeness, meeting cohort targets quickly, and understanding potential rates of attrition depending on the method of recruitment is required in the planning stages. Given the fact that traditional random sampling methods are no longer cost-effective to be used exclusively [15], balancing these processes and potential outcomes is particularly important.

Furthermore, the use of incentives is beginning to be highlighted as key factors in the recruitment and retention of young people [27,28]. Importantly, a systematic review found that study retention rates among the 10 studies identified increased with higher monetary values; however, whether cash incentives were more effective than gifts was not clear. This is in contrast with recent findings regarding the use of incentives in relation to cohort recruitment. In particular, an Australian study focused on contraceptive use and pregnancy intentions among similarly aged women found that only small incentives (AU $20 gift card) were required to recruit their demographically representative cohort [2]. In addition, a UK-based cohort study that used a combination of prize draw and gift voucher incentives found that the withdrawal rate was not influenced by whether an incentive was received [29]. In the ALSWH, incentives differed at each survey.

Interestingly, results for health and health behavior factors are in agreement with findings from the 4 ALSWH cohorts born between 1921 and 1926, 1946 and 1951, and 1973 and 1978 [30]. It appears to be that poor health and adverse health behaviors deter ongoing participation across generations and survey modalities. These findings reflect international literature that has demonstrated the impact of health and health behavior on participation in longitudinal health research [11,31-33]. Despite the advent of online recruitment approaches and Web-based surveying and advances in technology that support participant tracking [16], the issue of biases in retention remain.

XSL•FO
RenderX

Limitations

This study must be considered in light of its limitations. First, we examined rates of attrition among young women. Younger women, particularly in the ALSWH, have been found to be more mobile than women who are older [25]. In addition, these women were born in an age of technological advancement, particularly regarding online and social media. As a result, reasons for attrition may differ by age. It is also important to note that we had a restrictive 5-year age range for our cohort (ie, 18 to 23), which prevented us from being able to understand the nuances associated with age on recruitment strategy. This would require further examination with a broader age range. Second, there is the potential that unmeasured factors may have influenced rates of retention. Finally, although we examined associations between variables, causation cannot be implied.

In addition, the method of recruitment was determined via participant self-report with participants being only able to select 1 mode of invitation. There is the potential that participants were exposed to multiple approaches before joining the study [3] and that they selected the method that resonated the most. Future research is required to assess the number of strategies that are encountered before a commitment to study participation is made. Despite these limitations, the study is balanced by a number of strengths including being a large cohort, which is broadly representative of similarly aged Australians in the general population [3]. Such a cohort allows us to predict the subgroups of women that potentially should be targeted from the outset to remain in the study.

Conclusions

Although there is a need for more research into the factors that engage participants to commit to a cohort study and to follow through on that commitment, this study offers some insights. In particular, the mode of recruitment appears to create differing levels of commitment, perhaps reflecting the level of personal contact with the research team and time taken to consider enrolling in a study. In addition, demographic and health factors also impact on retention, suggesting that strategies to particularly re-engage those with poor health, adverse health behaviors, and from lower socioeconomic groups may be warranted. Perhaps cohort researchers would benefit from taking a leaf from qualitative methods and undertaking purposive re-engagement, along the lines of purposive sampling where a particular group is targeted for inclusion or, in this case, reinclusion. The more representative a cohort is and the less attrition it has, the more useful its results. However, a balance must be obtained between achieving representativeness, achieving rapid cohort recruitment, and mitigating the pitfalls of attrition based on recruitment method in the new era of cohort studies, where traditional recruitment methods are no longer exclusively viable options.

Acknowledgments

The research on which this study is based was conducted as part of the ALSWH by the University of Newcastle, NSW, and the University of Queensland. The authors are grateful to the women who provided the survey data. This work was supported by the Australian Government Department of Health, which has funded ALSWH since 1996. MLH is funded by an Australian Research Council Discovery Early Career Researcher Award and GDM is funded by a National Health and Medical Research Council Research Fellowship..

Conflicts of Interest

None declared.

Multimedia Appendix 1

Baseline characteristics at Survey 1 for women aged 18 to 23 years in 2012 to 2013, according to whether they responded at subsequent surveys.

[PDF File (Adobe PDF File), 116KB - jmir_v21i3e11286_app1.pdf]

Multimedia Appendix 2

Response rates of original sample (N=17,012) at follow-up surveys, according to baseline characteristics.

[PDF File (Adobe PDF File), 96KB - jmir_v21i3e11286_app2.pdf]

References

1. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
2. Harris ML, Loxton D, Wigginton B, Lucke JC. Recruiting online: lessons from a longitudinal survey of contraception and pregnancy intentions of young Australian women. Am J Epidemiol 2015 May 15;181(10):737-746. [doi: 10.1093/aje/kwv006] [Medline: 25883155]
3. Loxton D, Powers J, Anderson AE, Townsend N, Harris ML, Tuckerman R, et al. Online and offline recruitment of young women for a longitudinal health survey: findings from the Australian Longitudinal Study on Women's Health 1989-95 cohort. J Med Internet Res 2015 May;17(5):e109 [FREE Full text] [doi: 10.2196/jmir.4261] [Medline: 25940876]

XSL•FO
RenderX

4. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
5. Mishra GD, Hockey R, Powers J, Loxton D, Tooth L, Rowlands I, et al. Recruitment via the internet and social networking sites: the 1989-1995 cohort of the Australian Longitudinal Study on Women's Health. J Med Internet Res 2014 Dec;16(12):e279 [FREE Full text] [doi: 10.2196/jmir.3788] [Medline: 25514159]
6. Littman AJ, Boyko EJ, Jacobson IG, Horton J, Gackstetter GD, Smith B, Millennium Cohort Study. Assessing nonresponse bias at follow-up in a large prospective cohort of relatively young and mobile military service members. BMC Med Res Methodol 2010 Oct 21;10:99 [FREE Full text] [doi: 10.1186/1471-2288-10-99] [Medline: 20964861]
7. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health 2006 Aug;30(4):353-361. [doi: 10.1111/j.1467-842X.2006.tb00849.x] [Medline: 16956166]
8. Carter KN, Imlach-Gunasekara F, McKenzie SK, Blakely T. Differential loss of participants does not necessarily cause selection bias. Aust N Z J Public Health 2012 Jun;36(3):218-222. [doi: 10.1111/j.1753-6405.2012.00867.x] [Medline: 22672026]
9. Powers J, Loxton D. The impact of attrition in an 11-year prospective longitudinal study of younger women. Ann Epidemiol 2010 Apr;20(4):318-321. [doi: 10.1016/j.annepidem.2010.01.002] [Medline: 20149968]
10. Goldberg M, Chastang JF, Zins M, Niedhammer I, Leclerc A. Health problems were the strongest predictors of attrition during follow-up of the GAZEL cohort. J Clin Epidemiol 2006 Nov;59(11):1213-1221. [doi: 10.1016/j.jclinepi.2006.02.020] [Medline: 17027433]
11. Cunradi CB, Moore R, Killoran M, Ames G. Survey nonresponse bias among young adults: the role of alcohol, tobacco, and drugs. Subst Use Misuse 2005;40(2):171-185. [doi: 10.1081/JA-200048447] [Medline: 15770883]
12. de Graaf R, Bijl RV, Smit F, Ravelli A, Vollebergh WA. Psychiatric and sociodemographic predictors of attrition in a longitudinal study: The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Am J Epidemiol 2000 Dec 1;152(11):1039-1047. [doi: 10.1136/jech.2008.085290] [Medline: 11117613]
13. Powers J, Tavener M, Graves A, Loxton D. Loss to follow-up was used to estimate bias in a longitudinal study: a new approach. J Clin Epidemiol 2015 Aug;68(8):870-876. [doi: 10.1016/j.jclinepi.2015.01.010] [Medline: 25700941]
14. Welch KE, Leardmann CA, Jacobson IG, Speigle SJ, Smith B, Smith TC, et al. Postcards encourage participant updates. Epidemiology 2009 Mar;20(2):313-314. [doi: 10.1097/EDE.0b013e3181931558] [Medline: 19234420]
15. Harris ML, Herbert D, Loxton D, Dobson A, Wigginton B, Lucke JC. Recruiting young women for health surveys: traditional random sampling methods are not cost-effective. Aust N Z J Public Health 2014 Oct;38(5):495. [doi: 10.1111/1753-6405.12281] [Medline: 25269982]
16. Loxton D, Tooth L, Harris ML, Forder PM, Dobson A, Powers J, et al. Cohort profile: the Australian Longitudinal Study on Women's Health (ALSWH) 1989-95 cohort. Int J Epidemiol 2018 Dec 1;47(2):391-32e. [doi: 10.1093/ije/dyx133] [Medline: 29025118]
17. Measuring remoteness: Accessibility/Remoteness Index of Australia (ARIA) revised edition. Canberra: Department of Health and Aged Care; 2001. URL: https://www.health.gov.au/internet/main/publishing.nsf/Content/E2EE19FE831F26BFCA257BF0001F3DFA/$File/ocpanew14.pdf [accessed 2019-03-08] [WebCite Cache ID 76j0nkAvg]
18. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483 [FREE Full text] [Medline: 1593914]
19. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/S0033291702006074] [Medline: 12214795]
20. Australian Bureau of Statistics. 2007. 4817.0.55.001 - Information Paper: Use of the Kessler Psychological Distress Scale in ABS Health Surveys, Australia, 2007-08 URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/4817.0.55.001 [accessed 2019-03-08] [WebCite Cache ID 76j0pdage]
21. National Health and Medical Research Council. Australian Guidelines to Reduce Health Risks from Drinking Alcohol. Canberra: Commonwealth of Australia; 2009. URL: https://nhmrc.gov.au/about-us/publications/australian-guidelines-reduce-health-risks-drinking-alcohol#block-views-block-file-attachments-content-block-1
22. Obesity: Preventing and Managing the Global Epidemic. Geneva: World Health Organization; 1999. URL: https://apps.who.int/iris/handle/10665/42330
23. Brown WJ, Bauman AE, Bull FC, Burton NW. Development of evidence-based physical activity recommendations for adults (18-64 years). Canberra: Australian Government Department of Health; 2012. URL: https://www.health.gov.au/internet/main/publishing.nsf/content/F01F92328EDADA5BCA257BF0001E720D/$File/DEB-PAR-Adults-18-64years.pdf
24. Martinez O, Wu E, Shultz AZ, Capote J, López RJ, Sandfort T, et al. Still a hard-to-reach population? Using social media to recruit Latino gay couples for an HIV intervention adaptation study. J Med Internet Res 2014 Apr;16(4):e113 [FREE Full text] [doi: 10.2196/jmir.3311] [Medline: 24763130]
25. Lee C, Dobson AJ, Brown WJ, Bryson L, Byles J, Warner-Smith P, et al. Cohort profile: the Australian Longitudinal Study on Women's Health. Int J Epidemiol 2005 Oct;34(5):987-991 [FREE Full text] [doi: 10.1093/ije/dyi098] [Medline: 15894591]

XSL•FO
RenderX

26. Frandsen M, Thow M, Ferguson SG. The effectiveness of social media (Facebook) compared with more traditional advertising methods for recruiting eligible participants to health research studies: a randomized, controlled clinical trial. JMIR Res Protoc 2016 Aug 10;5(3):e161 [FREE Full text] [doi: 10.2196/resprot.5747] [Medline: 27511829]
27. Singer E, Ye C. The use and effects of incentives in surveys. Ann Am Acad Pol Soc Sci 2012 Nov 26;645(1):112-141. [doi: 10.1177/0002716212458082]
28. Herbert DL, Loxton D, Bateson D, Weisberg E, Lucke JC. Challenges for researchers investigating contraceptive use and pregnancy intentions of young women living in urban and rural areas of Australia: face-to-face discussions to increase participation in a web-based survey. J Med Internet Res 2013 Jan 21;15(1):e10 [FREE Full text] [doi: 10.2196/jmir.2266] [Medline: 23337208]
29. Toledano MB, Smith RB, Brook JP, Douglass M, Elliott P. How to establish and follow up a large prospective cohort study in the 21st century--lessons from UK COSMOS. PLoS One 2015 Jul;10(7):e0131521 [FREE Full text] [doi: 10.1371/journal.pone.0131521] [Medline: 26147611]
30. Dobson AJ, Hockey R, Brown WJ, Byles JE, Loxton DJ, McLaughlin D, et al. Cohort profile update: Australian Longitudinal Study on Women's Health. Int J Epidemiol 2015 Oct;44(5):1547a-1547f. [doi: 10.1093/ije/dyv110] [Medline: 26130741]
31. Holt C, Le D, Calvanelli J, Huang J, Clark EM, Roth DL, et al. Participant retention in a longitudinal national telephone survey of African American men and women. Ethn Dis 2015;25(2):187-192 [FREE Full text] [Medline: 26118147]
32. Gustavson K, von Soest T, Karevold E, Røysamb E. Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. BMC Public Health 2012 Oct 29;12(1):918. [doi: 10.1186/1471-2458-12-918] [Medline: 23107281]
33. Forcey DS, Walker SM, Vodstrcil LA, Fairley CK, Bilardi JE, Law M, et al. Factors associated with participation and attrition in a longitudinal study of bacterial vaginosis in Australian women who have sex with women. PLoS One 2014;9(11):e113452 [FREE Full text] [doi: 10.1371/journal.pone.0113452] [Medline: 25412421]

Abbreviations

ALSWH: Australian Longitudinal Study on Women's Health
BMI: body mass index
K10: Kessler Psychological Distress Scale
PR: prevalence ratio

Edited by G Eysenbach; submitted 13.06.18; peer-reviewed by J Blaney, A Göritz, C Blumenberg; comments to author 06.10.18; revised version received 30.11.18; accepted 03.01.19; published 25.03.19.

Please cite as:
Loxton D, Harris ML, Forder P, Powers J, Townsend N, Byles J, Mishra G
Factors Influencing Web-Based Survey Response for a Longitudinal Cohort of Young Women Born Between 1989 and 1995
J Med Internet Res 2019;21(3):e11286
URL: http://www.jmir.org/2019/3/e11286/
doi:10.2196/11286
PMID:30907739

©Deborah Loxton, Melissa L Harris, Peta Forder, Jennifer Powers, Natalie Townsend, Julie Byles, Gita Mishra. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Computer-Mediated Communication to Facilitate Synchronous Online Focus Group Discussions: Feasibility Study for Qualitative HIV Research Among Transgender Women Across the United States

Andrea L Wirtz[1], MHS, PhD; Erin E Cooney[1], MSPH; Aeysha Chaudhry[2], MSc; Sari L Reisner[2,3,4,5], ScD; American Cohort To Study HIV Acquisition Among Transgender Women[1]

[1]Department of Epidemiology, Center for Public Health and Human Rights, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[2]Division of General Pediatrics, Boston Children's Hospital, Boston, MA, United States

[3]Pediatrics, Harvard Medical School, Boston, MA, United States

[4]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States

[5]Fenway Health, The Fenway Institute, Boston, MA, United States

Corresponding Author:
Andrea L Wirtz, MHS, PhD
Department of Epidemiology
Center for Public Health and Human Rights
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St
E7644
Baltimore, MD, 21205
United States
Phone: 1 4105020800
Email: awirtz1@jhu.edu

Abstract

Background: Novel, technology-based methods are rapidly increasing in popularity across multiple facets of quantitative research. Qualitative research, however, has been slower to integrate technology into research methodology. One method, computer-mediated communication (CMC), has been utilized to a limited extent for focus group discussions.

Objective: This study aimed to assess feasibility of an online video conferencing system to further adapt CMC to facilitate synchronous focus group discussions among transgender women living in six cities in eastern and southern United States.

Methods: Between August 2017 and January 2018, focus group discussions with adult transgender women were conducted in English and Spanish by research teams based in Boston, MA, and Baltimore, MD. Participants were sampled from six cities: Baltimore, MD; Boston, MA; New York, NY; Washington, DC; Atlanta, GA; and Miami, FL. This was formative research to inform a technology-enhanced cohort study to assess HIV acquisition among transgender women. This analysis focused on the methodologic use of CMC focus groups conducted synchronously using online software that enabled video or phone discussion. Findings were based on qualitative observations of attendance and study team debriefing on topics of individual, social, technical, and logistical challenges encountered.

Results: A total of 41 transgender women from all six cities participated in seven online focus group discussions—five English and two Spanish. There was equal racial distribution of black/African American (14/41, 34%) and white (14/41, 34%) attendees, with 29% (12/41) identifying as Hispanic/Latina ethnicity. Overall, 29 of 70 (41%) eligible and scheduled transgender women failed to attend the focus group discussions. The most common reason for nonattendance was forgetting or having a scheduling conflict (16/29, 55%). A total of 14% (4/29) reported technical challenges associated with accessing the CMC focus group discussion. CMC focus group discussions were found to facilitate geographic diversity; allow participants to control anonymity and privacy (eg, use of pseudonyms and option to use video); ease scheduling by eliminating challenges related to travel to a data collection site; and offer flexibility to join via a variety of devices. Challenges encountered were related to overlapping conversations; variable audio quality in cases where Internet or cellular connection was poor; and distribution of incentives (eg,

XSL•FO
RenderX

cash versus gift cards). As with all focus group discussions, establishment of ground rules and employing both a skilled facilitator and a notetaker who could troubleshoot technology issues were critical to the success of CMC focus group discussions.

Conclusions: Synchronous CMC focus group discussions provide a secure opportunity to convene participants across geographic space with minimal time burden and without losing the standardized approach that is expected of focus group discussions. This method may provide an optimal alternative to engaging hard-to-reach participants in focus group discussions. Participants with limited technological literacy or inconsistent access to a phone and/or cellular data or service, as well as circumstances necessitating immediate cash incentives may, however, require additional support and accommodation when participating in CMC focus group discussions.

(J Med Internet Res 2019;21(3):e12569) doi:10.2196/12569

KEYWORDS

transgender; qualitative research; formative research; technology

Introduction

Novel, technology-based methods are rapidly increasing in popularity across multiple facets of quantitative HIV research. For several years now, technology has been integrated into sampling and recruitment methods via the use of online advertisements on social media and dating apps, electronic and short message service (SMS) text messaging surveys for data collection, online incentive systems, and electronic gift cards. Outside of HIV research, several large cohorts, such as the National Institutes of Health's (NIH) Precision Medicine Initiative Cohort [1] and the Black Women's Health Study [2], recruit and follow over 50,000 to 1 million participants across the United States using predominantly technology-based methods. HIV research is keeping with these trends, enrolling large samples of participants for behavioral surveys [3-5] and intervention research [6,7]. The development of SMS text messaging and mobile phone apps has become an increasingly common method to provide and support prevention and care interventions [8,9], while gaming methods have been incorporated to keep participants, particularly youth, engaged in research activities and to promote and maintain health behaviors [10,11]. The combined benefits of improving overall study efficiency, potentially decreasing study costs associated with reducing space and staff needs for data collection, and improving convenience for participants suggest that integration of technology into research practices will continue to emerge and evolve.

Qualitative research, however, has been arguably slower to integrate technology into its research methodology. One method, computer-mediated communication (CMC), has been utilized to a limited extent for focus group discussions. CMC, or *online* focus groups, has been used with populations that include, but are not limited to, cancer survivors [12,13], transplant patients [14], adolescents [15], gay and bisexual men [16], and transgender men and women [17,18]. This method provides a platform in which participants may join a virtual group to respond to a series of open-ended questions that follow a typical semistructured guide.

Previously, CMC focus group discussions have used online discussion forums in an asynchronous format that allows participants to type responses to posted questions [16]. In the asynchronous format, a group of participants may respond to the questions and to prior posts by other participants, but not all participants may be participating at the same time. The limited research of this method has suggested that CMC provides multiple benefits in terms of reducing cost and barriers associated with finding an ideal time and physical space that is convenient for participants to meet. Comparisons of in-person versus CMC focus groups have found that visual anonymity and perceived distance across the Internet stimulates group discussion and disclosure, allowing for greater sharing of ideas during CMC [19]. While typed responses provide more direct responses to questions (ie, less digression from topics) and candid sharing of sensitive information in an anonymous and confidential environment [16], discomfort with typing and low literacy levels may prohibit some participants from fully describing their viewpoints or experiences. Conversely, synchronous CMC focus groups that use audio or video options allow participants to join the discussion at the same time and respond to the question, as well as to each other's comments, in real time. Online meeting platforms, which allow participants to join securely and free of charge from a variety of online or telephone mediums, may solve some of the initial limitations of typed, asynchronous CMC qualitative research, while still providing the same benefits.

This study aimed to assess feasibility of an online conferencing system to adapt CMC methods to facilitate focus group discussions in a synchronous audio and video format among transgender women living in six cities in eastern and southern United States.

Methods

Setting

Between August 2017 and January 2018, a total of seven synchronous focus group discussions with adult transgender women were conducted by research teams in Boston, MA, and Baltimore, MD. Participants were sampled from across six cities, including Baltimore, MD; Boston, MA; New York, NY; Washington, DC; Atlanta, GA; and Miami, FL. The focus group discussions were conducted as formative research to develop and implement a technology-enhanced cohort study to assess HIV acquisition among transgender women in these cities.

Synchronous CMC focus groups were selected for use with this study population, given past observations of challenges faced by participants in terms of transportation to and from the physical location of the focus group discussions [20]. Interest

XSL•FO
RenderX

in convening focus groups across the six study cities further supported the decision to use synchronous CMC methods.

Recruitment, Eligibility Screening, and Enrollment

To reach a diverse sample, participant recruitment was conducted via multiple avenues, including word-of-mouth, advertisements shared at or by local community-based organizations and clinics, and advertisements shared via social media. Candidate participants were informed that the study was to "understand more about HIV and other health concerns among transgender women across the US." Participants were informed that the purpose of this study was to obtain their feedback, opinion, and suggestions on the methods that would be used in an upcoming study among transgender women.

Participants who were interested in joining a focus group contacted the study coordinator who screened participants by phone to assess eligibility. Participants were deemed eligible if they met the following criteria: aged 18 years or older; reported an identity along the trans feminine spectrum, which included reporting being both assigned male sex at birth and reporting transgender, gender nonconforming, or female gender identity; and no prior participation in the focus group discussion. We used maximum variation sampling [21] in an effort to ensure representation across geographic site, race and ethnicity, and age.

Focus group discussions were available in English or Spanish languages, and the screening and consent processes followed the individual's preferred language. Eligible candidates consented and were then scheduled for a focus group at a time and day that was convenient to them. The verbal consent script informed participants of the purpose of the study, the reason they were invited to participate, and the risks and benefits to participants. Participants were informed that they could use their real name or a nickname or alias during the focus group discussion to protect their privacy and could opt to use a Web camera during the focus group but were informed that this could reduce their privacy.

Study staff informally assessed technological literacy during initial screening telephone calls. Study staff were therefore able to tailor instructions and troubleshooting tips based on participant comfort and preference, with the intention of including participants with varying levels of technological literacy. Participants were provided with verbal instructions via the telephone and were emailed instructions on how to join via multiple platforms. Simple instructions that were augmented by screenshots and other images described how the participant could join the focus group by cell phone, landline, mobile app, or via Web-based connection. Email and text message reminders were also sent to participants the day of the CMC focus group to remind them of their upcoming discussion. Additional effort focused on providing support to participants who appeared to have lower levels of technological literacy or lower levels of comfort with technology. For these participants, staff members often recommended calling into the focus group discussion, rather than use a computer or tablet.

Focus Group Implementation

Groups were comprised of 5-10 participants across geographic locations. Focus group discussions lasted approximately 60-90 minutes and were led by a facilitator with the support of a notetaker from the study team. The notetaker played the traditional *notetaker* role utilized in focus group discussions, but also provided technical support during the focus group discussion. This included making sure that participants were able to securely connect to the conference system by telephone or computer, including ensuring that speakers and microphones were enabled; troubleshooting any login challenges; reminding participants that they had the option to use their name or alias, as well as to enable or disable their video depending on their privacy preferences; and documenting names of participants to ensure that the appropriate study incentive was sent following the discussion. Focus groups were implemented using the Zoom online meeting platform (Zoom Video Communications, Inc), which is a Health Insurance Portability and Accountability Act (HIPAA)-compliant system that provides end-to-end encryption and optional audio and video recording [22]. Once connected to the group, participants could communicate individually with study staff using the private message option within the system for any additional support.

The focus group facilitator specified ground rules at the start of the discussion; reminded participants of key aspects of the consent form, including consent to record; and verified group consent to record before the audio recording began. The focus group discussion was then implemented using a semistructured guide to ask open-ended questions that targeted the primary research questions of interest. Following the conclusion of the focus group discussions, participants were sent an Amazon gift card in the amount of US $40. After encountering challenges to enrolling participants in Miami, FL, we subsequently revised our methodology to offer cash incentives.

Focus group guides were developed to gain insight from transgender women community members for the development of the subsequent cohort. The guide was developed to collect information on the feasibility, acceptability, and potential barriers of various recruitment and enrollment methods, as well as study branding and marketing.

Analysis

All focus group discussions were recorded via the Zoom software platform; audio recordings were submitted to a professional transcription company to be transcribed. Transcripts were reviewed for accuracy by staff members who were present during the focus group. Focus group transcripts were complemented by detailed notes and debriefing meetings with study staff following the discussions. Trained qualitative analysts dually coded the transcripts using NVivo qualitative analysis software (QSR International) and met to discuss agreement and resolve discrepancies in coding. Spanish transcripts were coded in Spanish by Spanish-speaking staff and key quotes were translated into English.

This analysis focuses on the methodologic use of CMC focus groups conducted synchronously using online software that enables video or phone discussion for the purpose of informing

XSL•FO
RenderX

future use of this method. Findings are based on qualitative observations of attendance and study team debriefing on topics of individual, social, technical, and logistical challenges encountered by the study team or by participants.

Ethical Review

This study was approved by the Institutional Review Board at the Johns Hopkins School of Public Health. The collaborating partners, The Fenway Institute and Boston Children's Hospital, ceded review to the Johns Hopkins School of Public Health as the prime recipient of the funding. The study was also supported by a community review provided by a Community Advisory Board comprised of transgender women community members from across the six participating cities.

Results

Sample

A total of 41 transgender women, of the 70 who were eligible, participated in seven online focus group discussions—five English and two Spanish—between August 2017 and January 2018. There was equal racial distribution of black/African American (14/41, 34%) and white (14/41, 34%) attendees, with 29% (12/41) identifying as Hispanic/Latina ethnicity. Participants were from all six cities, though slightly more from Baltimore and Washington, DC, participated in the discussions (9/41, 22% and 11/41, 27%, respectively) (see Table 1).

A total of 29 of 70 (41%) eligible and scheduled transgender women failed to attend the focus group discussions. Over half (16/29, 55%) were black/African American and 34% (10/29) were mixed race or other. Nonattendees were fairly distributed across the cities, with the exception that a higher proportion were from Washington, DC (see Table 1). The most common reason for nonattendance was forgetting or having a scheduling conflict (16/29, 55%). A total of 14% (4/29) reported technical challenges associated with accessing the CMC focus group discussion.

Benefits

We observed several benefits associated with the CMC focus group discussions. The use of an online forum facilitated geographic diversity by bringing together participants from multiple cities, which increased the racial and ethnic diversity of the sample. The geographic distance between participants, coupled with the ability to use pseudonyms, and the option to use video provided additional measures for anonymity and study privacy. We observed that a few participants, however, had prior social connections despite geographic distances, particularly through online venues and social networks. The use of CMC also eased scheduling by eliminating challenges related to participants' travel to a data collection site. These benefits were augmented by the flexibility of the system to allow participants to join the discussion via telephone (ie, cell or landline), mobile phone or tablet (ie, through an app), or computer.

Table 1. Focus group participant characteristics, stratified by attendance.

Characteristic	Attendees (N=41)	Nonattendees (N=29)	Total eligible participants (N=70)
Age in years, mean (SD)	39.3 (13.6)	40.1 (12.9)	39.6 (13.2)
Race/ethnicity, n (%)			
White	14 (34)	3 (10)	17 (24)
Black	14 (34)	16 (55)	30 (43)
Asian	2 (5)	0 (0)	2 (3)
Hispanic/Latina	12 (29)	10 (34)	22 (31)
More than one race or other	11 (27)	10 (34)	21 (30)
City, n (%)			
Atlanta	7 (17)	4 (14)	11 (16)
Baltimore	9 (22)	4 (14)	13 (19)
Boston	4 (10)	4 (14)	8 (11)
New York City	5 (12)	3 (10)	8 (11)
Miami	5 (12)	4 (14)	9 (13)
Washington, DC	11 (27)	10 (34)	21 (30)
Reasons for not attending, n (%)			
Forgot or scheduling conflict	N/A[a]	16 (55)	N/A
Technological challenge	N/A	4 (14)	N/A
Unknown	N/A	9 (31)	N/A

[a]N/A: not applicable.

XSL•FO
RenderX

Challenges

Several challenges were encountered during CMC focus group discussions, but they were predominantly minor in nature. These tended to include overlapping conversations and variable audio quality. Because some participants elected to keep their video turned off, it was easy for multiple participants to speak at the same time. Most participants immediately recognized when this occurred and would often take turns allowing each other to speak. CMC focus group discussions were ultimately facilitated in approximately the same way that in-person focus group discussions are facilitated and the facilitators made a point to refer back to anyone who had not been able to provide a response. Audio quality varied in cases where Internet or cellular connection was poor or when participants situated their computer or phone in a way such that the microphone was muffled. These were rare events and tended to be resolved when it was brought to the participant's attention that they could not be heard. Ground rules were established at the beginning of the discussions to remind participants to allow each other to speak and to keep their phones or computers muted when not speaking.

Additional Considerations

Our experience with Spanish-speaking participants in Miami provided a unique challenge. Upon screening and enrollment, most candidate participants informed study staff that they did not have a phone or computer by which they could join the CMC focus groups. Ultimately, it was decided that our collaborating partners in Miami would host the focus group discussion in their office where trained research staff from Boston and Baltimore were able to connect via the video conference software. This situation proved to be the most difficult for the use of CMC focus group discussions. In prior CMC focus groups, it had been possible to view the facial expressions and body language of participants who elected to use video, and individual use of phones and computers permitted relatively clear and audible discussions. However, the setup of the camera limited the view of some participants who were sitting in the group. The setup also made it more difficult to hear participants who were sitting further away from the computer. Ultimately, the CMC focus group was able to be implemented and provided useful formative information for this population. After discussion with local staff in Miami, it was understood that many of the participants did, indeed, have cell phones, which they brought to the discussion, but simply preferred to receive cash rather than an Amazon gift card incentive.

Discussion

Principal Findings

The use of CMC focus group discussions provides an efficient and secure opportunity to convene participants across geographic space. In this case, the use of CMC focus group discussions facilitated geographic diversity, empowered participants to control their privacy during the focus group discussion, and provided flexibility in scheduling by eliminating challenges related to traveling to a data collection site. These benefits are particularly important for transgender women and other hard-to-reach populations who have multiple competing priorities and who may prefer more privacy in research participation. The benefits and challenges of CMC focus group discussions were balanced against the benefits and challenges that are associated with in-person focus group discussions (see Table 2). With consideration to unique aspects of CMC focus group discussions, we found that CMC focus group discussions are feasible to implement in overcoming challenges and barriers typically associated with in-person focus group discussions and with no apparent bearing on data or recording quality. The personal skills of facilitators and notetakers are as crucial to successful implementation of CMC focus group discussions as they are in person [23]. While the Zoom conference platform was used for these focus group discussions, any electronic platform that meets the following criteria may suffice for this purpose: allows synchronous audio and/or video communication, provides appropriate security measures, and is free of charge to participants.

In the context of this particular study, the greatest benefits of CMC focus group discussions were the flexibility in scheduling and the geographic reach offered by this method. Rarely are there situations in which busy participants can find a mutually agreeable time to participate in a focus group. CMC focus group discussions, however, reduced the burden of identifying a mutually agreeable location for participants to meet, as well as reduced time for transport to and from the location of the discussion. For transgender women, this may have the added effect of reducing travel time and costs, given expense of parking and that use of public transportation can add an additional 2-hour round trip to attend an hour-long focus group discussion in these cities. CMC focus groups also obviate security risks that are associated with traveling to unfamiliar locations and potential exposure to stigmatization and violence among this population [24]. Participants further seemed to enjoy speaking to others from different cities and these cross-city discussions provided the opportunity to assess and discuss differences across geographic locations.

XSL•FO
RenderX

Table 2. Comparison of traditional, in-person focus group discussions with CMC[a] focus group discussions and considerations for implementation of CMC focus groups.

Study component	Traditional, in-person focus group discussions	CMC focus group discussions	Considerations for CMC focus group discussions
Geographic distribution	Constrained to local setting	Can be conducted across geographic distance, pulling participants and staff across distances	Time differences
Scheduling	Requires mutually agreeable date, time, and location; additional time needed to travel to and from study site	Requires mutually agreeable date and time	Send reminder of upcoming discussion via email, call, or text Send clear instructions to access the call via phone or computer
Logistics	Requires sufficient space to facilitate a group discussion of 8-10 participants and additional staff Requires participants to travel to specified locations (with consideration for travel reimbursement, public transport, and safety)	Requires telephone or Internet access among participants and staff Requires consideration of where and how participants connect online or by phone in terms of safety and privacy	Clear instructions provided during telephone screening and sent via email with screenshots Utilize notetaker to help troubleshoot any connection challenges that participants experience
Description of focus groups	Generally understood by participants; often viewed as similar to support groups	Described as *online* focus groups, but less understood and potentially daunting for participants with lower technological literacy	Clarify that individuals can participate by logging in online (with or without video) or calling in via telephone
Data security	Focuses on security of audio recording and transcripts	Focuses on security of audio recording within the electronic system and transcripts	Continue to monitor platform for any changes to data security features
Privacy	Participants requested not to disclose their real names and use pseudonyms Participants may recognize each other from the community	Participants requested not to disclose their real names and use pseudonyms; participants given the option to turn their Web cameras on or off Participants may recognize each other from the online community	Host permissions on the back end allow staff to change the name of participants who fail to use pseudonyms
Rapport	Developed during screening, consent, and throughout the discussion	May be more limited due to impersonal conference call-style setting	The use of video by study staff allows participants to view the staff and facilitates building of rapport
Participant engagement	Facilitator observes body language to guide the conversation and attempts to engage participants who do not participate in the discussion	Body language is only visible for participants who elect to utilize their cameras, though facilitator attempts to engage participants who do not participate in the discussion	Facilitator may observe mood and emotion from participants' tone of voice to guide the conversation
Incentives	Can be provided in person in physical or electronic format	Geographic distance requires use of electronic payment or gift card or mailing the physical gift card or check	Consider local preferences for cash, mailed debit cards, or electronic payment methods
Recording and transcription	Recording may be of poor quality if participants are seated far away from each other, if participants interrupt each other, or if only one microphone is positioned in the middle of a large space; cost of transcription may increase due to poor audio quality	Software provides optional recording; cost of transcription may increase due to feedback and poor connection	Remind participants of alternative method to connect if connection quality is poor
Cost	Cost associated with space, transport, refreshments, and participant incentives	Cost associated with Zoom subscription (US $10/month) and participant incentives	Consider participant access to phone and computer

[a]CMC: computer-mediated communication.

Comparison With Prior Work

While this study did not directly compare the quality of CMC focus group discussions to in-person focus group discussions, we did not perceive that any less information was shared in CMC focus group discussions when compared to prior experiences with in-person focus group discussions with transgender women. A qualitative study by Woodyatt and colleagues directly compared in-person focus group discussions to typed CMC focus group discussions with gay and bisexual men [16] and provides further insight into the differences between CMC and in-person focus group discussions. They

XSL•FO
RenderX

found that the anonymous setting of the typed CMC discussions allowed those participants to discuss sensitive topics of intimate partner violence more candidly than those participating in person [16]. Woodyatt and colleagues, however, also reported observing a greater number of group conflicts during the typed CMC discussions compared to in-person focus group discussions; they attributed this to the complete anonymity of the typed CMC discussions [16].

Reflecting on these other variations of CMC focus group discussions, the use of synchronous CMC discussions with audio and video, as in our study with transgender women, creates a form of focus group discussion that falls on the spectrum between in-person and typed CMC focus group discussions. Providing an anonymous setting allows for open discussion of sensitive topics that are also observed with typed CMC discussions, but the audio and participant-controlled settings that allow participants to enable video may facilitate social connection and minimize group conflicts. While participants may disagree with one another's perspectives, no group conflicts in terms of insults or heated disagreements were observed in this study.

Response rates in qualitative research among transgender women are highly variable, if reported at all [25], and were no different for this study. The participation rate in this study was moderate, with 41 of 70 (59%) candidate participants attending their scheduled discussion. Follow-up with candidates who had not attended their scheduled calls suggested that scheduling conflicts were the greatest source of their absence, with fewer absences reportedly due to technology issues. The capability to connect to the discussion through multiple mediums likely alleviated the technological barriers that may have been associated with a platform requiring participants to connect only through one medium (ie, phone or Internet). These capabilities also overcome limitations associated with typed CMC focus group discussions that would restrict participation to individuals with higher levels of literacy and typing abilities, as well as those with access to computers [17,19,26]. Nonetheless, participants with limited technological literacy or inconsistent access to a phone and/or cellular data or service, as well as circumstances necessitating immediate cash incentives, may require additional support and accommodation when connecting to CMC focus group discussions.

The preference of participants in Miami to convene in person for the CMC focus group discussion highlighted the importance and underestimated role of participant incentives. Institutional administrative requirements and service costs of other electronic payment methods led us to select the use of Amazon gift cards, which do not have additional service fees and are easy to track for reporting requirements. We frequently had to explain the Amazon business model and types of retail to participants across sites, in efforts to correct perceptions that Amazon is an online bookstore and to inform participants of the various ways in which purchases could be sent to an individual. The participants from Miami preferred to be paid directly in cash, however, for more immediate use or because they did not have reliable Internet connection to make purchases or stable housing to receive packages. The format and amount of participant incentives are as important as any other decision that is made in the implementation of focus group discussions, given their potential impact on selection bias. However, the use of CMC focus group discussions restricts the available options, as physical incentives cannot be provided directly to participants. For study populations that have greater, more immediate financial needs and less of an understanding of, or lower access to, electronic payment methods, identifying the appropriate incentive requires careful consideration and input from key community informants.

Limitations

This study describes the feasibility of incorporating new technology into qualitative research for the purposes of improving access, reducing logistical constraints, and convening diverse participants across geographic space. Further research is needed, however, to fully articulate the benefits of this method in terms of efficiency, participant experiences, and data quality associated with this method. Technological literacy was not assessed among participants at enrollment. Incorporating a screening tool to assess technological literacy would provide important information in terms of potential biases associated with this method. To our knowledge, brief screening tools to measure technological literacy for public health research are not currently available, but are urgently needed for this research, as well as for the rapidly expanding use of technology in research methodologies.

Conclusions

Synchronous CMC focus group discussions using a secure online conferencing system provide a new methodologic tool for qualitative research among hard-to-reach participants and those residing across geographically distant settings. These focus group discussions draw on the privacy benefits offered by typed CMC focus groups, as well as the collegiality of group dynamics offered by in-person focus group discussions. As with all focus group discussions, establishment of ground rules and employing a skilled facilitator is critical to the success of the discussion. Unique to CMC focus group discussions, we found that it is also critical to hire a notetaker who is also technologically skilled to troubleshoot and provide support to participants as they connect to the focus group discussions. Researchers who employ synchronous CMC focus group discussions with audio and video should also consider technological literacy, acceptability of noncash incentives, feasibility of distributing incentives, and imperceptibility of nonverbal communication during implementation planning. For study participants with limited time or concerns related to meeting in person, synchronous CMC focus group discussions offer a valuable option to ensure that they continue to be engaged in qualitative research.

XSL•FO
RenderX

Acknowledgments

We would like to express our gratitude to the transgender women who participated in this study. This study would not be possible without their participation. This research is supported by an award from the National Institute of Allergy and Infectious Disease (NIAID), National Institute of Mental Health (NIMH), and National Institute of Child Health and Human Development (NICHD) (award number: UG3-AI133669). This research has been facilitated by the infrastructure and resources provided by the Johns Hopkins University Center for AIDS Research, an NIH-funded program (P30AI094189), which is supported by the following NIH cofunding and participating institutes and centers: NIAID; National Cancer Institute (NCI); NICHD; National Heart, Lung, and Blood Institute (NHLBI); National Institute on Drug Abuse (NIDA); NIMH; National Institute on Aging (NIA); Fogarty International Center (FIC); National Institute of General Medical Sciences (NIGMS); National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK); and Office of AIDS Research (OAR). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

The following are members of the collaborative author, American Cohort to Study HIV Acquisition Among Transgender Women (LITE): Sari Reisner (multiple Principal Investigator [mPI]) (Harvard University, Boston Children’s Hospital); Andrea Wirtz (mPI) (Johns Hopkins University); Keri Althoff (Johns Hopkins University); Chris Beyrer (Johns Hopkins University); Erin Cooney (Johns Hopkins University); Oliver Laeyendecker (Johns Hopkins University); Tonia Poteat (University of North Carolina); Ken Mayer (Fenway Health); Asa Radix (Callen-Lorde Community Health Center); Christopher Cannon (Whitman Walker Health); W David Hardy (Whitman Walker Health); Jason Schneider (Emory University and Grady Hospital); Sonya Haw (Emory University and Grady Hospital); Allan Rodriguez (University of Miami); and Andrew Wawrzyniak (University of Miami). ALW, SLR, and EEC developed the study concept; EEC, SLR, and ALW implemented the focus group discussions; EEC and AC implemented the qualitative analysis; ALW wrote the first draft of the manuscript; and all authors reviewed and provided scientific input to the manuscript.

Conflicts of Interest

None declared.

References

1. National Institutes of Health. About the All of Us Research Program URL: https://allofus.nih.gov/about/about-all-us-research-program [accessed 2019-02-17] [WebCite Cache ID 76GFGoYeF]
2. Boston University. Black Women's Health Study URL: https://www.bu.edu/bwhs/history/ [accessed 2019-02-17] [WebCite Cache ID 76GFLZLZt]
3. Hickson F, Tomlin K, Hargreaves J, Bonell C, Reid D, Weatherburn P. Internet-based cohort study of HIV testing over 1 year among men who have sex with men living in England and exposed to a social marketing intervention promoting testing. Sex Transm Infect 2015 Feb;91(1):24-30. [doi: 10.1136/sextrans-2014-051598] [Medline: 25270006]
4. Cheung DH, Suharlim C, Guadamuz TE, Lim SH, Koe S, Wei C. Earlier anal sexarche and co-occurring sexual risk are associated with current HIV-related risk behaviors among an online sample of men who have sex with men in Asia. AIDS Behav 2014 Dec;18(12):2423-2431 [FREE Full text] [doi: 10.1007/s10461-014-0821-0] [Medline: 24920344]
5. Matthews JE, Stephenson R, Sullivan PS. Factors associated with self-reported HBV vaccination among HIV-negative MSM participating in an online sexual health survey: A cross-sectional study. PLoS One 2012;7(2):e30609 [FREE Full text] [doi: 10.1371/journal.pone.0030609] [Medline: 22363453]
6. Bauermeister JA, Pingel ES, Jadwin-Cakmak L, Harper GW, Horvath K, Weiss G, et al. Acceptability and preliminary efficacy of a tailored online HIV/STI testing intervention for young men who have sex with men: The Get Connected! program. AIDS Behav 2015 Oct;19(10):1860-1874 [FREE Full text] [doi: 10.1007/s10461-015-1009-y] [Medline: 25638038]
7. Villegas N, Santisteban D, Cianelli R, Ferrer L, Ambrosia T, Peragallo N, et al. Pilot testing an Internet-based STI and HIV prevention intervention with Chilean women. J Nurs Scholarsh 2015 Mar;47(2):106-116 [FREE Full text] [doi: 10.1111/jnu.12114] [Medline: 25410132]
8. Catalani C, Philbrick W, Fraser H, Mechael P, Israelski DM. mHealth for HIV treatment & prevention: A systematic review of the literature. Open AIDS J 2013;7:17-41 [FREE Full text] [doi: 10.2174/1874613620130812003] [Medline: 24133558]
9. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: Systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
10. Enah C, Moneyham L, Vance DE, Childs G. Digital gaming for HIV prevention with young adolescents. J Assoc Nurses AIDS Care 2013;24(1):71-80. [doi: 10.1016/j.jana.2012.03.005] [Medline: 22871481]
11. 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]

XSL•FO
RenderX

12. Nilsson J, Jervaeus A, Lampic C, Eriksson LE, Widmark C, Armuand GM, et al. 'Will I be able to have a baby?' Results from online focus group discussions with childhood cancer survivors in Sweden. Hum Reprod 2014 Dec;29(12):2704-2711 [FREE Full text] [doi: 10.1093/humrep/deu280] [Medline: 25344069]
13. Wettergren L, Eriksson LE, Nilsson J, Jervaeus A, Lampic C. Online focus group discussion is a valid and feasible mode when investigating sensitive topics among young persons with a cancer experience. JMIR Res Protoc 2016 May 09;5(2):e86 [FREE Full text] [doi: 10.2196/resprot.5616] [Medline: 27161146]
14. Boateng B, Nelson MK, Huett A, Meaux JB, Pye S, Schmid B, et al. Online focus groups with parents and adolescents with heart transplants: Challenges and opportunities. Pediatr Nurs 2016;42(3):120-123, 154. [Medline: 27468513]
15. Doornwaard SM, den Boer F, Vanwesenbeeck I, van Nijnatten CHCJ, Ter Bogt TFM, van den Eijnden RJJM. Dutch adolescents' motives, perceptions, and reflections toward sex-related Internet use: Results of a Web-based focus-group study. J Sex Res 2017 Dec;54(8):1038-1050. [doi: 10.1080/00224499.2016.1255873] [Medline: 27982701]
16. Woodyatt CR, Finneran CA, Stephenson R. In-person versus online focus group discussions: A comparative analysis of data quality. Qual Health Res 2016 May;26(6):741-749. [doi: 10.1177/1049732316631510] [Medline: 26935719]
17. Reisner SL, Randazzo RK, White Hughto JM, Peitzmeier S, DuBois LZ, Pardee DJ, et al. Sensitive health topics with underserved patient populations: Methodological considerations for online focus group discussions. Qual Health Res 2018 Aug;28(10):1658-1673 [FREE Full text] [doi: 10.1177/1049732317705355] [Medline: 29298574]
18. Miner MH, Bockting WO, Romine RS, Raman S. Conducting Internet research with the transgender population: Reaching broad samples and collecting valid data. Soc Sci Comput Rev 2012 May 01;30(2):202-211 [FREE Full text] [doi: 10.1177/0894439311404795] [Medline: 24031157]
19. Reid DJ, Reid FJ. Online focus groups: An in-depth comparison of computer-mediated and conventional focus group discussions. Int J Mark Res 2005 Jan 30;47(2):131-162. [doi: 10.1177/147078530504700204]
20. Owen-Smith AA, Woodyatt C, Sineath RC, Hunkeler EM, Barnwell T, Graham A, et al. Perceptions of barriers to and facilitators of participation in health research among transgender people. Transgend Health 2016;1(1):187-196 [FREE Full text] [doi: 10.1089/trgh.2016.0023] [Medline: 28861532]
21. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
22. Zoom. URL: https://zoom.us/ [accessed 2019-02-17] [WebCite Cache ID 76GFwsuQO]
23. Sim J. Collecting and analysing qualitative data: Issues raised by the focus group. J Adv Nurs 1998 Aug;28(2):345-352. [Medline: 9725732]
24. Wilson E, Arayasirikul S, Johnson K. Access to HIV care and support services for African American transwomen living with HIV. Int J Transgend 2013;14(4):182-195 [FREE Full text] [doi: 10.1080/15532739.2014.890090] [Medline: 24817835]
25. Badger F, Werrett J. Room for improvement? Reporting response rates and recruitment in nursing research in the past decade. J Adv Nurs 2005 Sep;51(5):502-510. [doi: 10.1111/j.1365-2648.2005.03521.x] [Medline: 16098167]
26. Hamilton RJ, Bowers BJ. Internet recruitment and email interviews in qualitative studies. Qual Health Res 2006 Jul;16(6):821-835. [doi: 10.1177/1049732306287599] [Medline: 16760538]

Abbreviations

CMC: computer-mediated communication
FIC: Fogarty International Center
HIPAA: Health Insurance Portability and Accountability Act
mPI: multiple Principal Investigator
N/A: not applicable
NCI: National Cancer Institute
NHLBI: National Heart, Lung, and Blood Institute
NIA: National Institute on Aging
NIAID: National Institute of Allergy and Infectious Disease
NICHD: National Institute of Child Health and Human Development
NIDA: National Institute on Drug Abuse
NIDDK: National Institute of Diabetes and Digestive and Kidney Diseases
NIGMS: National Institute of General Medical Sciences
NIMH: National Institute of Mental Health
OAR: Office of AIDS Research
SMS: short message service

XSL•FO
RenderX

Edited by G Eysenbach; submitted 22.10.18; peer-reviewed by G Yang, R Kachur; comments to author 29.11.18; revised version received 20.12.18; accepted 30.12.18; published 29.03.19.

Please cite as:
Wirtz AL, Cooney EE, Chaudhry A, Reisner SL, American Cohort To Study HIV Acquisition Among Transgender Women
Computer-Mediated Communication to Facilitate Synchronous Online Focus Group Discussions: Feasibility Study for Qualitative HIV Research Among Transgender Women Across the United States
J Med Internet Res 2019;21(3):e12569
URL: http://www.jmir.org/2019/3/e12569/
doi:10.2196/12569
PMID:30924782

©Andrea L Wirtz, Erin E Cooney, Aeysha Chaudhry, Sari L Reisner, American Cohort To Study HIV Acquisition Among Transgender Women (LITE). Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Electronic Health Behaviors Among US Adults With Chronic Disease: Cross-Sectional Survey

Lillian Madrigal[1], MPH; Cam Escoffery[1], CHES, MPH, PhD

Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Lillian Madrigal, MPH
Rollins School of Public Health
Emory University
1518 Clifton Rd
Atlanta, GA, 30322
United States
Phone: 1 404 727 3956
Email: lmadrig@emory.edu

Abstract

Background: With increased access to technology and the internet, there are many opportunities for utilizing electronic health (eHealth), internet, or technology-delivered health services and information for the prevention and management of chronic diseases.

Objective: The aim of this paper was to explore (1) the differences in technology use, (2) Web-based health information seeking and use behaviors, (3) attitudes toward seeking health information on the Web, and (4) the level of eHealth literacy between adults aged 18 and 64 years with and without chronic disease.

Methods: A cross-sectional internet survey was conducted in March 2017 with 401 US adults. Participant responses were examined to understand associations between chronic disease status and eHealth behaviors such as internet health-seeking behaviors and Web-based behaviors related to health, tracking health indicators with a mobile app, patient portal use, and preferences for health information.

Results: About 1 in 3 (252/401, 37.2%) participants reported at least 1 chronic disease diagnosis. Seventy-five percent (301/401) of all participants reported having ever searched for health information on the Web. Participants with a chronic disease reported significantly higher instances of visiting and talking to a health care provider based on health information found on the Web (40.0% [48/120] vs 25.8% [46/178], χ^2_2=6.7; *P*=.01; 43.3% [52/120] vs 27.9% [50/179]; χ^2_2=7.6; *P*=.006). The uses of health information found on the Web also significantly differed between participants with and without chronic diseases in affecting a decision about how to treat an illness or condition (49.2% [59/120] vs 35.0% [63/180], χ^2_3=6.7; *P*=.04), changing the way they cope with a chronic condition or manage pain (40.8% [49/120] vs 19.4% [35/180], χ^2_2=16.3; *P*<.001), and leading them to ask a doctor new questions or get a second opinion (37.5% [45/120] vs 19.6% [35/179], χ^2_2=11.8; *P*<.001). Chronic disease participants were significantly more likely to be tracking health indicators (43.9% [65/148] vs 28.3%, [71/251] χ^2_2=10.4; *P*=.006). In addition, participants with chronic disease diagnosis reported significantly higher rates of patient portal access (55.0% [82/149] vs 42.1% [106/252], χ^2_2=6.3; *P*=.01) and use (40.9% [61/149] vs 21.0% [53/252], χ^2_2=18.2; *P*<.001). Finally, both groups reported similar perceived skills in using the internet for health information on the eHealth Literacy Scale (eHEALS). The majority of participants responded positively when asked about the usefulness of health information and importance of accessing health resources on the Web.

Conclusions: The high rates of reported information seeking and use of internet-based health technology among participants with chronic disease may reflect the uptake in eHealth to help manage chronic disease conditions. Health care providers and educators should continue to seek ways to interact and support patients in their management of chronic disease through eHealth platforms, including capitalizing on Web-based resources, patient portals, and mobile phone apps for disease education and monitoring.

(J Med Internet Res 2019;21(3):e11240) doi:10.2196/11240

XSL•FO
RenderX

KEYWORDS

information seeking behavior; chronic disease; health promotion; eHealth; mobile apps; patient portals

Introduction

Background

Electronic health (eHealth) encompasses wearable devices that sync automatically to Web-based dashboards, eHealth records accessed through mobile apps, social media networks, where patients can share experiences 24/7, Web-based medical diagnosis search engines, videoconference meetings with health care providers miles away, and much more. eHealth is defined as "health services and information delivered or enhanced through the Internet and related technologies [1]." Today, more than ever, adults have access to all sorts of internet-based technologies that can assist them with their health promotion and medical care.

It is now common practice for individuals to use the internet to seek information about their health conditions [2,3]. Furthermore, 1 area, in particular, where there is great potential for eHealth is with chronic disease management. Chronic disease self-management is the ability of patients to handle living with a chronic illness, including symptoms, treatment, physical and social consequences, and lifestyle changes [4]. Components of disease self-management encompass medical, role, symptom, and lifestyle management [5,6]. There is also a growing number of Web-based and mobile phone apps tools to help individuals manage their conditions and communicate with their health care providers. There are apps that assist patients with blood pressure monitoring, checking medical records, encouraging daily exercise, and reminding patients to take their medications [7].

Health care providers and health systems can support self-management by discussing goals and progress on patients' self-management behaviors, offering self-management education, and following up on self-management goals and behaviors [8,9]. These practices could be enhanced by technology-based support tools such as disease indicator tracking and/or Web-based offerings or apps that support patient education. In general, technologies like smartphones tend to be used by younger adults, whereas people with chronic disease tend to be older, leading to disparities in internet-based technology use between people with and without chronic disease [3,10]. However, with growing accessibility to the internet, Web-based patient portals, and smartphones, there is the potential to use eHealth to improve chronic disease management, cost, and outcomes, as well as reduce barriers to care because of issues of mobility and distance to health care providers.

Objectives

The purpose of this paper is to explore the differences between those with and without chronic disease(s) in their technology use, Web-based health information seeking and use behaviors, attitudes toward seeking health information on the Web, and level of eHealth literacy. We aim to answer the research questions (1) What is the difference in the prevalence of participation in Web-based health-related activities between adults with and without a chronic disease? and (2) Are there differences in level of eHealth literacy, engagement in eHealth behaviors, and health information resource preferences between these 2 groups?

Methods

Sample and Design

Data for this study were collected in March 2017 from a 1-time, cross-sectional internet survey of US adults drawn from Lightspeed Research (Lightspeed), an international, Web-based consumer survey company that recruits respondents by opt-in emails through multiple methods. Potential panelists registered with a unique email address and completed an in-depth demographic registration. Potential participants were eligible if they were adults (older than 18 years), had internet access, spoke English, and lived in the US participant locations verified by internet protocol addresses. Eligible respondents were invited via email to take the survey using a Health Insurance Portability and Accountability Act compliant version of SurveyMonkey. All participants were asked to read an informed consent form and give passive consent by clicking to begin the survey. No personally identifiable information was requested in the survey. Participants who completed the survey were given points by Lightspeed. This study was approved by the Emory University Institutional Review Board. At the end of the month-long recruitment period, 403 participants completed the survey, but only 401 were analyzed for this study as 2 individuals did not answer the question about chronic disease status.

Measures

The survey contained a total of 109 items. Participants were asked about ownership of different devices (ie, smartphones, tablets, and computers), internet access, frequency of internet use, as well as engagement of eHealth behaviors such as tracking of health indicators, use of mobile apps for health, health information-seeking behaviors, and other Web-based activities related to health. These questions were adapted from the Pew Health and Internet Surveys [11,12]. Participants were also asked about how they used health information (eg, had conversation with family, changed behavior, and made a decision about a condition).

Participant eHealth literacy was measured using the eHealth Literacy Scale (eHEALS), an 8-item self-reported measure of perceived eHealth literacy [13]. Participants rated their level of agreement with statements on their knowledge, comfort, and perceived skills at locating, evaluating, and applying eHealth information to health problems using a 5-point Likert-type scale (1=strongly disagree, 5=strongly agree, and range 8-40). Higher scores reflect greater perceived levels of eHealth literacy.

Next, the participants were asked about how useful the internet is in making decisions about their health and the importance of accessing resources on the internet; questions were adopted from the Health Information National Trends Survey [14]. Participants rated the extent to which they trust different sources of health information (ie, internet, television, and government

XSL•FO
RenderX

agency) on a scale of 1=not at all to 4=a lot. Participants also indicated how they preferred to receive health information from different sources (ie, person, print, and website).

Finally, demographic information on gender, race, Hispanic origin, income and education level, employment status, chronic illness diagnosis, reading level, geographic location, and rurality of their residence in the United States were assessed. Important to this study is the chronic disease diagnosis item. This was a 2-part item, first asking, "Have you been diagnosed or treated by a professional for a chronic disease?," and then prompting participants who responded "yes" to select all chronic diseases for which they have ever been treated or diagnosed. Due to various definitions of *chronic disease* [15], this study included a broad list of 14 common chronic diseases as well as an "other" option for participants to specify additional conditions. In addition, 403 participants completed the survey. Out of which, 2 participants did not respond to the chronic illness diagnosis question and were removed from the study analysis resulting in a total of 401 participants.

Data Analyses

The collected data were downloaded from SurveyMonkey and analyzed in SAS software version 9 (SAS Institute Inc). Descriptive statistics were run and used to report the chronic disease prevalence and types, demographics of the participants, levels of technology ownership and use, health monitoring, information seeking, other eHealth behaviors, and eHealth literacy and attitudes about the internet. Difference among these eHealth seeking and use variables and having chronic disease were run through independent chi-square tests and independent *t* tests. Independent chi-square tests were calculated to compare the frequency of the categorical variables between those with and without chronic disease. Independent *t* tests were used for all continuous variables.

The range, mean score, and SD were calculated for the perceived eHealth literacy level. We computed a total score for the eHEALS and calculated a Cronbach alpha to measure reliability of the total scale for the total sample and by chronic disease status groups. We examined the difference between perceived levels of eHealth literacy and chronic disease status through independent sample *t* tests. The level of eHealth behavior engagement was also compared between chronic disease status groups. eHealth behaviors were split into 2 groups: "informational" and "participatory." Informational eHealth behaviors are those that include seeking information on the Web, whereas participatory behaviors are those that include some kind of active engagement from the participant (eg, to post, share, or comment on health-related issues via social media, to join or develop Web-based health communities, or maintain healthy lifestyles) [16]. Finally, we compared eHEALS scores in both groups by eHealth behaviors using independent sample *t* tests. The level of significance for all tests was set at $P<.05$.

Results

Respondent Characteristics

About 1 in 3 (149/401, 37.2%) participants reported that they had been diagnosed with at least 1 chronic disease (see Table 1). Of the 37.2% of participants who reported a chronic disease diagnosis, the most frequently reported were high blood pressure (77/149, 51.7%), high cholesterol (60/149, 40.3%), and diabetes (48/149, 32.2%). The other chronic diseases commonly reported were arthritis, depression, asthma, anxiety, heart disease, and bronchitis. Both genders were represented at 50% and participants' ages ranged from 18 to 90 years (mean 50.7, SD 17.1). Race reflected the racial diversity of the United States with 66.8% (268/401) white, 19.2% (77/401) black, and 14.0% (56/401) other races, and 19.5% (78/399) were of Hispanic origin. About 47% (190/400) of participants had a college degree or higher, 43.8% (175/400) were employed either full-time or part-time, and 36.8% (147/399) reported household incomes over US $75,000. Many of the participants were married (202/399, 50.6%) and lived in urban (136/398, 34.2%) or suburban (199/398, 50.0%) areas, from all regions in the United States.

There were significant differences in age, employment, household income, and marital status between the groups with and without chronic disease diagnosis. The average age was significantly higher in the chronic disease group (mean 57.5, SD 15.5) compared with (mean 46.6, SD 16.6) the no chronic disease group ($t_{401}=-6.49$; $P<.001$). The chronic disease group was significantly less likely to be employed and contained a larger retired population ($\chi^2_3=14.2$; $P<.001$), with a larger proportion of the chronic disease population making an income of less than US $25,000 ($\chi^2_3=8.3$; $P=.02$). The chronic disease group had a significantly higher percentage of married individuals ($\chi^2_3=6.7$; $P=.04$) than the group with no chronic diseases.

Technology Ownership and Access

A large majority of participants owned laptops (288/401, 71.8%) and smartphones (288/401, 71.8%), followed by desktop computers, tablets, and digital versatile disc players (see Table 1). No chronic disease participants (190/252, 75.4%) reported owning a smartphone significantly more than chronic disease participants (98/149, 65.8%), ($\chi^2_2=4.3$; $P=.04$). Almost all participants reported having access to a computer (398/400, 99.3%) and using the internet several times a day (205/401, 51.1%) or almost constantly (148/401, 36.9%).

XSL•FO
RenderX

Table 1. Participant characteristics.

Characteristic	Chronic disease	No chronic disease	Total	*P* value
Chronic disease diagnosis, n (%)	149 (37.2)	252 (62.8)	401 (100)	__[a]
Type of chronic disease, n (%)				
High blood pressure	77 (51.7)	—	—	—
High cholesterol	60 (40.3)	—	—	—
Diabetes	48 (32.2)	—	—	—
Arthritis	38 (25.5)	—	—	—
Depression	27 (18.1)	—	—	—
Asthma	20 (13.4)	—	—	—
Anxiety	18 (12.1)	—	—	—
Heart disease	13 (8.7)	—	—	—
Bronchitis	12 (8.1)	—	—	—
Other chronic disease	64 (43.0)	—	—	—
Age (years), mean (SD)	57.5 (15.5)	46.6 (16.6)	50.7 (17.1)	<.001
Gender (n=399), n (%)				.49
Male	78 (52.3)	120 (47.6)	198 (49.4)	—
Female	70 (47.0)	130 (51.6)	200 (49.9)	—
Other	—	1 (0.4)	1 (0.2)	—
Race (N=401), n (%)				.87
White	102 (68.5)	166 (65.9)	268 (66.8)	—
Black	27 (18.1)	50 (19.8)	77 (19.2)	—
Other	20 (13.4)	36 (14.3)	56 (14.0)	—
Hispanic origin (n=399)	30 (20.1)	48 (19.2)	78 (19.5)	.82
Level of school (n=400), n (%)				.43
High school graduate, GED[b], or less	24 (16.1)	49 (19.5)	73 (18.3)	—
Some college	58 (38.9)	79 (31.5)	137 (34.3)	—
College	42 (28.2)	82 (32.7)	124 (31.0)	—
Graduate	25 (16.8)	41 (16.3)	66 (16.5)	—
Employment (n=400), n (%)				<.001
Employed, full-time or part-time	52 (35.1)	123 (48.8)	175 (43.8)	—
Retired	54 (36.5)	50 (19.8)	104 (26.0)	—
Other	42 (28.4)	79 (31.3)	121 (30.3)	—
Household income is US $ (n=399), n (%)				.02
Less than 24,999	43 (29.3)	45 (17.9)	88 (22.1)	—
25,000-74,999	50 (34.0)	114 (45.2)	164 (41.1)	—
75,000 or more	54 (36.7)	93 (36.9)	147 (36.8)	—
Marital status (n=399), n (%)				.04
Married	82 (55.0)	120 (48.0)	202 (50.6)	—
Divorced or separated or widowed	28 (18.8)	34 (13.6)	62 (15.5)	—
Single	39 (26.2)	96 (38.4)	135 (33.8)	—
Urban-rural location (n=398), n (%)				.08
Urban	52 (35.6)	84 (33.3)	136 (34.2)	—

XSL•FO
RenderX

Characteristic	Chronic disease	No chronic disease	Total	*P* value
Suburban	64 (43.8)	135 (53.6)	199 (50.0)	—
Rural	30 (20.5)	33 (13.1)	63 (15.8)	—
US geography (n=398), n (%)				.56
Northeast	29 (19.6)	64 (25.6)	93 (23.4)	—
South	36 (24.3)	62 (24.8)	98 (24.6)	—
Midwest	31 (20.9)	39 (15.6)	70 (17.6)	—
Southwest	16 (10.8)	27 (10.8)	43 (10.8)	—
West	36 (24.3)	58 (23.2)	94 (23.6)	—
Technology ownership (N=401), n (%)				—
Laptop	105 (70.5)	183 (72.6)	288 (71.8)	.64
Smartphone	98 (65.8)	190 (75.4)	288 (71.8)	.04
Desktop computer	96 (64.4)	138 (54.8)	234 (58.4)	.06
Tablet (Kindle, iPad)	79 (53.0)	148 (58.7)	227 (56.6)	.26
Digital versatile disk or Blu ray player	81 (54.4)	139 (55.2)	220 (54.9)	.88
Game console	46 (30.9)	98 (38.9)	144 (35.9)	.11
Cell phone	47 (31.5)	57 (22.6)	104 (25.9)	.05
Internet access (n=400), n (%)	149 (100.0)	249 (98.8)	398 (99.3)	.18
Frequency of internet use (N=401), n (%)				.27
Almost constantly	48 (32.2)	100 (39.7)	148 (36.9)	—
Several times a day	80 (53.7)	125 (49.6)	205 (51.1)	—
About once a day to less often	21 (14.1)	27 (10.7)	48 (12.0)	—

[a]Not applicable.

[b]GED: General Educational Development.

Health Information Seeking and Technology Use for Health

Many of the participants (301/401, 75.1%) reported having ever searched for health information on the internet and to have searched for health information on the Web in the past month (172/401, 42.9%; mean 3.1, SD 8.8). The top 4 search topics were diet/nutrition, exercises, medicines, and quick remedies. Participants with chronic diseases were significantly more likely to have searched about medicines than the no chronic disease participants (25.5% [38/149] vs 13.5% [34/252], t_{401}=3.06; *P*=.02).

About 34.1% of participants reported tracking any health indicator regularly (see Table 2). Chronic disease participants were significantly more likely to be tracking health indicators (43.9% [65/148] vs 28.3% [71/251], χ^2_2=10.4; *P*=.006). Participants reported tracking health indicators by keeping track in their head (82/401, 20.4%), on paper (62/401, 15.5%), with a phone app (56/401, 14.0%), medical device (48/401, 12.0%), website (38/401, 9.5%), wearable device (36/401, 9.0%), and/or computer program (33/401, 8.2%). Of the 24.2% (97/401) of participants who reported having health focused mobile phone apps, the top apps were related to exercise (73/401, 18.2%), diet (35/401, 8.7%), and weight (25/401, 6.2%). There were no meaningful differences between chronic disease and no chronic disease groups.

Other common health activities on the internet included reading about someone else's health experience (40.3% [60/149] chronic disease group vs 33.3% [84/252] no chronic disease group), watching a video about health (34.9% [52/149] vs 28.1% [70/252]), and surfing the Web to find others who have similar health conditions (27.5% [41/149] vs 24.2% [61/252]). Chronic disease participants reported significantly higher activity related to signing up for health email updates (29.5% [44/149] vs 19.8% [50/252]; χ^2_2=5.3; *P*=.02) and downloading health insurance forms or applying for health insurance (25.2% [37/147] vs 14.7% [37/252], χ^2_2=6.8; *P*=.009). Participants used the information they found on the Web in a variety of ways. Participants with chronic disease reported significantly higher instances of visiting a health care provider based on health information found on the Web (40.0% [48/120] vs 25.8% [46/178], χ^2_2=6.7; *P*=.01) and talking with a provider about health information found on the Web (43.3% [52/120] vs 27.9% [50/179], χ^2_2=7.6; *P*=.006) than those with no chronic diseases. The uses of health information found on the Web also significantly differed between participants with and without chronic diseases in affecting a decision about how to treat an illness or condition (49.2% [59/120] vs 35.0% [63/180], χ^2_3=6.7;

XSL•FO
RenderX

P=.04), changing the way they cope with a chronic condition or manage pain (40.8% [49/120] vs 19.4% [35/180], χ^2_2=16.3; *P*<.001), and leading them to ask a doctor new questions or get a second opinion (37.5% [45/120] vs 19.6% [35/179], χ^2_2=11.8; *P*<.001).

Finally, related to health system portal use, 46.9% (188/401) of participants reported having access to a patient portal or app with 28.4% (114/401) of participants who have ever used a patient portal in the last 12 months. Patients with a chronic disease diagnosis reported significantly higher rates of patient portal access (55.0% [82/149] vs 42.1% [106/252], χ^2_2=6.3; *P*=.01) and patient portal use in the last 12 months (40.9% [61/149] vs 21.0% [53/252], χ^2_2=18.2; *P*<.001). Of those with access to a patient portal, frequent uses included viewing test or lab results (89/188, 47.3%), emailing the doctor or doctor's office (73/188, 38.8%), and setting up an appointment on the Web (59/188, 31.4%). There were no meaningful differences in patient portal activities between those with or without a chronic disease diagnosis.

Electronic Health Literacy

Generally, both groups, those with and without chronic disease, reported similar perceived skills in using the internet for health information. The total eHEALS score of the chronic disease group was significantly higher by 1.23 points (t_{393}=−1.99; *P*=.03, see Table 3) than that of the group with no chronic diseases. The average total score for all participants was 29.89 (SD 5.95); it was 30.66 (SD 6.10) for those with chronic disease and 29.43 (SD 5.81) for those without chronic disease. The chronic disease participants generally reported slightly higher confidence in 7 out of the 8 eHEALS items. The scale had high internal consistency with a Cronbach alpha of .936.

Overall, both groups reported neutral to positive feelings regarding the usefulness of the internet in making decisions about health (mean 3.65, SD 0.98) as well as the importance of accessing health resources on the internet (mean 3.69, SD 1.07). Participants with chronic disease significantly reported slightly higher levels of confidence that they could get advice or information about health or medical topics if needed (mean 2.26 vs mean 2.48, t_{399}=2.23; *P*=.03).

Electronic Health Behavior Engagement and Electronic Health Literacy Rate Scale Scores

There were no significant differences in participant engagement in eHealth behaviors (Table 4). eHealth behaviors were split into 2 groups: "informational" and "participatory." Informational eHealth behaviors are those that include seeking information on the Web, whereas participatory behaviors are those that include some kind of active engagement from the participant (eg, to post, share, or comment on health-related issues via social media, to join or develop Web-based health communities, or maintain healthy lifestyles) [16]. The list of eHealth behaviors by type is included in Table 5. On average, the chronic disease group reported engaging in 4.57 different eHealth behaviors (SD 3.32), whereas the no chronic disease group reported engaging in 3.97 eHealth behaviors (SD 3.22). Both groups reported higher engagement in informational eHealth behaviors compared with participatory ones. This echoes the results described in Table 2.

XSL•FO
RenderX

Table 2. Participant health information–seeking and technology use for health behaviors.

Item	Chronic disease, n (%)	No chronic, disease n (%)	Total, N (%)	*P* value
Ever looked on the Web for health info (N=401)	120 (80.5)	181 (71.8)	301 (75.1)	.05
Looked on the Web for health info in the past month (N=401)	62 (41.6)	110 (43.7)	172 (42.9)	.69
Average number of times in past month looked for health information (n=297)	4.06 (SD 12.94)	2.43 (SD 4.05)	3.08 (SD 8.79)	.12
Do you track a health indicator? (n=399)	65 (43.9)	71 (28.3)	136 (34.1)	.006
Tracking method (N=401)				
In their head	34 (22.8)	48 (19.0)	82 (20.4)	.37
Paper	24 (16.1)	38 (15.1)	62 (15.5)	.78
App on phone	19 (12.8)	37 (14.7)	56 (14.0)	.59
Medical device	33 (22.1)	15 (6.0)	48 (12.0)	<.001
Website or on the Web	14 (9.4)	24 (9.5)	38 (9.5)	.97
Wearable device	6 (4.0)	30 (11.9)	36 (9.0)	.008
Computer program	17 (11.4)	16 (6.3)	33 (8.2)	.07
Do you use any health apps? (N=401)	36 (24.2)	61 (24.2)	97 (24.2)	>.99
Types of health apps used (N=401)				
Exercise	25 (16.8)	48 (19.0)	73 (18.2)	.57
Diet, food, calorie counter	13 (8.7)	22 (8.7)	35 (8.7)	>.99
Weight	7 (4.7)	18 (7.1)	25 (6.2)	.33
Blood pressure	10 (6.7)	9 (3.6)	19 (4.7)	.15
WedMD or health organization	11 (7.4)	12 (4.8)	23 (5.7)	.28
Menstrual cycle	8 (5.4)	7 (2.8)	15 (3.7)	.19
Sleep	1 (0.7)	14 (5.6)	15 (3.7)	.01
Blood sugar or diabetes	8 (5.4)	5 (2.0)	13 (3.2)	.06
Medication management	7 (4.7)	2 (0.8)	9 (2.2)	.02
Mood or feelings	5 (3.4)	2 (0.8)	7 (1.7)	.06
Health topics searched in the past month (N=401)				
Diet or nutrition	39 (26.2)	56 (22.2)	95 (23.7)	.37
Exercise	19 (12.8)	63 (25.0)	82 (20.4)	.003
Medicines	38 (25.5)	34 (13.5)	72 (18.0)	.002
Quick remedy	25 (16.8)	26 (10.3)	51 (12.7)	.06
Other health activities on the internet (N=401)				
Read someone else's commentary or experience about health issues	60 (40.3)	84 (33.3)	144 (35.9)	.16
Watched a video about health	52 (34.9)	70 (28.1)	122 (30.7)	.16
Gone on the Web to find others who might have health concerns similar to you	41 (27.5)	61 (24.2)	102 (25.4)	.46
Signed up to receive email updates	44 (29.5)	50 (19.8)	94 (23.6)	.02
Download forms or applied for health insurance on the Web	37 (25.2)	37 (14.7)	74 (18.5)	.009
Uses of health information (n=299)				
Have a conversation with friend or family member	46 (38.3)	77 (43.0)	123 (41.1)	.42
Changed behavior	23 (19.2)	41 (23.0)	64 (21.5)	.43
Made a decision about condition	50 (41.7)	68 (38.0)	118 (39.5)	.52

XSL•FO
RenderX

Item	Chronic disease, n (%)	No chronic, disease n (%)	Total, N (%)	*P* value
Visited a doctor or provider	48 (40.0)	46 (25.8)	94 (31.5)	.01
Talked with a doctor or provider	52 (43.3)	50 (27.9)	102 (34.1)	.006
Did the information found on the Web... (n=301)				
Affect a decision about how to treat an illness or condition	59 (49.2)	63 (35.0)	122 (40.5)	.04
Change your overall approach to maintaining your health or someone else's health	40 (33.3)	55 (30.6)	95 (31.6)	.59
Change the way you cope with a chronic condition or manage pain	49 (40.8)	35 (19.4)	84 (27.9)	<.001
Affect a decision about whether to see a doctor	36 (30.0)	46 (25.6)	82 (27.2)	.40
Lead you to ask a doctor new questions, or get a second opinion	45 (37.5)	35 (19.6)	80 (26.8)	<.001
Doctor has recommended a particular health or medical website to you? (n=399)	23 (15.5)	13 (5.2)	36 (9.0)	<.001
Health system portal use (N=401)				
Number of people with patient portal access (portal or app)	82 (55.0)	106 (42.1)	188 (46.9)	.01
People who have used the portal in the last 12 months	61 (40.9)	53 (21.0)	114 (28.4)	<.001
Uses of the patient portal for those with access (n=188)				
Email your doctor or office	28 (34.1)	45 (42.5)	73 (38.8)	.24
Set an appointment on the Web	17 (20.7)	42 (39.6)	59 (31.4)	.01
See tests or laboratory results	36 (43.9)	53 (50.0)	89 (47.3)	.41

XSL•FO
RenderX

Table 3. Participant health literacy and perceptions of internet health resources. Cronbach alpha: chronic disease=.940; no chronic disease=.933; total=.936.

eHEALS[a] item[b]	Chronic disease, mean (SD)	No chronic disease, mean (SD)	Total, mean (SD)	*P* value
I know how to find helpful health resources on the internet (n=399)	4.08 (0.78)	3.74 (0.89)	3.87 (0.86)	<.001
I know how to use the internet to answer my health questions (N=401)	3.94 (0.84)	3.81 (0.82)	3.86 (0.83)	.13
I know what health resources are available on the internet (n=400)	3.64 (0.97)	3.67 (0.84)	3.66 (0.89)	.73
I know where to find helpful health resources on the internet (n=399)	3.93 (0.90)	3.78 (0.86)	3.83 (0.88)	.10
I know how to use health information I find on the internet to help me (N=401)	3.92 (0.92)	3.80 (0.88)	3.85 (0.90)	.20
I have the skills I need to evaluate the health resources I find on the internet (n=400)	3.82 (0.94)	3.69 (0.89)	3.74 (0.91)	.17
I can tell high quality from low quality health resources on the internet (n=399)	3.69 (0.95)	3.48 (0.91)	3.56 (0.93)	.03
I feel confident in using information from internet to make health decision (n=399)	3.61 (0.94)	3.52 (0.94)	3.55 (0.94)	.33
Total eHEALS score (n=393)	30.66 (6.10)	29.43 (5.81)	29.89 (5.95)	.05
Usefulness of internet to making decisions about health[c] (n=400)	3.68 (1.01)	3.62 (0.95)	3.65 (0.98)	.53
Importance of accessing health resources on the internet[d] (n=399)	3.72 (1.17)	3.67 (1.01)	3.69 (1.07)	.68
Overall, how confident are you that you could get advice or information about health or medical topics if you needed it?[e] (N=401)	2.26 (1.00)	2.48 (0.99)	2.40 (1.00)	.03

[a]eHEALS: electronic health literacy scale.

[b]1=strongly disagree to 5=strongly agree.

[c]1=not useful at all to 5=very useful.

[d]1=not important at all to 4=very important.

[e]1=not confident at all to 5=completely confident.

Table 4. eHealth[a] behavior engagement.

Type of eHealth behavior	eHealth behaviors participants reported engaging in		*P* value
	Chronic disease (n=149), mean (SD)	No chronic disease (n=252), mean (SD)	
eHealth behaviors	4.57 (3.32)	3.97 (3.22)	.07
Informational eHealth behaviors	3.0 (2.06)	2.7 (2.07)	.11
Participatory eHealth behaviors	1.6 (1.55)	1.3 (1.49)	.12

[a]eHealth: electronic health.

Higher health literacy eHEALS scores were associated with engaging in different eHealth behaviors (Table 5). In Table 5, for both the chronic disease and no chronic disease groups, average eHEALS scores were compared among participants based on their eHealth engagement behavior. For all informational eHealth behaviors for both chronic disease status groups, there were significantly higher eHEALS scores in the group that reported engaging in the behavior. With the participatory eHealth behaviors, the average eHEALS scores were also higher in those that engaged in the behavior versus those that did not (eg, tracking health indicators using a mobile app, patient portal use in last 12 months). However, the statistical significance varied. One reason for this could be the overall lower reported engagement in the participatory eHealth behaviors.

XSL•FO
RenderX

Table 5. Comparison of eHealth[a] literacy scale scores by eHealth behavior engagement.

eHealth behavior	Chronic disease group					No chronic disease group				
	Those engaging in eHealth behavior		Those not engaging in eHealth behavior		*P* value	Those engaging in eHealth behavior		Those not engaging in eHealth behavior		*P* value
	Average eHEALS[b] score	n (%)	Average eHEALS score	n (%)		Average eHEALS score	n (%)	Average eHEALS score	n (%)	
Informational eHealth behaviors										
Web-based health information seeking (ever)	31.29	119 (80.4)	28.03	29 (19.6)	.009	30.70	178 (72.7)	26.04	67 (27.3)	<.001
Web-based health information seeking (in the last month)	31.97	87 (58.8)	28.79	61 (41.2)	.002	31.26	139 (56.7)	27.03	106 (43.3)	<.001
Web-based information seeking on the phone (ever)	32.82	45 (30.6)	29.76	102 (69.4)	.005	31.31	87 (35.5)	28.39	158 (64.5)	<.001
Signed up for health email updates or alerts	32.25	44 (29.9)	30.01	103 (70.1)	.04	31.71	49 (20.0)	28.86	196 (80.0)	.002
Went on the Web to read about other's experiences	31.93	60 (40.5)	29.78	88 (59.5)	.04	31.33	82 (33.5)	28.47	163 (66.5)	<.001
Went on the Web to watch health-related videos	32.54	52 (35.1)	29.64	96 (64.9)	.005	31.84	69 (28.4)	28.41	174 (71.6)	<.001
Went on the Web to find others with similar health issues	32.50	40 (27.0)	29.97	108 (73.0)	.03	31.62	60 (24.5)	28.72	185 (75.5)	<.001
Participatory eHealth behaviors										
Download forms or applied for health insurance on the Web	32.49	37 (25.2)	30.10	110 (74.8)	.04	30.92	36 (14.7)	29.17	209 (85.3)	.10
Web-based tracking of weight or diet indicators	33.41	22 (14.9)	30.17	126 (85.1)	.02	30.69	48 (19.6)	29.12	197 (80.4)	.09
Web-based tracking of other health indicators	34.17	30 (20.4)	29.68	117 (79.6)	<.001	31.55	33 (13.5)	29.12	211 (86.5)	.03
Posted on social networking	32.39	28 (18.9)	30.25	120 (81.1)	.09	31.27	41 (16.7)	29.06	204 (83.3)	.03
Posted on a Web-based discussion group	30.94	18 (12.2)	30.56	129 (87.8)	.80	31.45	20 (8.2)	29.25	225 (91.8)	.11
Tracking health indicators using a mobile app	32.29	35 (23.6)	30.15	113 (76.4)	.07	31.15	60 (24.5)	28.87	185 (75.5)	.008
Patient portal use in last 12 months	31.48	60 (40.5)	30.09	88 (59.5)	.17	31.75	83 (33.9)	28.24	162 (66.1)	<.001

[a]eHealth: electronic health.

[b]eHEALS: eHealth Literacy Scale.

XSL•FO
RenderX

Table 6. Distribution of participants' preferences for methods of receipt of electronic health information: "In which of these ways would you like to get information and advice about how to manage health conditions and make changes in health behaviors (diet, exercise)?"

Communication method	Chronic disease, n	No chronic disease, n	Total, N (%)
Get information from internet websites	58	87	145 (36.2)
Print materials (eg, brochures, tip sheets)	65	54	119 (29.7)
Health newsletters or information by email	49	57	106 (26.4)
In-person counseling with a patient educator	43	51	94 (23.4)
Communications using doctor's secure email	33	37	70 (17.5)
Get information from your doctor's home page or patient portal	30	32	62 (15.5)
Telephone sessions with a health coach or educator or provider	33	23	56 (14.0)
Health newsletters or information by mail	29	23	52 (13.0)
Watch Web-based videos on doctor's website, and YouTube	22	29	51 (12.7)
Use a health app on your tablet or smartphone	21	29	50 (12.5)

Channels for Health Information

Finally, the top 10 sources in which participants preferred to receive health information included via the internet as the top choice (58/401, 36.2%), followed by print materials (65/401, 29.7%), health newsletters or information via email (49/401, 26.4%), in-person counseling with a patient educator (43/401, 23.4%), and direct communication with the doctor via email (33/401, 17.5%; Table 6). There were no significant differences in methods for receiving health information between the 2 groups.

Discussion

Principal Findings

Overall, this study found that adults with and without a chronic disease diagnosis are going to the internet to seek health information. Participants with chronic disease appear to be slightly more activated in their eHealth behaviors in searching for health information, tracking health indicators, and using a patient portal. A number of other studies have also reported increased eHealth behaviors for people with chronic disease; these eHealth behaviors include looking for information regarding their conditions [8,17,18], using Web-based or computer tools to help manage their conditions [19-21], and using portal platforms to increase engagement in their own personal health information as well as increase communication with their health care providers [22,23].

The differences in the study population characteristics between the 2 groups reflect the national prevalence of chronic diseases in older populations [24]. The chronic diseases commonly reported by those who have been diagnosed with at least 1 chronic disease in our sample also reflect the most prevalent chronic diseases in the United States [24]. Study participants with chronic disease tended to be older, retired, at lower income levels, and married. These characteristics could influence their eHealth behaviors. For example, older adults do not typically use as much technology compared with younger populations [25]. This study also found some examples of this technology gap (eg, less smartphone ownership); however, it was not great and did not result in huge differences in eHealth behaviors. Predictors of eHealth behaviors within this population are discussed in a previously published paper [26]. The authors found that, generally, younger age, being female, and higher eHEALS scores were significant predictors of looking on the Web for health information in the past month and having a health-related mobile app.

This study's results indicate high access to the internet and eHealth resources. In 2010 and 2013, national survey data from the Pew Research Center found that people with chronic diseases were less likely to have internet access [8,18]. In this study, there were no differences between groups in access or frequency of internet use, probably because of the rise in ubiquitous internet access and use in recent years. In a 2018 national report by Pew, 89% of US adults reported regular access to the internet [27]. In addition, across all participants, the top most preferred way (36.2%) to receive health information was via the internet (see Table 6). Consistent with the other literature, eHealth, that is health services and information delivered or enhanced through the internet and related technologies, is currently the norm and most common way through which individuals seek health information in the United States [7,18,21,28]. Similar to the previous Pew studies, we also found that those with chronic conditions were more likely than other adults to engage in certain eHealth-related behaviors, such as tracking a health indicator or following up with a medical professional based on information they found on the Web about their condition [7,8]. However, there were no significant differences in the average number of eHealth behaviors in which participants reported engaging between those with and without a chronic disease (see Table 4). These findings demonstrate a consistency in technology-based behaviors for health among adults.

The use of eHealth resources requires eHealth literacy, "the ability to read, use computers, search for information, understand health information, and put it into context [13]." This study found moderately high rates of health literacy among all participants using the eHEALS (mean 29.89 out of a possible 40 points). The eHEALS has been tested for validity and reliability in many populations, including those with chronic disease [13,17,29]. In addition, we found that higher eHEALS average scores were associated with engagement in eHealth

XSL•FO
RenderX

behaviors, particularly with information-seeking behaviors. Although eHealth behaviors have become commonplace, future interventions may be necessary to improve literacy and capacity to use eHealth for the community to impact its disease management and/or well-being. Stellesfon et al explored levels of health literacy among people with chronic disease using the eHEALS and found that although participants reported moderate levels of eHealth literacy, they were not as confident in the ability to distinguish the quality of the health information found on the Web. The participants in this study within both groups also reported lower confidence in high- versus low-quality Web-based health information.

Beyond seeking health information, the rise of Web-based technology has allowed for greater management of personal health information, especially through patient portals. Coughlin et al define patient portals as "Web-based, patient-centered health care information systems linked to a patient's electronic medical record." Patient portals have many functions including, but not limited to, the ability to communicate with health care providers, checking health records and lab results, requesting prescription refills, viewing educational materials, and scheduling appointments [30]. Since the 2009 US Health Information Technology for Economic and Clinical Health Act, health care providers and systems have dramatically increased the availability of electronic health records and access to patient portals [31]. In 2014, the Office of the National Coordinator for Health Information Technology reported that about 40% of US adults have access to their electronic health records via Web-based patient portals and around 55% of those adults have actually gone on the Web to access them [32]. In this study, approximately half (188/401, 46.9%) of the participants reported having access to a patient portal, and of those with access, 60.1% (114/188) reported using the patient portal in the last 12 months.

Since 2009, patient portals have been studied extensively to understand current access and usage trends, patient and provider attitudes toward patient portals, and the benefits of patient portals. However, as uptake increases, more research is still needed in this area, especially as it pertains to the management of chronic diseases. In the Coughlin et al review, several studies found benefits to using health care system patient portals to assist patients manage their chronic disease(s) and improve patient outcomes. Other studies focused on patients with chronic disease have reported that portals have the ability to improve access and communication with health care providers [22,23]. Not only do the patients have a direct way to message their health care provider, but having access to their own health data and knowing that their provider has their complete health history allow for patients to discuss concerns in more detail during their visits. This study's participants with chronic disease had higher access and use of patient portals than the no chronic disease group. Participants used the patient portal mostly for viewing test or lab results (89/188, 47.3%), emailing the doctor or doctor's office (73/188, 38.8%), and setting up an appointment on the Web (59/188, 31.4%). These results support previous findings that patient portals are a good potential avenue for intervention with this population and are enhancing the way patients interact with their providers as well as understand their own health data. This portal technology has the potential to increase the efficiency and quality of health care; however, much more research is needed to rigorously test the benefits and increase both provider and patient uptake of these systems [23]. Continued research is necessary to better understand the purpose of portal use, frequency of use, and the relationship between use and better health outcomes.

Another key eHealth behavior regarding personal health information has been the tracking and management of health indicators. With the rise of wearable and smartphone technologies, there are many more ways for people to collect and record data on their own health behaviors and outcomes. A 2012 survey conducted by the Pew Research Center's Internet and American Life Project on mobile health behaviors found that 69% of US adults kept track of at least 1 health indicator (ie, weight, diet, exercise routine, or symptom). They also found that people living with 1 or more chronic conditions were no more likely than other US adults to track their weight, diet, or exercise routine, but are more likely to track disease-related health indicators (eg, blood pressure). Their study found that 49% of trackers say they keep track of progress "in their heads," 34% say they track the data on paper, like in a notebook or journal, and 21% say they use some form of technology to track their health data [33]. In this study, only 34.1% (136/399) of participants reported tracking any health indicator regularly, which is lower. Chronic disease participants were significantly more likely to be tracking health indicators and participants similarly reported the same top 3 tracking methods (in their head, on paper, with some form of technology). In 2012, Pew Research Center also reported that 19% of smartphone owners have at least 1 health app on their phone [33]. Recently, with the increased prevalence of smartphone technology, we expected to see much higher rates of use of smartphone apps to track health indicators, but only 24.2% (97/401) of the study participants reported having any health-focused smartphone apps. These results may be partially because our broader sample of different age groups that may differ from other research in this area. This is a critical area of chronic disease management for patients to track important health indicators (ie, weight, A1c, blood pressure, and steps for physical activity) and share relevant data to their providers as necessary for their health management. Health apps and passive technology such as wearable devices can assist patients with these tasks [34].

Recent studies that have examined mobile app use within chronic disease populations have found that although a wide offering of chronic disease management apps is on the market, they are lacking in their functionality, clinical utility, and usability [35]. In 2018, Escoffery and colleagues came to similar conclusions after reviewing apps for epilepsy self-management. They found that although a number of apps existed, many were very limited, and there was a lack of theory and evidence-grounded health behavior techniques applied to these apps to assist with chronic disease management [36]. Other studies have found variability in app design and education and indicated the need for detailed education and self-management features [37]. In addition to this lack of quality, there is a lack of use by those who are at highest risk or poorest health status and behaviors [38]. Much further research is needed in the development, testing, and accessibility of quality mobile health

XSL•FO
RenderX

apps for people with chronic diseases focused on education, monitoring of health status, social support, and promotion of self-management.

Limitations

Limitations of this study include that the sample may not be representative of all US adults. To reduce this selection bias, we attempted to recruit a cross-section of the population across gender, racial, ethnic, and educational groups. In addition, as the survey was conducted via a Web-based platform, there is a bias toward individuals with access to and familiar with the internet. However, our finding is consistent with another national study with approximately 89% of US adults with regular access to the internet and their high frequency of internet use (26% reporting that they are almost constantly on the Web) [27]. Reporting bias because of self-reported data is another limitation, especially for our key variable of interest: chronic disease status. Our survey took a broader approach to defining chronic disease and simply asked participants to self-report if they have ever been diagnosed or treated for a chronic disease. We did not medically verify disease diagnosis. In addition, the sample size is smaller than similar survey studies [10,38,39] and comparisons between chronic disease and no chronic disease groups are limited by the small numbers. Finally, the eHEALS only includes skills related to certain eHealth behaviors and does not yet include eHealth behaviors relate to mobile apps and social media.

Conclusions

This study presents current technology ownership, eHealth literacy, information seeking, eHealth behaviors, and impacts of eHealth information seeking among a sample of US adults with chronic diseases. The observed high information seeking and use of internet-based health technology among participants with chronic disease may reflect uptake in eHealth to help manage chronic disease conditions. Health care providers and educators should continue to seek ways to interact and support patients in their management of chronic disease through eHealth platforms, including Web-based resources, patient portals, and smartphone apps for disease monitoring.

Acknowledgments

This study was conducted by Emory Rollins School of Public Health. The survey was administered by Lightspeed Research (lightspeedresearch.com) for US $4000.

Conflicts of Interest

None declared.

References

1. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
2. Fox S, Duggan M. Pew Research Center Internet & American Life Project. 2013 Jan 15. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2018-05-28] [WebCite Cache ID 6zkMcnfJN]
3. Pew Research Center Science and Society. 2013 Nov 26. The Diagnosis Difference URL: http://www.pewinternet.org/2013/11/26/the-diagnosis-difference [accessed 2018-05-28] [WebCite Cache ID 6zkMxXiGa]
4. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [Medline: 12401421]
5. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. J Am Med Assoc 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
6. Institute of Medicine (US) Committee on the Public Health Dimensions of the Epilepsies. In: England MJ, Liverman CT, Schultz AM, Strawbridge LM, editors. Epilepsy Across The Spectrum: Promoting Health And Understanding. Washington DC: National Academies Press; 2012.
7. Landro L. How apps can help manage chronic diseases. Wall St J 2017 Jun 25:- [FREE Full text]
8. Glasgow RE, Davis CL, Funnell MM, Beck A. Implementing practical interventions to support chronic illness self-management. Jt Comm J Qual Saf 2003 Nov;29(11):563-574. [Medline: 14619349]
9. Coleman MT, Newton KS. Supporting self-management in patients with chronic illness. Am Fam Physician 2005 Oct 15;72(8):1503-1510 [FREE Full text] [Medline: 16273817]
10. Ernsting C, Dombrowski SU, Oedekoven M, O´Sullivan JL, Kanzler M, Kuhlmey A, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Apr 5;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
11. Fox S. Pew Research Center. 2013 Dec 4. Health and technology in the U.S URL: http://www.pewinternet.org/2013/12/04/health-and-technology-in-the-u-s/ [accessed 2018-05-28] [WebCite Cache ID 6zkNMjxq4]
12. Fox S. Pew Research Center's Internet & American Life Project. 2015 Jan 15. The social life of health Information URL: http://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/ [accessed 2018-05-28] [WebCite Cache ID 6zkNTIaaH]
13. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

XSL•FO
RenderX

14. National Cancer Institute. 2015. Health Information National Trend Survey URL: https://hints.cancer.gov/default.aspx [accessed 2018-05-28] [WebCite Cache ID 75e9hJyCN]
15. Bernell S, Howard S. Use your words carefully: what is a chronic disease? Front Public Health 2016;4:159 [FREE Full text] [doi: 10.3389/fpubh.2016.00159] [Medline: 27532034]
16. Chen W, Lee KH. More than search? Informational and participatory eHealth behaviors. Comput Human Behav 2014 Jan;30:103-109. [doi: 10.1016/j.chb.2013.07.028]
17. 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 2018 Dec;33(12):1410-1424. [doi: 10.1080/10410236.2017.1353868] [Medline: 28872905]
18. Fox S, Purcell K. Pew Research Center's Internet & American Life Project. 2010 Mar 24. Chronic Disease and the Internet URL: http://www.pewinternet.org/2010/03/24/chronic-disease-and-the-internet/ [accessed 2018-05-28] [WebCite Cache ID 6zkNhjlfD]
19. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' acceptance of smartphone health technology for chronic disease management: a theoretical model and empirical test. JMIR Mhealth Uhealth 2017 Dec 6;5(12):e177 [FREE Full text] [doi: 10.2196/mhealth.7886] [Medline: 29212629]
20. James DC, Harville C. eHealth literacy, online help-seeking behavior, and willingness to participate in mHealth chronic disease research among African Americans, Florida, 2014-2015. Prev Chronic Dis 2016 Dec 17;13:E156 [FREE Full text] [doi: 10.5888/pcd13.160210] [Medline: 27854421]
21. Fox S, Duggan M. Pew Research Center's Internet & American Life Project. 2013 Jan 28. Tracking for Health URL: http://www.pewinternet.org/2013/01/28/tracking-for-health/ [accessed 2018-05-28] [WebCite Cache ID 6zkNoGYKH]
22. 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]
23. Sorondo B, Allen A, Fathima S, Bayleran J, Sabbagh I. Patient portal as a tool for enhancing patient experience and improving quality of care in primary care practices. EGEMS (Wash DC) 2016;4(1):1262 [FREE Full text] [doi: 10.13063/2327-9214.1262] [Medline: 28203611]
24. Centers for Disease Control and Prevention. Center for Chronic Disease Prevention and Health Promotion. 2017 Jun 28. Chronic Disease Overview URL: https://www.cdc.gov/chronicdisease/overview/index.htm [accessed 2018-05-28] [WebCite Cache ID 6zbFIANdi]
25. Anderson M, Perrin A. Pew Research Center. 2017 May 17. Tech Adoption Climbs Among Older Adults URL: http://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2018-05-28] [WebCite Cache ID 6zkQzo2OW]
26. Escoffery C. Gender similarities and differences for e-Health behaviors among U.S. adults. Telemed J E Health 2018 May;24(5):335-343. [doi: 10.1089/tmj.2017.0136] [Medline: 28813630]
27. Perrin A, Jiang J. Pew Research Center. 2018 Mar 14. About a quarter of U.S. adults say they are 'almost constantly' online URL: http://www.pewresearch.org/fact-tank/2018/03/14/about-a-quarter-of-americans-report-going-online-almost-constantly/ [accessed 2018-11-12] [WebCite Cache ID 73scM8HYh]
28. 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]
29. Paige SR, Krieger JL, Stellefson M, Alber JM. eHealth literacy in chronic disease patients: an item response theory analysis of the eHealth literacy scale (eHEALS). Patient Educ Couns 2017;100(2):320-326. [doi: 10.1016/j.pec.2016.09.008] [Medline: 27658660]
30. Coughlin SS, Prochaska JJ, Williams LB, Besenyi GM, Heboyan V, Goggans DS, et al. Patient web portals, disease management, and primary prevention. Risk Manag Healthc Policy 2017 Apr 7;10:33-40 [FREE Full text] [doi: 10.2147/RMHP.S130431] [Medline: 28435342]
31. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015 Feb 20;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]
32. Patel V, Barker W, Siminerio E. The Office of the National Coordinator for Health Information Technology. Washington DC; 2015 Oct. Trends in Consumer Access and Use of Electronic Health Information URL: https://tinyurl.com/yy8u2g26 [accessed 2018-05-28] [WebCite Cache ID 75eAwM16p]
33. Fox S, Duggan M. Pew Research Center's Internet & American Life Project. 2012 Nov 8. Mobile Health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [accessed 2018-05-28] [WebCite Cache ID 6zkR4I2B9]
34. 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;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
35. 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 1;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578] [Medline: 27920321]

XSL•FO
RenderX

36. Escoffery C, McGee R, Bidwell J, Sims C, Thropp EK, Frazier C, et al. A review of mobile apps for epilepsy self-management. Epilepsy Behav 2018 Apr;81:62-69. [doi: 10.1016/j.yebeh.2017.12.010] [Medline: 29494935]
37. Huckvale K, Morrison C, Ouyang J, Ghaghda A, Car J. The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. BMC Med 2015 Mar 23;13(1):58 [FREE Full text] [doi: 10.1186/s12916-015-0303-x] [Medline: 25857569]
38. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
39. Kim S. An exploratory study of inactive health information seekers. Int J Med Inform 2015 Feb;84(2):119-133 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.10.003] [Medline: 25453277]

Abbreviations

eHEALS: eHealth Literacy Scale
eHealth: electronic health

Edited by G Eysenbach; submitted 07.06.18; peer-reviewed by B Arnoldussen, D Zhang; comments to author 27.09.18; revised version received 14.11.18; accepted 10.12.18; published 05.03.19.

Please cite as:
Madrigal L, Escoffery C
Electronic Health Behaviors Among US Adults With Chronic Disease: Cross-Sectional Survey
J Med Internet Res 2019;21(3):e11240
URL: https://www.jmir.org/2019/3/e11240/
doi:10.2196/11240
PMID:30835242

©Lillian Madrigal, Cam Escoffery. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 05.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Characterizing the Digital Health Citizen: Mixed-Methods Study Deriving a New Typology

John Powell[1*], PhD, FFPH; Ulrike Deetjen[2*], MA (Oxon), DPhil (Oxon)

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

[2]Oxford Internet Institute, University of Oxford, Oxford, United Kingdom

[*]all authors contributed equally

Corresponding Author:
John Powell, PhD, FFPH
Nuffield Department of Primary Care Health Sciences
University of Oxford
Radcliffe Observatory Quarter
Woodstock Road
Oxford, OX2 6GG
United Kingdom
Phone: 44 1865617768
Email: john.powell@phc.ox.ac.uk

Abstract

Background: A key challenge for health systems harnessing digital tools and services is that of digital inclusion. Typically, digital inequalities are conceptualized in relation to unequal access or usage. However, these differences do not fully explain differences in health behavior as a result of health-related internet use.

Objective: Our objective was to derive a new typology of health internet users based on their antecedent motivations and enablers, to explain how individuals' different orientations influence their health behavior.

Methods: We used a mixed-methods design using (1) qualitative data from 43 semistructured interviews about individuals' general and health-related internet use, and how this influenced their health perception and their help-seeking decisions, and (2) quantitative data from the Oxford Internet Surveys (OxIS), a household survey of 2150 adults in England about their internet use and other characteristics. We used the interview data to identify constructs that described motivations and enablers affecting how internet use shaped respondents' health perception and health service use. We then used these constructs to identify variables in OxIS, which provided a quantitative measure of these constructs. We then undertook a hierarchical cluster analysis of these constructs, using the numerical variables, to derive a proposed typology of health information seekers.

Results: Both the qualitative findings and the subsequent cluster analysis suggested the existence of 6 types of individuals, categorized as learners, pragmatists, skeptics, worriers, delegators, and adigitals. Learners had a strong desire to understand health better. They used the internet to make decisions about whether they needed to see a professional and to learn about their and others' health. Pragmatists primarily used the internet to decide whether seeing a doctor was worthwhile. Skeptics were skeptical of physicians and the medical system and valued the internet for solving health problems that doctors may not be able to deal with. Worriers found it difficult to interpret health information online, described health information seeking online as frightening, and reported a critical attitude toward online health information despite seeking it frequently. Delegators comprised nonusers and users valuing the internet as an information source, but not necessarily wanting or being able to use the internet themselves. Adigitals comprised many nonusers, but also users, who did not see the internet as a useful information tool and presented strong views on its low suitability for health care.

Conclusions: This research supports a shift in the understanding of the digital divide in health, away from only access and usage issues, toward also conceptualizing an outcomes divide, whereby different types of health behavior result from the differing orientations of internet users accessing online health information. This new typology can be used to inform digital inclusion policies in health systems.

(J Med Internet Res 2019;21(3):e11279) doi:10.2196/11279

XSL•FO
RenderX

KEYWORDS

eHealth; health information seeking; perceived health; health service use; health outcomes; digital divide; digital inequalities

Introduction

Background

Health systems are under increasing pressure to save money and improve quality [1]. There are hopes that over the next decade, digital health, and specifically the internet harnessed as a health service tool, can address these aims by shaping individuals' service use and their health perceptions. As new digital tools and new models of health service provision emerge, there are implicit assumptions that health consumers will take on new roles and responsibilities as digital health citizens [2].

In this brave new world of technology-enhanced health care, where digital technology is increasingly becoming a determinant of health, a key challenge is that of digital inclusion. Even though the digital divide in terms of internet access may have been reduced with rising internet penetration, inequalities remain in the ability to make meaningful use of online resources and to obtain benefits from doing so [3,4]. Achieving health outcomes depends on enablers of internet skills [5-7], particularly in terms of making sense of information quantity and quality [8], and health literacy, in the sense of being able to translate the findings into health-promoting behaviors [9-11]. In addition, motivations differ: not everybody can and wants to assume responsibility for their health [12,13]. Differences in outcomes are also reflected in the general shift toward referring to digital inequalities rather than to a single digital divide [14,15], and to distinguishing between the *access divide* around the turn of the millennium, the *usage divide* over the last 10 years [3], and, more recently, a third-level digital divide in relation to the benefits that individuals obtain online [4,16].

Objective

Following these developments, this research drew on empirical data to propose a new typology of individuals' orientations toward online health information. These orientations shape how individuals use the internet in the health context in the first place, but also, as we argue in this paper, they can explain individuals' differing health behaviors that occur as a consequence [17,18] in terms of their health perceptions and health service use (with implications for health outcomes). We propose the *outcomes divide* as a conceptualization of digital inequalities in the context of health outcomes: that is, even accounting for differences in access and usage, there remain inequalities in the health outcomes achieved through the use of digital tools by these different types of individuals.

We derived this typology by combining findings from both qualitative and quantitative work. We did this by identifying antecedent factors in individuals' motivations (their attitudes and desire to use health information from online sources) and enablers (their ability and interest in accessing the internet, and making sense of the information found). This is in line with various studies that found that antecedent factors must be incorporated when analyzing internet use [19,20], in particular with respect to individuals' motivations and enablers [3]. Motivations and enablers form a complex relationship with how individuals obtain outcomes online: attitudes, awareness of technology, desire for information, job requirements, skills, and social contacts shape how individuals use technology and what they need to get out of it—and their health behaviors and subsequent outcomes in turn shape their future expectations of use [17,18].

Methods

Research Design

We undertook a mixed-methods research design using both face-to-face interviews and quantitative analysis of a survey dataset, as part of a larger study about the relationship between internet use and health outcomes. We conducted 43 face-to-face interviews about individuals' general and health-related internet use, and how this influenced their health perception and their help-seeking decisions. We recruited many of these interviewees through their participation in the survey we used for the quantitative analysis: the Oxford Internet Surveys (OxIS; 2013 report). This is a random-sample survey conducted biannually since 2003, based on randomly sampled output areas (statistical areas of about 300 individuals formed based on sociodemographic homogeneity), and then randomly selected individuals within these. Using a traditional pen-and-paper method, OxIS collects data on online and offline activities, attitudes, and skills for 2150 internet users and nonusers in England (for details of the survey methods and questions asked, see the OxIS website [21]). The Central University Research Ethics Committee of the University of Oxford approved this study (number: OII C1A 14-003).

We recruited 31 interview participants from respondents to OxIS following a 2-stage sampling process: first, we purposively selected 14 output areas from OxIS to obtain areas with diverse urban and rural characteristics and area classifications; second, we contacted all OxIS participants in these output areas between July and November 2014, and the lead author interviewed all those who agreed to participate in this follow-up research. Using former OxIS participants as interviewees for the qualitative part allowed for comparison of the qualitative and quantitative elements for single individuals. We identified an additional 12 interviewees through approaching further interviewees in public spaces of the output area (purposive sampling) and by contacting others based on recommendations of OxIS participants to include particularly information-rich cases and improve the coverage of the population (snowball sampling). As a result, the interviewees included in the final sample covered the full sociodemographic spectrum in terms of sex (there were 25 female and 18 male participants), age (range 19-85 years), educational attainment, National Statistics Socio-economic Classification (NS-SEC; from 1-4, with 1 denoting the highest socioeconomic status), and long-term health conditions.

The interviews lasted about 50 minutes on average and followed a topic guide to ask about general internet use first, then health-related internet use and specific instances of

XSL•FO
RenderX

health-related use, and effects and outcomes of internet use, before concluding with final reflections, next steps, and sociodemographic information. Except for 2 interviews in which the interviewees requested to talk on the phone, all interviews were conducted in person, mainly at the individuals' homes, and partially in public cafés. Informed consent was obtained from each participant in writing (and orally for the 2 telephone interviews). For the analysis, the interviews were audiorecorded and transcribed by 1 of the authors (UD). All interviews were manually coded by 1 of the authors (UD). Emerging codes were discussed with 2 project supervisors (the coauthor and another researcher) who also read samples of the transcripts. The coding was inductive, emerging from the data, and therefore not biased toward the topics from the quantitative data, which also allowed for unexpected themes to emerge [22].

Analytic Approach

Our mixed-methods analytic approach was to identify constructs from the qualitative interviews that described motivations and enablers affecting how internet use shaped the interviewees' health perception and health service use, and then to identify variables in OxIS that provided a quantitative measure of these constructs. While this approach has limitations in that we were restricted to variables already collected in the OxIS dataset, using previously collected OxIS data nevertheless had advantages in that the dataset was a well-constructed existing sample where we could link our qualitative findings to the quantitative data.

Having identified variables in the quantitative dataset that broadly matched our emergent qualitative constructs, we then derived a proposed typology of health information seekers using a cluster analysis of the numerical variables. We used hierarchical clustering, an algorithmic approach that groups individuals with similar observations into clusters based on the distance between their values. It is hierarchical, as it starts with the 2 values closest together and groups those, then continues in a stepwise fashion grouping the next closest, until all are grouped (ie, eventually, if the process is continued, there will be only 1 cluster). The underlying distance measure we used was Ward's linkage with squared euclidean distances, since it provided better results than other hierarchical methods in general [23], and specifically for OxIS [24]. We evaluated the best fit of the clustering solution based on within- or between-cluster distance and entropy [25], and by comparing the dendrogram cluster output with the emergent findings from the qualitative data.

Results

Motivations and Enablers

In line with the central role of antecedent factors for understanding internet use and its outcomes, 5 main motivations for health-related internet use emerged from the qualitative interview data: convenience and speed of access at all times; preparing for appointments; "translating" health professionals' advice through nonmedical terminology online; building up further health-related knowledge; and connecting with others to get peer advice. The qualitative findings also suggested that these enablers could be traced back to 4 interrelated prerequisites for using the internet: devices to connect to the internet; skills influencing ability to read and use online information (including both technical skills and health literacy); interest in using the internet; and appropriate opportunities for use.

Having identified these influence factors from the qualitative interview work, we mapped these onto variables that had been measured in the OxIS data, creating 8 constructs for which we had quantitative data (the mapping was not perfect, as discussed in the Limitations section below). First, constructs relating to *internet usefulness*— the internet being an efficient means of finding information, making life easier and helping to save time—reflect the internet's convenience. Second, the motivation relating to interpreting and extending professional advice and building up further knowledge is encapsulated in individuals' *learning attitudes*. Third, the motivations for building up further knowledge also reflects a certain level of *online enjoyment*: enjoying reading and understanding all about certain topics online. Fourth, people's attitudes toward medical professionals were revealed in the motivation to check on the doctor, for which the quantitative concept of *trust in medical doctors* may be an acceptable reflection.

With respect to enablers, fifth, self-rated *internet skills* conceptualize the skills dimension. We recognize that it would additionally have been desirable to include a specific measure of health literacy (as separate from technical skills) to account for the ability to find and carefully interpret medical information online [26]; however, this was not available in the OxIS data. Sixth, for the attitudinal aspect captured in the dimensions of interest and usage opportunities, *internet interest* reflects an individual's desire to access and use the internet. Seventh, *technology attitude*, about how individuals view the general upsides and downsides of technology, relates to the wider attitudinal aspects in relation to enablers (especially "interest in using the internet" from our qualitative work). Eighth, we included the OxIS variable for *self-efficacy*, capturing to what extent individuals consider themselves as actively shaping their health, because this theme surfaced in several of our interviews and also reflects the extant health behavior literature [27].

A Typology of Health Information Seekers

Having identified constructs qualitatively, and where possible identified quantitative variables from the OxIS data that measured these, our next step was to conduct the cluster analysis using OxIS. On the basis of the dendrogram shown in Figure 1, we could have justified different numbers of groups, but we chose 6 clusters not only as being supported by the cluster analysis, but also because this reflected (and supported) the qualitative findings because, over the course of the interviews, the emergent findings had indicated that individuals could be grouped into 6 different types in terms of their motivations and enablers.

The average within-cluster distance was smaller (2.5) than the between-cluster distance (4.3), with an entropy value of 1.7, so that the quality of the solution was in the range of other cluster models [25].

Based on the distribution of the constructs within each type, we named the 6 types *the learners*, *the pragmatists*, *the skeptics*,

XSL•FO
RenderX

the worriers, *the delegators*, and *the adigitals* (Figure 2). Despite a user-nonuser split visible in the dendrogram, these types cut across users and nonusers: 0.7% (3/422) to 4.1% (10/245) of individuals in the first 4 types were nonusers (of whom nearly all cited a lack of devices or skills as their main reason for nonuse), while 28.4% (65/229) of the delegators and 90.6% (491/542) of the adigitals did not use the internet. All groups included those with long-term health conditions (between 50/514, 9.7% and 46/229, 20.0% across the first 5 types), particularly the adigitals (262/542, 48.3%). The frequency of health information seeking also differed, with the worriers looking up health information most frequently (μ=1.8 on a 5-item Likert scale).

Figure 1. Dendrogram for hierarchical clustering of typology. Percentages are from weighted Oxford Internet Surveys data.

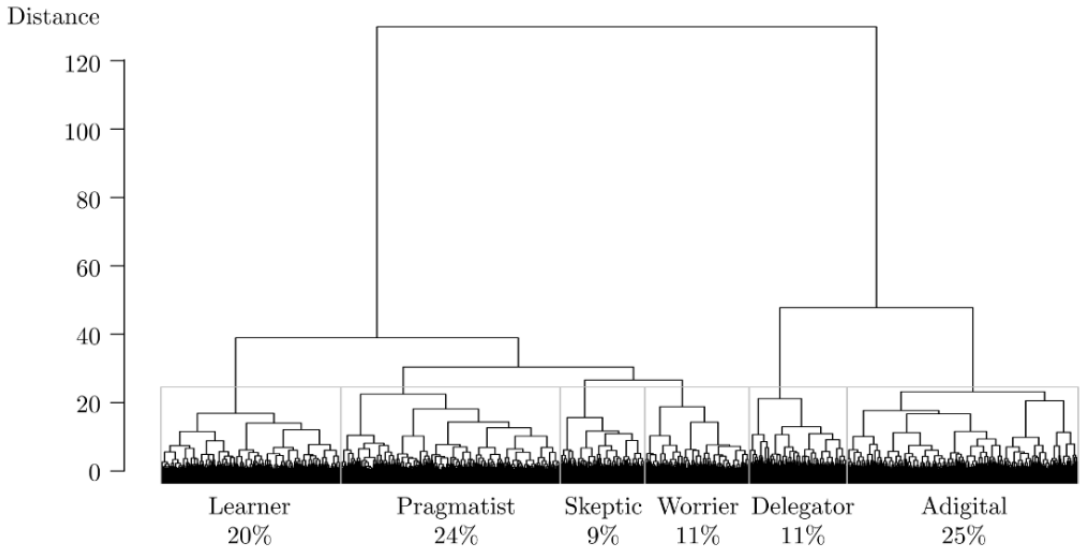


Figure 2. Typology of health information seekers showing cluster dimensions. All values are mean [SD]. The diagram shows the divergence from the arithmetic mean for each of the clustering dimensions. All constructs are measured on 5-item Likert scales.

	Learner	Pragmatist	Skeptic	Worrier	Delegator	Adigital
Learning attitude	**4.3** [0.4]	**3.7** [0.6]	**3.7** [0.5]	**3.8** [0.5]	**3.6** [0.7]	**2.9** [0.7]
Online enjoyment	**4.0** [0.6]	**2.8** [0.8]	**3.7** [0.6]	**3.7** [0.7]	**3.0** [1.0]	**2.6** [0.8]
Trust in doctors	**4.3** [0.7]	**4.1** [0.6]	**2.5** [0.8]	**4.1** [0.6]	**3.6** [1.2]	**3.9** [1.0]
Self-efficacy	**4.3** [0.5]	**3.9** [0.7]	**3.6** [0.8]	**2.7** [0.7]	**3.8** [0.9]	**3.7** [0.9]
Technology attitude	**4.1** [0.6]	**3.4** [0.6]	**3.4** [0.5]	**3.7** [0.6]	**3.1** [0.6]	**2.4** [0.6]
Internet usefulness	**4.6** [0.4]	**4.0** [0.6]	**4.0** [0.5]	**3.9** [0.5]	**3.8** [0.7]	**3.2** [0.8]
Internet skills	**4.4** [0.7]	**3.8** [0.9]	**3.7** [0.7]	**4.2** [0.7]	**3.4** [0.7]	**1.2** [0.4]
Internet interest	**4.6** [0.5]	**4.3** [0.5]	**4.1** [0.4]	**4.6** [0.5]	**1.7** [0.5]	**1.1** [0.3]

XSL•FO
RenderX

Table 1. Logistic regression for types of health information seekers(N=2150; largest condition index=4).

Independent variables	Types of health information seekers											
	Learner		Pragmatist		Skeptic		Worrier		Delegator		Adigital	
	Odds ratio	*P* value	Odds ratio	*P* value	Odds ratio	*P* value	Odds ratio	*P* value	Odds ratio	*P* value	Odds ratio	*P* value
Learning attitude	1.83	<.001	0.94	.50	0.55	<.001	0.51	<.001	1.44	.049	0.15	<.001
Online enjoyment	5.32	<.001	0.45	<.001	1.07	.58	1.13	.33	1.35	.09	0.13	<.001
Trust in doctors	2.76	<.001	1.21	.009	0.10	<.001	3.80	<.001	0.97	.80	0.72	.22
Self-efficacy	5.14	<.001	3.02	<.001	0.44	<.001	0.05	<.001	0.81	.13	1.14	.70
Technology attitude	1.86	<.001	0.54	<.001	0.56	.001	0.79	.21	0.82	.34	0.58	.31
Internet usefulness	1.80	<.001	0.71	<.001	0.43	<.001	2.70	<.001	2.68	<.001	0.30	.003
Internet skills	3.99	<.001	0.60	<.001	3.04	<.001	0.55	.01	73.69	<.001	0.00	<.001
Internet interest	4.14	<.001	17.98	<.001	14.15	<.001	20.03	<.001	0.00	<.001	0.00	<.001
Age[a]	1.04	.76	1.15	.11	1.18	.27	0.88	.37	0.59	.004	0.82	.70
Sex[b]	1.06	.52	1.00	.98	1.06	.61	1.90	.34	1.25	.11	0.94	.86
Education[c]	0.88	.25	1.17	.06	1.09	.498	0.94	.65	0.98	.92	0.86	.68
NS-SEC[d]	1.04	.73	0.98	.75	1.10	.43	0.98	.91	0.83	.25	0.78	.54
Long-term health condition[e]	0.92	.49	0.90	.19	0.92	.55	0.93	.56	0.78	.08	0.95	.87
Pseudo-R^{2}[f]	0.73	—[g]	0.48	—	0.62	—	0.71	—	0.77	—	0.98	—

[a]Measured as a continuous variable (in years).

[b]Binary (male or female).

[c]The highest level of qualifications attained (none, primary, secondary, further, higher).

[d]1 of 5 categories in the National Statistics Socioeconomic Classification (NE-SEC; professional, intermediate, manual, unemployed, student)

[e]Binary (yes or no).

[f]Following Cragg and Uhler's definition.

[g]Not applicable.

To delineate the types using the clustering dimensions and further sociodemographic characteristics, Table 1 presents a logistic regression on each type. It shows that sociodemographic features are less informative for determining individuals' types and illustrates that individuals are not deterministically assigned a type, but rather have a higher probability of falling into one type or another based on their antecedent factors.

Finally, we returned to the qualitative interview data, both to confirm the validity of the clusters (that they were indeed reflected in the qualitative findings) and to illustrate the types by describing a representative of each one in more detail, as follows (real names have been replaced with pseudonyms).

The *learners* had a strong desire to understand health better. They used the internet to make decisions about whether they needed to see a professional and to learn about their and others' health. Miriam (age 58 years, NS-SEC class 2, education level 2 of 4), who had had a minor stroke and arthritis, is an example of this group. Before and after seeing the doctor, she usually read everything she could find, in line with the generally high learning attitude of the learners ("The doctors give you the basic information that you need to know, this is what it's called, [but] it explains a lot more on the internet."). Consequently, she also reduced her health service use ("I still do go back to that and rather than go to the doctors and say, well this has happened…I would go on there and reread through it." *)*. Her high level of online enjoyment and skills was also reflected in Miriam's enjoyment in looking up health information, also for other people's conditions ("I do like to read up on these, so I see if there's something I haven't got, just see how people...deal with a situation"). Miriam also set herself up as a lay expert for other people in her social circle ("If somebody's brought something up that something's happened to them, I tell them that I would give them my point of view once I've read up on it on the internet.").

The *pragmatists* primarily used the internet to decide whether it was worth seeing a doctor. For example, Hugh (age 45 years, NS-SEC class 1, education level 4) primarily wanted to make quick decisions about health service use and the required urgency, with a low level of enjoyment ("I would use it just to kind of get a handle on whether it's worth going to the doctor or not, not for kind of detailed self-diagnosis."). Hugh also showed high skills and high trust in doctors combined with professional respect ("I might challenge [the doctor], but I let him...give his diagnosis first.…You have to rely on the expertise of the professional first."). Like other pragmatists, Hugh did not want to share and discuss health problems online but valued official information and online health services ("I wouldn't be

XSL•FO
RenderX

any more concerned about privacy or security than with internet banking."). Like Hugh, many pragmatists showed a high understanding of the need for new approaches to health provision ("I'm all for putting less pressure on the health service through faster [and] more efficient forms of medical support...[but] in Britain people are very proud of their health service...[it's] sacred ground.").

In contrast, the *skeptics* are skeptical of physicians and the medical system and value the internet for solving health problems that doctors may not be able to deal with, as Brian (age 53 years, NS-SEC class 1, education level 3) openly showed ("The doctors are more and more useless now as time goes on…[online] there's forums and you can cross-reference things a bit better, rather than [depend on] the opinion of one person like the GP."), also due to the availability of user-written information. Brian thought the internet reduced his health service use overall *(* "Probably if I didn't have the information via the internet, maybe I would need to go and see the doctor more often." *)*, although he also provided examples of increased health service use *(* "I have high blood pressure, and [the doctor] said that's nothing to worry about.…So I took my own blood pressure readings...went back to the doctor, and he said, okay, I'll give you some blood pressure tablets."). In general, the skeptics mainly saw the doctor's role as a provider of medication.

The *worriers* found it difficult to interpret health information online, describing health information seeking online as frightening and reporting a critical attitude toward it despite seeking it frequently. Helen (age 43 years, NS-SEC class 1, education level 4) had epilepsy and enjoyed browsing through health topics ("It's not necessarily about epilepsy; it's other things. I can spend ages on it, going on things that aren't relevant to me, but I can also really forget most information."). However, health information made her feel afraid ("I'm afraid of what I might find. If you're on your own and look at a website, and find something really bad, [it's] really dangerous."). In line with the low self-efficacy of this group, Helen did not proactively want to address epilepsy ("I haven't brought up [my children's potential epilepsy] with my doctors. Because I think I'm afraid to do it. And I haven't looked on websites because it's very personal."). She exhibited a high level of trust in her doctors ("It's really important that you work with your doctors and your specialist and not go on the website, because it could really make it worse."), showed a strong normative attitude about health information online ("I only really look up official websites...I've always been told by people not to look up health care...because you always see the horror stories"), and said that she would value recommendations about specific websites from her doctors .

The *delegators* were composed of nonusers and users valuing the internet as an information source, but not necessarily wanting or being able to use the internet themselves. Kathleen (age 75 years, NS-SEC class 1, education level 0) had elaborate networks to access health information ("I can go and get it from the library....If I really do want more, I've got a friend in London who's got a computer, and she would...phone me back and tell me or she'll send it down to me.") and valued the comfort of doing so ("I don't know whether I really do need to have a computer, because anything like that I can find out."). Kathleen also actively read health information in the local newspaper, showing her interest in health and feeling responsible for her well-being. While trust in doctors differed among delegators, as evidenced by the high standard deviation in this group, Kathleen placed high trust in her doctor, and—illustrating again the delegation aspect—valued that he followed up newspaper articles that she took into the consultation ("He doesn't just do it like from what it says in the paper...but takes it home and googles it on his computer.").

The *adigitals* comprised many nonusers, but also users, who did not see the internet as a useful information tool and presented strong views on its low suitability for health care. Charlotte (age 78 years, NS-SEC class 2, education level 0) was a nonuser and generally did not like to work with technology, also due to health-related reasons ("My fingers have never worked in a way that I can use a keyboard of any sort, piano, computer keyboard, I can't separate them. So I lost patience."). For health information, she either would ask the doctor or, above all, thought that she knew what was best for her ("In most ways you know your own body....[I follow] just my own instincts. And I found out that they never let me down, fortunately."). While the adigitals did not show a consistent picture of trust in medical professionals, they generally expressed concerns about how people use the internet for health information ("They'll worry themselves into goodness knows what and they do the same on the computer—as soon as she sneezes she looks it up on the computer.").

Discussion

Principal Findings

In line with the shifting digital divide from the *access* to the *usage* divide [3,15], this research supports a further shift to conceptualizing an *outcomes* divide or outcomes inequalities, whereby different types of health behavior result from the differing orientations of internet users accessing online health information. In contrast to the existing literature [4], the outcomes divide we propose in this paper may only partially be traced back to sociodemographic factors, as understanding outcomes requires a more nuanced view not necessarily following the user characteristics underlying the several existing digital divide conceptualizations [28]. This is not to downplay the importance of sociodemographic factors and other structural conditions that can shape internet use. For example, previous work has shown the strong influence of socioeconomic status on internet-related attitudes and behaviors [29]. But in our findings these factors, while important at an overarching level, did not have different influences on our various categorization of types, where we were particularly interested in determining how motivations and enablers clustered into types.

This study showed the central role of antecedent factors to internet use for influencing behaviors. The qualitative data confirmed motivations [30,31] and enablers [32] found in previous research, and indicated that outcomes may be shaped by types of health information seekers formed based on these motivations and enablers. This confirms that previously existing health behaviors translate to the online realm [29,30,33], and

XSL•FO
RenderX

emphasizes attitudes and skills as mediators for internet outcomes [17]. In that sense, online health resources become part of normal health practices, help seeking, and everyday life information seeking [34-36]. This ties in with the wider argument that individuals use technologies to satisfy existing needs, with technological innovation merely creating new ways of doing so [37], as reflected in theories about the social shaping of technology [38].

The qualitative data suggested that 2 of our 6 types, the learners and the pragmatists, use the internet efficiently in health-related contexts, both to increase the appropriateness of their health service use and, in the case of the learners, to gain self-efficacy for self-care and extending professional advice. In that sense, individuals in both groups consistently gain benefits from using the internet.

Then again, outcomes do not necessarily have to be positive. For the skeptics, the suggested relationship to perceived health was negative, which may indicate that using the internet was less beneficial than the skeptics thought. In support of this, other research found that low-trust individuals tended to substitute physician services with health information online [39,40], and while those skeptical of medical care had lower health service use, they also often showed worse health behaviors and lower health perception [41,42]. For the worriers, internet use was barely associated with any changes in perceived health and health service use. While health information seeking does not necessarily lead to higher health service use and worse health perceptions, the findings indicated that this group did not necessarily realize any outcomes, partially because they stopped looking up health information as a consequence.

Finally, the effects of internet use were lowest for the delegators and the adigitals, although with higher effect sizes for the delegators. Some nonusers in the delegator group used the internet more intensely (via intermediaries) than users, building support networks with different individuals for different purposes. While both the delegators and the adigitals largely comprised nonusers, this shows how internet outcomes may not follow the lines of the user-nonuser split. This is further corroborated by the relatively similar outcomes for users and nonusers in the delegator group, which may partially be due to their preference for outsourcing health-related information seeking.

The typology introduced in this research therefore presents a tool for systematizing orientations toward health information seeking to conceptualize the outcomes divide. This is similar to other typologies in research on health and internet use [43,44], which serve to “shift study of the Internet away from an overly narrow focus on comparing users and non-users, and [focus] more research and debate on other variations among users and non-users that have equally significant implications for the future of the Internet” [24] (pg 9).

Limitations

A limitation of our approach is that, for 6 of the 31 OxIS interviewees, the quantitative and qualitative classifications did not correspond, so that we manually reclassified these based on the qualitative data. There are three reasons for this mismatch. First, in line with the process-based models of perceived health [45], individuals may transcend the type boundaries over time, which became evident in 2 interviewees’ altered views of the health system after major health incidences. Second, the typology lacks some specificity due to the absence of some items in the OxIS question set. These include the lack of a measure of health literacy, no measure of acute health incidents (only including the presence or absence of long-term health conditions), and only limited information on situational factors such as devices and use opportunities. Even though health information seeking is similar to other informational activities online [46], 2 interviewees had particular attitudes about health online that differed from their general internet views. Third, 2 interviewees showed different attitudes in the interview for no evident reason, which highlights the constructed nature of survey and interview data, and the challenges of their triangulation.

Future Research and Implications

Further research should attempt to replicate and refine the developed typology, ideally with health-related dimensions by including more specific constructs of health literacy and health-related self-efficacy [7,47,48]. Due to transitions of individuals between types over time, the typology should also be based on longitudinal measurements on the same individuals from multiple points in time, as individuals may develop and fall into different types following different events in their lives. For the broader context of internet research, it would be relevant to understand whether the established typology also describes orientations that are relevant for other internet-based outcomes. This would cross-validate the results of this research, and thereby provide theoretical support for a more general outcomes divide beyond the health context.

These type-based findings have implications for policy and practice, particularly for health systems seeking to maximize digital inclusion. A multifaceted approach is required to address the differing needs of the 6 types. Nonusers of the learner and pragmatist type in particular should be provided with access to digital resources, as most individuals of these types cited the nonavailability of devices as the main reason for nonuse. In addition, worriers may benefit from additional guidance: they highly appreciated medical professionals and suggested that doctors should recommend specific websites reflecting their preference for professional guidance and managing uncertainty [49]. They would, for example, use initiatives that provide official endorsement or certification of digital resources and health apps. This stands in contrast to the skeptics, who—independently of the internet—might benefit from building up trust in doctors and the medical system to change their health behaviors, evaluate the appropriateness of health service use, and ensure compliance with medical recommendations [40,41].

Particularly for delegators, but also for all others with a lower level of skills, the social environment is of crucial importance for internet-based outcomes. As the type name implies, obtaining value from the internet depends on being able to delegate. Here, addressing social and digital inclusion becomes a joint priority, where not everyone has to be online, but everyone should benefit from online resources. Finally, for the adigitals, it may be most

XSL•FO
RenderX

important to address motivations for internet use. This applies to users and nonusers alike: 79.8% (392/491) of nonusers of this type referred to a lack of interest as the main reason for nonuse, whereas users mainly explained why the internet was not suitable for health-related matters.

Conclusion

This research showed how health internet users may be conceptualized based on a typology of 6 orientations toward online health information seeking. The findings illustrate that the digital divide is increasingly more complex to delineate [50], indicate that previously existing health behaviors translate to the online realm [32], and support the shift toward an outcomes divide in terms of the benefits that individuals of differing types may obtain online [4]. This research also showed that health behaviors (and, by deduction, possibly health outcomes) are primarily shaped by antecedent factors such as motivations and enablers [3,17], rather than sociodemographic factors [4]. This research therefore makes, to our knowledge, one of the first empirical contributions to an emerging literature assessing how differences in outcomes represent the next stage of continuously shifting digital inequalities.

Acknowledgments

JP is funded by the UK National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care Oxford at Oxford Health NHS Foundation Trust. UD was funded by the Clarendon Fund, the UK Economic and Social Research Council and Balliol College, University of Oxford, Oxford, UK. The views expressed are those of the authors and not necessarily those of the funders, the NHS or the Department of Health and Social Care.

Conflicts of Interest

None declared.

References

1. Crawford R, Emmerson C. NHS and social care funding: the outlook to 2021/22. Research report. London, UK: Nuffield Trust and Institute for Fiscal Studies; 2012. URL: https://www.nuffieldtrust.org.uk/files/2017-01/nhs-social-care-funding-outlook-2021-22-web-final.pdf [accessed 2018-06-12] [WebCite Cache ID 707ZNg6Io]
2. Powell J, Newhouse N, Boylan A, Williams V. Digital health citizens and the future of the NHS. Digit Health 2016 Oct 10;2:205520761667203. [doi: 10.1177/2055207616672033] [Medline: 29942569]
3. van Deursen AJ, van Dijk JA. The digital divide shifts to differences in usage. New Media Soc 2013 Jun 07;16(3):507-526. [doi: 10.1177/1461444813487959]
4. van Deursen AJ, Helsper EJ. The third-level digital divide: who benefits most from being online? In: Robinson L, Cotten SR, Schulz J, Hale TM, Williams A, editors. Communication and Information Technologies Annual: Digital Distinctions and Inequalities. Volume 10. Bingley, UK: Emerald Group Publishing Limited; 2015:29-52.
5. 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]
6. Birru MS, Monaco VM, Charles L, Drew H, Njie V, Bierria T, et al. Internet usage by low-literacy adults seeking health information: an observational analysis. J Med Internet Res 2004 Sep 3;6(3):e25 [FREE Full text] [doi: 10.2196/jmir.6.3.e25] [Medline: 15471751]
7. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 04;18(10):e264 [FREE Full text] [doi: 10.2196/jmir.6349] [Medline: 27702738]
8. 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]
9. Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008 Dec;67(12):2072-2078. [doi: 10.1016/j.socscimed.2008.09.050] [Medline: 18952344]
10. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med 2004 Dec;19(12):1228-1239 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.40153.x] [Medline: 15610334]
11. Parker R. Health literacy: a challenge for American patients and their healthcare providers. Health Promot Int 2000;15(4):277-283. [doi: 10.1093/heapro/15.4.277]
12. Henwood F, Wyatt S, Hart A, Smith J. 'Ignorance is bliss sometimes': constraints on the emergence of the 'informed patient' in the changing landscapes of health information. Sociol Health Illn 2003 Sep;25(6):589-607 [FREE Full text] [Medline: 12919447]
13. Powell J, Boden S. Greater choice and control? Health policy and the online health consumer. Policy Internet 2012;4(2):1-23. [doi: 10.1515/1944-2866.1180]
14. Robinson L, Cotten SR, Ono H, Quan-Haase A, Mesch G, Chen W, et al. Digital inequalities and why they matter. Inf Commun Soc 2015 Mar 16;18(5):569-582. [doi: 10.1080/1369118X.2015.1012532]

XSL•FO
RenderX

15. DiMaggio P, Hargittai E. From the 'digital divide' to 'digital inequality': studying Internet use as penetration increases. Working paper series 15. Princeton, NJ: Princeton University Center for Arts and Cultural Policy Studies; 2001. URL: https://www.princeton.edu/~artspol/workpap/WP15%20-%20DiMaggio%2BHargittai.pdf [accessed 2018-06-12] [WebCite Cache ID 707aF9Jah]
16. 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 Inform 2017 Dec;34(8):1607-1624. [doi: 10.1016/j.tele.2017.07.007]
17. Helsper EJ. A corresponding fields model for the links between social and digital exclusion. Commun Theor 2012 Oct 15;22(4):403-426. [doi: 10.1111/j.1468-2885.2012.01416.x]
18. Straubhaar J, Tufekci Z, Spence J, Rojas V. Digital inequity in the Austin Technopolis: an introduction. In: Straubhaar J, Spence J, Tufekci Z, Lentz RG, editors. Inequity in the Technopolis: Race, Class, Gender, and the Digital Divide in Austin. Austin, TX: University of Texas Press; 2012.
19. Reisdorf BC, Axelsson A, Maurin H. Living offline - a qualitative study of internet non-use in Great Britain and Sweden. Sel Pap Internet Res 2012;2:1-28. [doi: 10.2139/ssrn.2721929]
20. Anderson B, Tracey K. Digital living: the impact (or otherwise) of the Internet on everyday British life. In: Wellman B, Haythornthwaite C, editors. The Internet in Everyday Life. Malden, MA: Blackwell Publishers Ltd; 2008.
21. About the Oxford Internet Surveys. Oxford, UK: Oxford Internet Institute; 2018. URL: https://oxis.oii.ox.ac.uk/ [accessed 2019-01-16] [WebCite Cache ID 75TTqJMvU]
22. Silverman D. Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction. 3rd edition. London, UK: SAGE; 2006.
23. Ferreira L, Hitchcock DB. A comparison of hierarchical methods for clustering functional data. Commun Stat Simul Comput 2009 Dec 09;38(9):1925-1949. [doi: 10.1080/03610910903168603]
24. Dutton WH, Blank G, Groselj D. Cultures of the Internet: the Internet in Britain. Oxford Internet Survey 2013 Report. Oxford, UK: Oxford Internet Institute, University of Oxford; 2013. URL: http://oxis.oii.ox.ac.uk/wp-content/uploads/sites/43/2014/11/OxIS-2013.pdf [accessed 2018-06-12] [WebCite Cache ID 707aqGCfE]
25. Zhao Y, Karypis G. Empirical and theoretical comparisons of selected criterion functions for document clustering. Mach Learn 2004 Jun;55(3):311-331. [doi: 10.1023/B:MACH.0000027785.44527.d6]
26. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr 29;13(2):e35. [doi: 10.2196/jmir.1581] [Medline: 21531690]
27. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
28. Yu L. Understanding information inequality: making sense of the literature of the information and digital divides. J Libr Inf Sci 2016 Jun 30;38(4):229-252. [doi: 10.1177/0961000606070600]
29. Hale TM. Is there such a thing as an online health lifestyle? Examining the relationship between social status, internet access, and health behaviors. Inf Commun Soc 2013 May;16(4):501-518. [doi: 10.1080/1369118X.2013.777759]
30. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
31. Chapple A, Evans J, Ziebland S. An alarming prognosis: how people affected by pancreatic cancer use (and avoid) internet information. Policy Internet 2012;4(2):1-20. [doi: 10.1515/1944-2866.1162]
32. van Dijk J, Hacker K. The digital divide as a complex and dynamic phenomenon. 2000 Presented at: 50th Annual Conference of the International Communication Association; Jun 15, 2000; Acapulco, Mexico.
33. Rogers A, Mead N. More than technology and access: primary care patients? views on the use and non-use of health information in the Internet age. Health Soc Care Community 2004;12(2):102-110. [doi: 10.1111/j.0966-0410.2004.00473.x] [Medline: 19777718]
34. Nettleton S, Burrows R, O'Malley L. The mundane realities of the everyday lay use of the internet for health, and their consequences for media convergence. Sociol Health Illn 2005 Nov;27(7):972-992 [FREE Full text] [doi: 10.1111/j.1467-9566.2005.00466.x] [Medline: 16313525]
35. Kivits J. Everyday health and the internet: a mediated health perspective on health information seeking. Sociol Health Illn 2009 Jul;31(5):673-687 [FREE Full text] [doi: 10.1111/j.1467-9566.2008.01153.x] [Medline: 19220804]
36. Savolainen R. Everyday life information seeking: approaching information seeking in the context of “way of life”. Libr Inf Sci Res 1995 Jun;17(3):259-294. [doi: 10.1016/0740-8188(95)90048-9]
37. Gershuny J. Social Innovation and the Division of Labour. Oxford, UK: Oxford University Press; 1983.
38. MacKenzie DA, Wajcman J. The Social Shaping of Technology. Buckingham, UK: Open University Press; 1999.
39. Dwyer DS, Liu H. The impact of consumer health information on the demand for health services. Q Rev Econ Finance 2013 Feb;53(1):1-11. [doi: 10.1016/j.qref.2012.11.002]
40. Lee C, Hornik RC. Physician trust moderates the internet use and physician visit relationship. J Health Commun 2009;14(1):70-76 [FREE Full text] [doi: 10.1080/10810730802592262] [Medline: 19180372]
41. Fiscella K, Franks P, Clancy CM. Skepticism toward medical care and health care utilization. Med Care 1998 Feb;36(2):180-189. [Medline: 9475472]

XSL•FO
RenderX

42. Sutherland LR, Verhoef MJ. Why do patients seek a second opinion or alternative medicine? J Clin Gastroenterol 1994 Oct;19(3):194-197. [Medline: 7806827]
43. Clatworthy J, Buick D, Hankins M, Weinman J, Horne R. The use and reporting of cluster analysis in health psychology: a review. Br J Health Psychol 2005 Sep;10(3):329-358. [doi: 10.1348/135910705X25697] [Medline: 16238852]
44. Christodoulides G, Michaelidou N, Theofania Siamagka NT. A typology of internet users based on comparative affective states: evidence from eight countries. Eur J Mark 2013 Feb 08;47(1/2):153-173. [doi: 10.1108/03090561311285493]
45. Wyke S, Adamson J, Dixon D, Hunt K. Consultation and illness behaviour in response to symptoms: a comparison of models from different disciplinary frameworks and suggestions for future research directions. Soc Sci Med 2013 Jun;86:79-87. [doi: 10.1016/j.socscimed.2013.03.007] [Medline: 23608096]
46. Blank G, Groselj D. Dimensions of internet use: amount, variety, and types. Inf Commun Soc 2014 Feb 28;17(4):417-435. [doi: 10.1080/1369118X.2014.889189]
47. Estacio EV, Whittle R, Protheroe J. The digital divide: examining socio-demographic factors associated with health literacy, access and use of internet to seek health information. J Health Psychol 2017 Feb 01:1359105317695429. [doi: 10.1177/1359105317695429] [Medline: 28810415]
48. 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;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
49. Fergus TA. Cyberchondria and intolerance of uncertainty: examining when individuals experience health anxiety in response to Internet searches for medical information. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):735-739. [doi: 10.1089/cyber.2012.0671] [Medline: 23992476]
50. van Dijk JA. The Deepening Divide: Inequality in the Information Society. Thousand Oaks, CA: SAGE Publications; 2005.

Abbreviations

NS-SEC: National Statistics Socio-economic Classification
OxIS: Oxford Internet Surveys

Edited by G Eysenbach; submitted 12.06.18; peer-reviewed by E Sillence, T Hale, B Arnoldussen, S Ghaddar; comments to author 06.09.18; revised version received 01.11.18; accepted 22.11.18; published 05.03.19.

Please cite as:
Powell J, Deetjen U
Characterizing the Digital Health Citizen: Mixed-Methods Study Deriving a New Typology
J Med Internet Res 2019;21(3):e11279
URL: https://www.jmir.org/2019/3/e11279/
doi:10.2196/11279
PMID:30835238

©John Powell, Ulrike Deetjen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 05.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Impact of Superfast Broadband, Tailored Booklets for Households, and Discussions With General Practitioners on Personal Electronic Health Readiness: Cluster Factorial Quasi-Randomized Control Trial

Philip Abbott-Garner[1], BSc, PhD; Janet Richardson[1], PhD; Ray B Jones[1], PhD

School of Nursing and Midwifery, University of Plymouth, Plymouth, United Kingdom

Corresponding Author:
Ray B Jones, PhD
School of Nursing and Midwifery, University of Plymouth
Rolle Building, University of Plymouth
Drake Circus
Plymouth,
United Kingdom
Phone: 44 1752661577
Email: ray.jones@plymouth.ac.uk

Abstract

Background: Electronic health (eHealth) may improve health outcomes, but many people remain digitally excluded. Personal readiness to use the internet for health may be limited by lack of internet infrastructure, personal skills, social support, service provision, and cost. The impact of interventions to reduce these barriers is unknown. From 2011, the British Government supported the implementation of "superfast" broadband (*Superfast*) across the rural county of Cornwall. This provided the opportunity to assess the impact of interventions at regional, practice, and household levels.

Objective: This study aimed to assess the impact of 3 interventions on personal eHealth readiness: (1) regional-level implementation of Superfast, (2) practice-level discussions with general practitioners to encourage greater internet use in health service provision, and (3) household-level tailored booklets providing information to help improve personal skills in eHealth.

Methods: This was a cluster quasi-randomized factorial controlled trial. Implementation of *Superfast* was monitored, and postcodes were classified as having early or late availability. An algorithm selected 78 from 16,385 eligible postcodes to minimize the possibility of overlap between general practices and ensure a balance of urban and rural areas; 1388 households were randomly selected from the 78 postcodes and allocated to the 8 (2 × 2 × 2) study arms. A modified version of the Personal eHealth Readiness Questionnaire was used to compare scores (0 to 10) and 4 components (personal, provision, support, and economic) from baseline (August 2013) to the 18-month follow-up between the 8 arms, to assess the impact of interventions. We compared SDs of scores to assess changes in eHealth inequalities.

Results: eHealth readiness improved over 18 months from 4.36 out of 10 to 4.59 out of 10 (t_{235}=4.18; P<.001; CI=0.13 to 0.35), resulting from increases in personal and provision components of the score (t_{255}=3.191; P=.002 and t_{258}=3.410; P=.001). However, there were no significant differences between the 3 interventions, either singly or in combination using intention-to-treat analysis. The proportion of internet users did not significantly increase (79.2%, 205/259 to 81.5%, 211/259) and mobile use was significantly greater (50.5%, 101/199 to 64.8%, 129/199). There was no change in eHealth inequality.

Conclusions: People in Cornwall became more ready to adopt eHealth services, increasing both their personal ability to use eHealth and their methods of access. The implementation of *Superfast* may have contributed to this; we are certain that our other 2 interventions did not. This increased eHealth readiness did not cause a larger digital divide. The study illustrates the complexity of conducting a randomized controlled trial to assess the impact of interventions at regional, practice, and household levels. Our method may be of use to others.

Trial Registration: ClinicalTrials.gov NCT00102401; https://clinicaltrials.gov/ct2/show/NCT02355808 (Archived by WebCite at http://www.webcitation.org/75oEz0E1x)

(J Med Internet Res 2019;21(3):e11386) doi:10.2196/11386

XSL•FO
RenderX

KEYWORDS

eHealth; randomized controlled trial; digital divide; broadband implementation; eHealth readiness; eHealth inequalities; tailored booklet; cluster trial

Introduction

Setting

Cornwall, a county in south west England, is a rural area with a population density of 1.5 persons per hectare versus the average of 4.1 persons per hectare in England [1]. It has a dispersed settlement pattern of numerous towns, villages, and hamlets; 27% of the population lives in urban areas, 29% in towns and larger villages, and 44% elsewhere [2]. Access to health care, transport, employment, information and communication technology (ICT), training, community facilities, and services such as shops and schools is a problem. Cornwall's population is older than the national average (29.7% aged older than 60 years compared with 22.3% nationally [1]). Before the Superfast Cornwall project (discussed below), internet infrastructure was poor. Maximum download speed averaged 5 to 6 Mbps in urban areas [3], and some *not spot areas* had no internet access. Internet reliability was poor, meaning access could often fluctuate during the day.

Superfast Cornwall Project

This program funded by the European Union, British Telecom (BT), and Cornwall Council aimed to provide superfast broadband (*Superfast)* infrastructure to Cornwall and the Isles of Scilly. The program ran from 2011 to 2015, during this duration, fiber optic broadband had been introduced to 95% of homes and businesses [3]. *Superfast*, defined as an infrastructure capable of delivering internet speeds higher than 24 Mbps [4], aimed to provide a faster and more reliable service with speeds of up to 330 Mbps. Introducing *Superfast* was a significant engineering task costing approximately £132 million and requiring the installation of 130,000 km of fiber optic cable [3].

Benefits of Electronic Health and Digital Divide

Systematic reviews have shown the potential positive impact of electronic health (eHealth) [5,6] in areas such as the management of long-term conditions [7-11], internet-delivered cognitive behavioral therapy [12-14], smoking cessation [15,16], and cost reductions [17-22]. Some are concerned that as we introduce more eHealth, the *digital divide* will increase. Although the proportion of nonusers had declined from 35% in 2003, 22% of the British population had still not used the internet in 2013 [5]. Age remained the biggest predictor of nonuse; in 2011, only 33% of those aged 65 years and older in the United Kingdom used the internet. People who stay offline have reduced opportunities [23], and this divide could increase with the implementation of *Digital First* across the National Health Service [24].

Barriers to Electronic Health

Differential access to information and computer technologies can be examined at the personal level [25], categorizing barriers as (1) provision (including the impact of lack of suitable infrastructure), (2) personal, (3) interpersonal, and (4) economic.

Provision Barrier

Poor internet access is a barrier to eHealth use [26-29]; in 2014, the average broadband speed in some rural areas was 5 Mbps compared with 27 Mbps in urban areas [30]. Slow-speed internet obviously compromises viewing of Web-based videos and images [31,32]. Variation is not just caused by hardware but can result from differences by geography or patient group in NHS services; for example, most renal patients in the United Kingdom have had access to their Web-based renal medical record for many years [33] but few, if any stroke patients had such access [34]. Video consultations had been used for dermatology [22,35] but not widely adopted in general practice.

Personal Barrier

Physical and psychological attributes can also be barriers, such as lack of ICT skills [36-39], distrust of internet [40] or health information it provides [41-43], and lack of motivation to access eHealth services [44,45]. Someone's current health may increase motivation to use the internet for health information [46,47], even as it limits their ability to do so [48]. Although video use is increasing, much internet health information is text based, meaning low-literacy populations can struggle to use information effectively [26,28,49,50].

Interpersonal (Social) Barrier

Some factors limiting eHealth use may be moderated if people have social support [34]. Many nonusers have some form of indirect access to the internet via other individuals (proxy users). In the United Kingdom, in 2013, approximately 70% of nonusers reported having access to a proxy user but only 20% actually used them to access the internet [51]. Nonusers who do not have access to or choose not to use the internet may lack a strong support structure to help them overcome fears and apprehension [52]. With decreased social connection, some may also lack exposure to the internet and other technologies [53]. Furthermore, they may not perceive the usefulness in adopting internet use or have limited motivation to do so [54,55].

Economic Barrier

UK national figures indicated that lower-income households were less likely to access the internet [56]. Although homes may be technically capable of internet connection, families may not be able to afford it; someone relying on accessing the internet at their local library may be restricted by transport costs [34]. Women diagnosed with breast cancer were less likely to use the internet for health if they had a lower income, even after controlling for other predictors [57]. Lung cancer patients with higher income were more likely to seek Web-based health information about their condition [58].

Measuring Electronic Health Readiness

The degree to which people are prepared and able to use eHealth can be termed "eHealth readiness" [59]. eHealth readiness has been approached in various ways with some focusing on the readiness of a whole sector or system; Legare et al [60]

XSL•FO
RenderX

identified 6 different assessment tools [61-66] for this approach. Others have assessed the eHealth literacy of individuals, for example, the eHealth literacy scale (eHEALS) [67]. Jones [34] took a pragmatic compromise examining eHealth readiness of individuals, but including in this their opportunities from infrastructure, economics, and social support. The Personal eHealth Readiness Questionnaire (PERQ) [68] was designed to measure the impact of interventions that aimed to improve eHealth readiness and reduce eHealth inequalities. PERQ uses a similar approach to eHEALS, adopting the use of scales as opposed to a binary measure, but included further variables to cover the full range of individuals from noninternet users to frequent internet users. PERQ has 4 subcomponents: provision, personal, support, and economic.

Measuring Electronic Health Inequalities

As older people have lower use of ICT, some observers assume that the digital divide will disappear with newer generations [29,69]. However, reduced ability to adopt new technology with age may continue [29]. Economic barriers may remain if ICT costs are too high for future generations of older adults. In addition to the ethical argument for addressing eHealth inequalities, such inequalities make the adoption of more cost-effective health delivery difficult if both eHealth and more traditional services must be provided [34]. We need, therefore, to develop interventions that help reduce eHealth inequalities and have a way of measuring them. The SD of the PERQ eHealth readiness score provides a measure of eHealth inequalities.

Assessing the Impact of Superfast on Electronic Health Readiness

Although poor internet infrastructure is recognized as a barrier to eHealth, there was no clear evidence that improving internet infrastructure alone is enough to improve uptake of eHealth services. A simple before-after comparison does not allow the attribution of likely improvement to the infrastructure change without some form of *control group*. However, many would argue that an infrastructure change on its own is unlikely to radically improve uptake of eHealth but that some form of education, awareness raising, behavioral and organizational change is also needed. The implementation of *Superfast* provided an opportunity to assess the impact of an improved internet infrastructure alongside interventions at practice and household levels.

Assessing the Impact of a Combination of Interventions on Electronic Health Readiness

Individuals possess different levels of eHealth readiness and, as discussed, may experience a wide range of separate and shared barriers. Barriers to the implementation of eHealth exist at multiple levels: (1) individual level, (2) clinician or service level, and (3) regional infrastructure level. It is unlikely that a single standardized intervention will be effective across these levels. Rather, it is likely that a combination of interventions, targeted at multiple levels, would prove most effective at reducing eHealth inequalities.

The *Superfast* project not only allowed for the impact assessment of an infrastructural change but also provided the opportunity to identify and assess the effectiveness of other interventions targeted at the personal and service levels. These interventions were designed to increase eHealth use, both singly and in combination. With limited resources, we sought to assess the impact of this infrastructure change in combination with individual- and provider-level interventions.

Methods

Design

A cluster, quasi-randomized, factorial (2 × 2 × 2) controlled trial design was used to examine the impact of 3 interventions: (1) regional-level improvement of physical infrastructure (*Superfast*), (2) practice-level discussions with general practitioners (GPs) to encourage greater use of the internet in health service provision, and (3) household-level tailored booklets (TBs) providing information to help improve personal skills in eHealth. Households within Cornwall were allocated to each of the 8 arms of the study. eHealth readiness and inequality were compared pre-and postintervention to measure the impact (singly and in combination) of each of the 3 interventions.

The study was approved by the Plymouth University Faculty of Health and Human Sciences Ethical Committee and obtained local research and development approval from the Royal Cornwall Shared Research Management Service. The trial was registered at the US National Institutes of Health (ClinicalTrials.gov) # NCT02355808 on April 2, 2015.

Sampling and Randomization

The initial sampling unit was the postcode. All 20,088 postcodes in Cornwall (excluding the Isles of Scilly for practical reasons) were included; 2958 listed as having a population of zero and postcodes without any population data were excluded. To more clearly define the presence or absence of *Superfast*, we excluded 745 postcodes with *Superfast* coverage of between 0% and 49% as these were in the process of receiving *Superfast* at the time of sampling. The remaining 16,385 postcodes, therefore, either had *Superfast* available or did not.

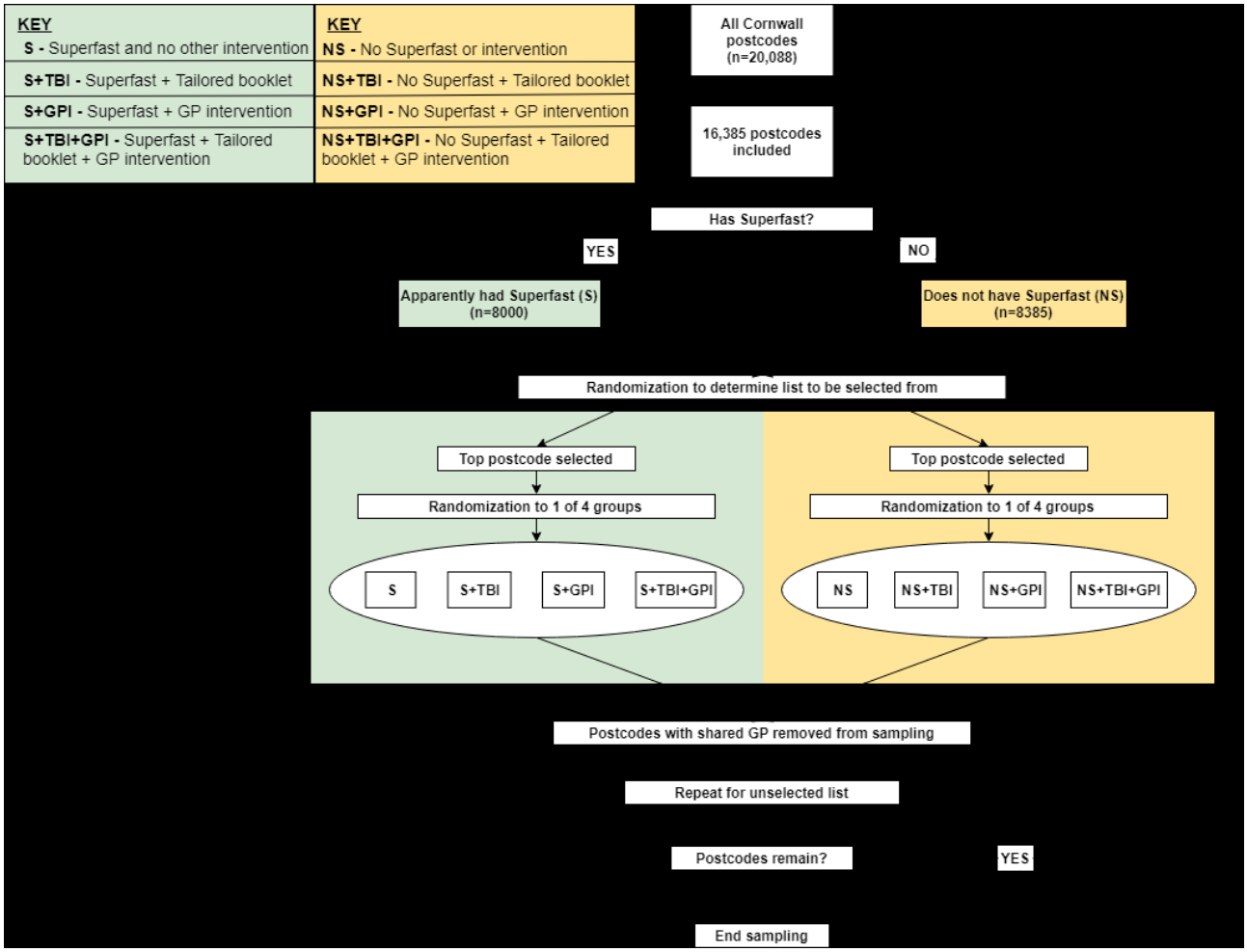
Providing an intervention at the primary care level via GP practices introduced the likelihood of contamination between intervention groups. GP practices often serve a large geographical area; any intervention at this level would affect several postcode clusters. This meant that random selection of postcodes, without accounting for the intervention area, would likely allocate postcodes with shared practices to separate intervention groups. The sampling method sought to reduce the likelihood of contamination by eliminating postcode clusters at the practice level.

GP practices were included based on longitude and latitude data from NHS choices [70]. GPs in Cornwall and those in Devon on the Cornish border, who were the closest GP to a Cornish postcode, were included.

We designed a method to (1) reduce potential contamination between the 8 arms of the study, (2) account for the rollout of *Superfast*, and (3) ensure similar allocation of urban and rural areas.

XSL•FO
RenderX

Figure 1. Sampling and randomization method needed to identify and randomly allocate general practice intervention within Superfast intervention. GP: general practitioner.



The 16,385 postcodes that met the initial inclusion criteria were allocated to 2 separate lists based on their *Superfast* coverage and sorted based on population (highest population at top):

- Has *Superfast* (S): postcodes with coverage ≥50% (n=8000)
- Did not have *Superfast* (NS): postcodes with coverage of 0% (n=8385).

Postcodes within these lists were randomized to intervention groups:

- No further intervention
- General practice intervention (GPI)
- Tailored booklet intervention (TBI)
- Tailored Booklet and GPI (TB+GPI).

The following process of selection and randomization (using Excel random number generation) then took place until no postcodes remained (Figure 1):

1. A randomization took place to identify which list (S, NS) would be selected first.
2. The first postcode (highest population) from the list (S, NS) was selected and was randomly allocated to either the one of 4 groups (S, S+TBI, S+GPI, S+TBI+GPI) or to the other 4 groups (NS, NS+TBI, NS+GPI, NS+TBI+GPI). There were 8 groups in total for 3 interventions (2 × 2 × 2).
3. Any postcode that shared the same geographically closest GP practice as the selected postcode was then eliminated.
4. The remaining top postcode on the second list was then selected and randomly allocated to 1 of the 4 groups.
5. This process was repeated from step 1 until no postcodes remained on either list.

Through this process, 78 from 16,385 postcodes were selected and randomly allocated to 1 of the 4 intervention groups within their level of *Superfast* coverage. Using Zoopla [68], a website giving estimated house values across the United Kingdom, all households within the postcode were listed and 18 randomly selected, using Excel number generation, and included within the study. In postcodes with less than 18 households, all households were included in the sample. The final sample consisted of 1388 households from 78 postcodes served by 78 different GP practices.

Sample Power

Most limitations of sample size were imposed by the infrastructure intervention (Superfast arm), which was limited to the county of Cornwall. This then limited the number of GP practices that could be allocated without *contamination* between the randomization arms. The number of households found in rural postcodes limited the number of households. A sample size calculation based on the desired magnitude of effect was, therefore, not conducted. Instead, a calculation was made to

XSL•FO
RenderX

estimate the possible magnitude of effect that could be found with 80% power.

With an assumed response rate of 50%, it was estimated that the smallest effect size that could be found between the 2 arms of *Superfast* ("has" and "does not have") was 0.52, assuming 80% power and 95% significance. The smallest effect size that could be found between each of the 8 arms of the study was 1.05, with 80% power and 95% significance.

Outcome Measures

A before versus after assessment of eHealth readiness using the PERQ [34] was conducted on households within the sample over an 18-month period. The PERQ (Multimedia Appendix 1) was modified slightly by improving the wording and layout based on recommendations in the original paper [34]. The 4 subcomponents were combined to create an overall eHealth readiness score (0 to 9). The SD of readiness scores was taken to represent eHealth inequality.

Interventions

Regional: Implementation of Superfast (Had Superfast/Did Not Have Superfast)

Before the implementation of *Superfast*, households were likely to have had internet connectivity ranging from none (not-spots) to maximum speeds of 5 to 6 Mbps. After implementation of Superfast, Cornwall reported that 95% (241,000) premises had *Superfast*, with nearly 90% able to connect at speeds of over 24 Mbps [3]. It was not possible for the study to allocate postcodes to receive or not receive *Superfast*. This process was dependent on the Superfast Cornwall timescale for the rollout; therefore, this arm of the study was a *natural experiment*. Clusters were categorized into areas with or without *Superfast*, based on the rollout at the time of sampling.

Practice: General Practice Intervention

The aim of GPI was to engage selected practices to encourage GPs (1) to adopt more eHealth services and (2) to actively promote the existing services to their patients and aid them in adopting such services. The hypothesis was that achieving these outcomes should impact patients within the area, resulting in increased eHealth readiness. With this intervention:

1. The researcher contacted (by post) selected practices in September to October 2014 to arrange meetings. This letter explained the project and sought permission to attend practice meetings to discuss their use of eHealth services. If this was not possible, the researcher tried to meet with a practice member or establish an email conversation.
2. GPs were given suggestions as to how they might expand their current use of eHealth services to use additional eHealth services or better promote their existing services, using examples of GPs in their area or nationally.
3. GPs were also asked to comment on the services they offered, perceived benefit or detriment, and ease of adoption.

Meetings were conducted for 15 min and covered 6 topics: Web-based appointment booking; Web-based repeat prescriptions, Web-based access to medical records, information prescription, phone triage, and video consultations. Meetings were tailored to consider the current services provided by the GP practice; if discussed services were currently implemented, the conversation would focus on the difficulty the GP experienced to implement and any perceived benefits or limitations of the system. We included 39 GP practices in the GPI arm.

Household: Tailored Booklet

Participants randomized to the booklet intervention received a tailored eHealth information booklet by post. A total of 16 A5 pages were created using information from national and local services. Some pages were included for all participants; other pages were based on responses to the PERQ. Creation of TBs used a decision tree to identify which A5 pages to include (Multimedia Appendix 2).

This booklet was addressed to the individual who completed and returned the survey. This process identified individual needs and then tailored a booklet to address those needs. For example, a noninternet user reporting that they would use the internet more for health if they could get someone to help them received a booklet showing resources such as UK Web-based centers. On the other hand, someone who reported that they lacked confidence in using the internet received information about Web-based internet training, such as Learn My Way [71].

As a cluster trial, all households in the intervention postcodes (clusters) received an eHealth information booklet; those households not randomly selected to complete the PERQ received a general rather than personalized booklet. Tailoring for these households could only use geographical data, for example, showing a person what is available in their area based on their postcode. In this case, booklets were addressed to the household as opposed to an individual.

Data Analysis

Data were entered into IBM SPSS version 23 for analysis. The main dependent variable (primary research question) under investigation was eHealth readiness in the form of a continuous variable, calculated from PERQ responses. Analysis was conducted using parametric tests to analyze eHealth readiness and the 4 subvariables that contributed to its calculation. On matched data, paired *t* tests were used to compare baseline with follow-up. To examine differences between groups, independent *t* tests were conducted on the change of continuous variables. Finally, a univariate general linear model was used to investigate the main effect of the 3 intervention conditions, added as fixed effects, on the change in eHealth readiness.

Some secondary analyses were conducted. The PERQ contained several categorical response questions that were relevant to the study and provided insight into eHealth behavior. For categorical data, nonparametric tests in the form of chi-squared tests for independent samples and McNemar, for paired data, were conducted. To provide further insight, in some cases, additional categorical variables were created from continuous variables to analyze proportions, for example, increased, decreased, no change.

XSL•FO
RenderX

Results

Response Rate

Of the 1388 households surveyed, 394 (28.4%) responded to the baseline PERQ (October 2013). At follow-up (March 2015), 259 households, 65.6% of original responders, replied to the PERQ (Figure 2).

Regional: Implementation of Superfast

With the Superfast Cornwall project completed, more accurate rollout data were released to the researcher. These data contained precise *go live* dates for all clusters (postcodes) included in the study, allowing households to be categorized by the number of months *Superfast* had been available in their area. Households in areas where *Superfast* had been available for more than 24 months at follow-up were categorized as "early receivers." Areas that had *Superfast* for 23 or less months at follow-up were categorized as "late receivers."

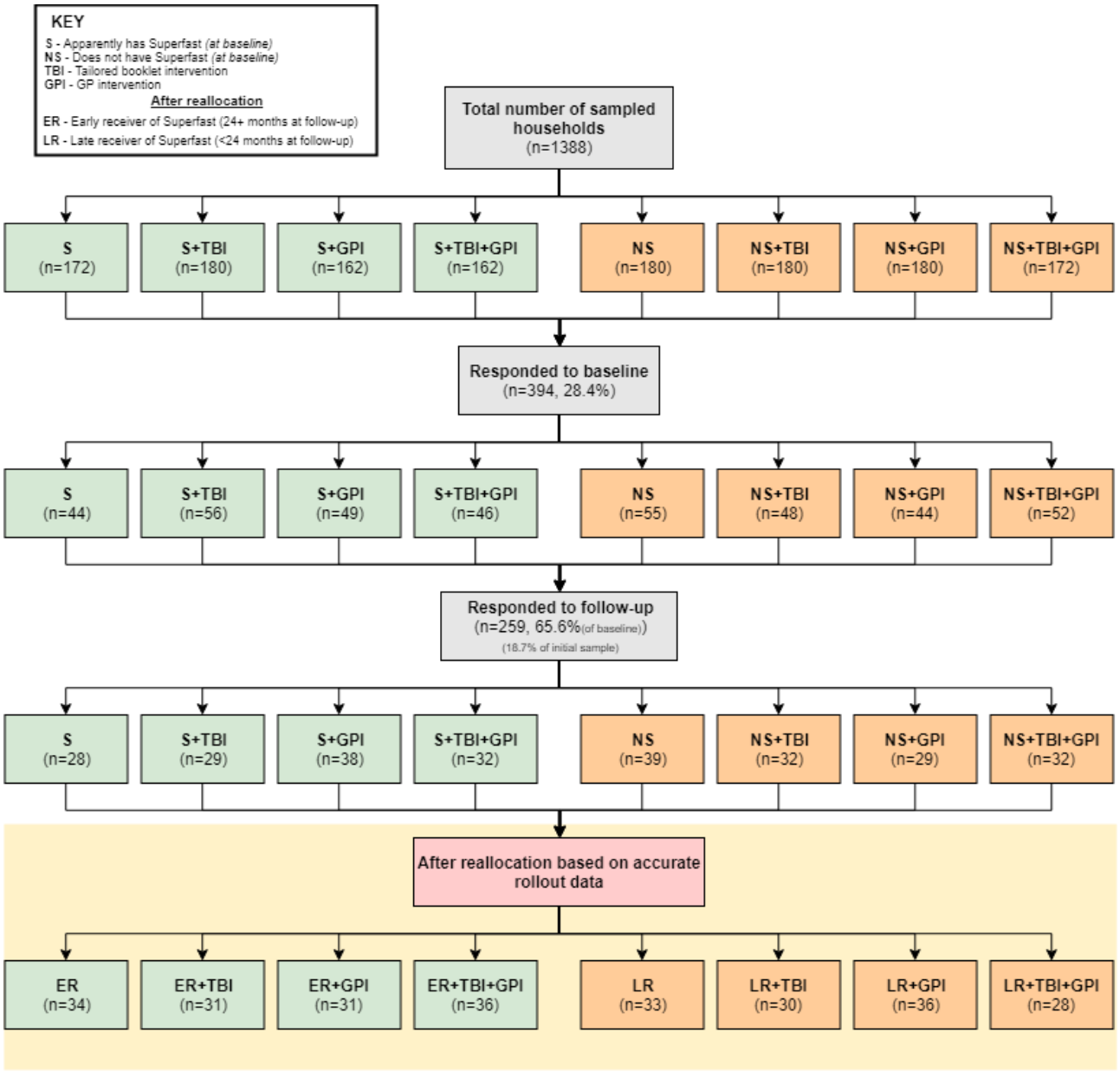
There was no significant difference between the change in readiness (0.26 vs 0.21; P=.66) or provision scores (0.16 vs 0.23; P=.53) or between the proportion of households increasing in readiness scores between early and late receivers.

There was no difference in the perception of speed within households between baseline and follow-up (McNemar=2.46; P=.25). However, changes in speed perception did significantly differ between *Superfast* arms, with 12 households (14.8%, 12/81) from "late receivers" and 5 (5.2%, 5/96) from "early receivers" reporting faster internet (χ^2_1=4.7, P=.03).

Practice: General Practice Intervention

Of the 38 GPs contacted to take part in the study, 8 (21%) agreed to take part, 3 (8%) refused due to busy schedules, and the majority (71%, 27/38) did not respond. The researcher attended 5 face-to-face meetings and had email correspondence with the remaining 3 GPs.

Figure 2. Consolidated Standards of Reporting Trials diagram of trial numbers for matched households showing early and late receivers of Superfast.



XSL•FO
RenderX

The GPI had no effect on household eHealth readiness, neither when considered as mean score (mean=0.18 vs mean=0.29; t_{234}=−1.01; *P*=.31; CI=−0.34 to 0.11) nor when considered as proportion of households increasing in readiness scores (32.5%, 38/117 vs 34.5%, 41/119; (χ^2_2=.6, *P*=.74).

Overall, 18.7% (38/203) of respondents across all arms, had been given information to help them use the internet for their health by a nurse, doctor, or another health care professional, but there was no difference between those in the GPI arm and others (18.3%, 19/104 vs 19.2%, 19/99; (χ^2_1=.03, *P*=.87).

As many GP practices did not take part in the study, we did an *as treated* analysis comparing households from GPs who had agreed to the intervention with other households, but there was still no difference. We also counted the number of practices offering Web-based access to medical records. Only 6 GPs within Cornwall had started to offer Web-based access to medical records at follow-up, previously none had offered this facility; however, there was no difference between those in the GPI versus others.

Household: Tailored Booklet Intervention

There was no significant effect of the booklet intervention on the change of readiness scores (t_{234}=−.106; *P*=.92). The proportion of households increasing in readiness scores was 33.5% (79/236) overall with no difference between those receiving (36.5%, 42/115) and not receiving the booklet intervention (37/121, 30.6%; (χ^2_2=1.2, *P*=.56).

The PERQ calculates a separate skill score based on responders' reported self-ability to complete 6 internet-related tasks. Overall, 32% (82/259) showed an increase in skills scores, but there was no difference between those who received a booklet compared with nonreceivers (33.6%, 42/125 vs 30.6%, 41/134; (χ^2_2=1.6, *P*=.46).

One area of the booklet focused specifically on the eHealth services offered by local GPs' websites to attempt to increase knowledge and use of these services. At baseline, a total of 54 (54/204, 26.5%) households reported that they "Didn't Know" if their local GP had a website; of these, 27 had become aware of their local GPs' website and the services it offered, but there was no difference between those receiving or not receiving the booklet (51.9%, 14/27 vs 48.1%, 13/27; χ^2=0.09, df=2; *P*=.96; χ^2_2=.09, *P*=.96).

Only 5 internet-using households (5.2%, 5/97) who had received the booklet acknowledged receiving "a booklet in the post regarding using the internet for health."

Interventions in Combination

A univariate general linear model was used to investigate the main and combined effect of the 3 interventions (*Superfast*, GP, booklet), added as fixed effects, on the change in eHealth readiness. A full-factorial interaction effect was also examined between *Superfast* × booklet × GP for the outcome of change in readiness. The model showed no significant main effect of either *Superfast* (*P*=.677), GP (*P*=.237), or booklet (*P*=.928) on the change in readiness scores.

Change in Internet Use

The proportion of internet users (79.2%, 205/259) at baseline did not significantly increase, being 81.5% (211/259) at 18-month follow-up. A fifth of respondents who reported that they had not used the internet at baseline (20.4%, 11/54) reported that they had used the internet at follow-up. Only 5 internet users at baseline (2.4%, 5/205) reported not to have used the internet in the previous 3 months at follow-up.

More households at follow-up had used their smartphones or mobile devices to access the internet compared with baseline (64.8%, 129/199 vs 50.5%,101/199; *P*<.001; Figure 3). A total of 34 internet users who had never used a mobile device to access the internet at baseline reported using a mobile device for internet access at follow-up. Only 6 households reporting that they had stopped using a mobile device for internet access.

Health-Related Internet Use

Most internet-using households had used the internet for at least one health-related activity; there was no significant difference between baseline (70.2%, 144/205) and follow-up (68.7%, 145/211; *P*=.29) for health-related internet use (Figure 4). However, at follow-up, internet users showed a significant increase in self-reported confidence in using the internet for health-related tasks (mean=7.39 vs mean=7.78; t_{197}=2.88; *P*=.004; CI=0.12 to 0.65).

No significant differences were found for the uses of the internet for health-related tasks between baseline and follow-up for internet-using households. The most common health-related activity was using a search engine to search for health topics (67% vs 66%), followed by using email for health (11% vs 12%) and discussing health topics on a forum (6% vs 6%).

Use of social media for health remained low; only 6 households (3%) at baseline and 4 at follow-up (2%) reported use.

Self-Perceived Barriers

Just over half (54.0%, 128/237) reported that they "have or would use the internet for health and have no real barriers to that use." There was no change at follow-up (58.6%, 139/237). The most common reported barrier at follow-up was "No need for health information" (11.0%, 27/246) and "I have no interest in using the internet" (11.0%, 27/246). Only 3 households (1.2%, 3/246) reported that they "Would use the internet more for health if I could get a good internet connection."

XSL•FO
RenderX

Figure 3. Where and how 211 internet-using households had accessed the internet in the last 3 months at baseline (October 2013) and follow-up (March 2015).

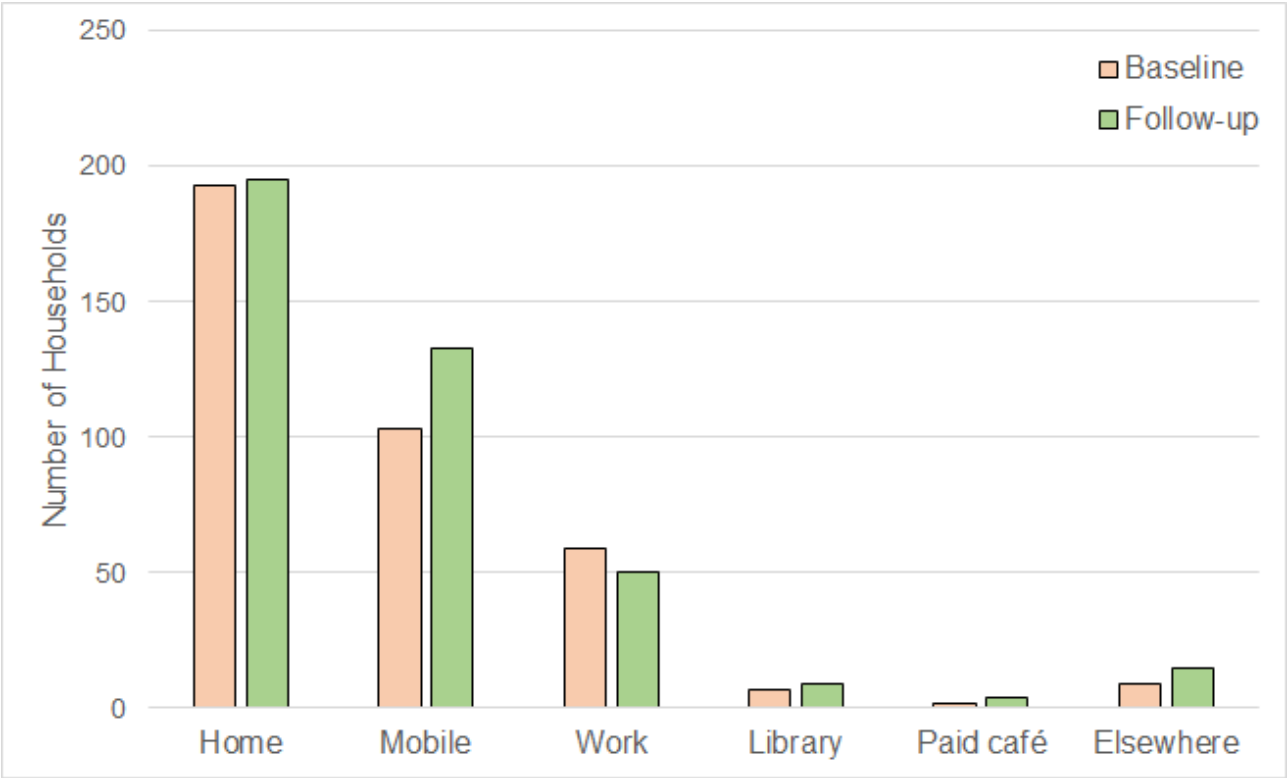
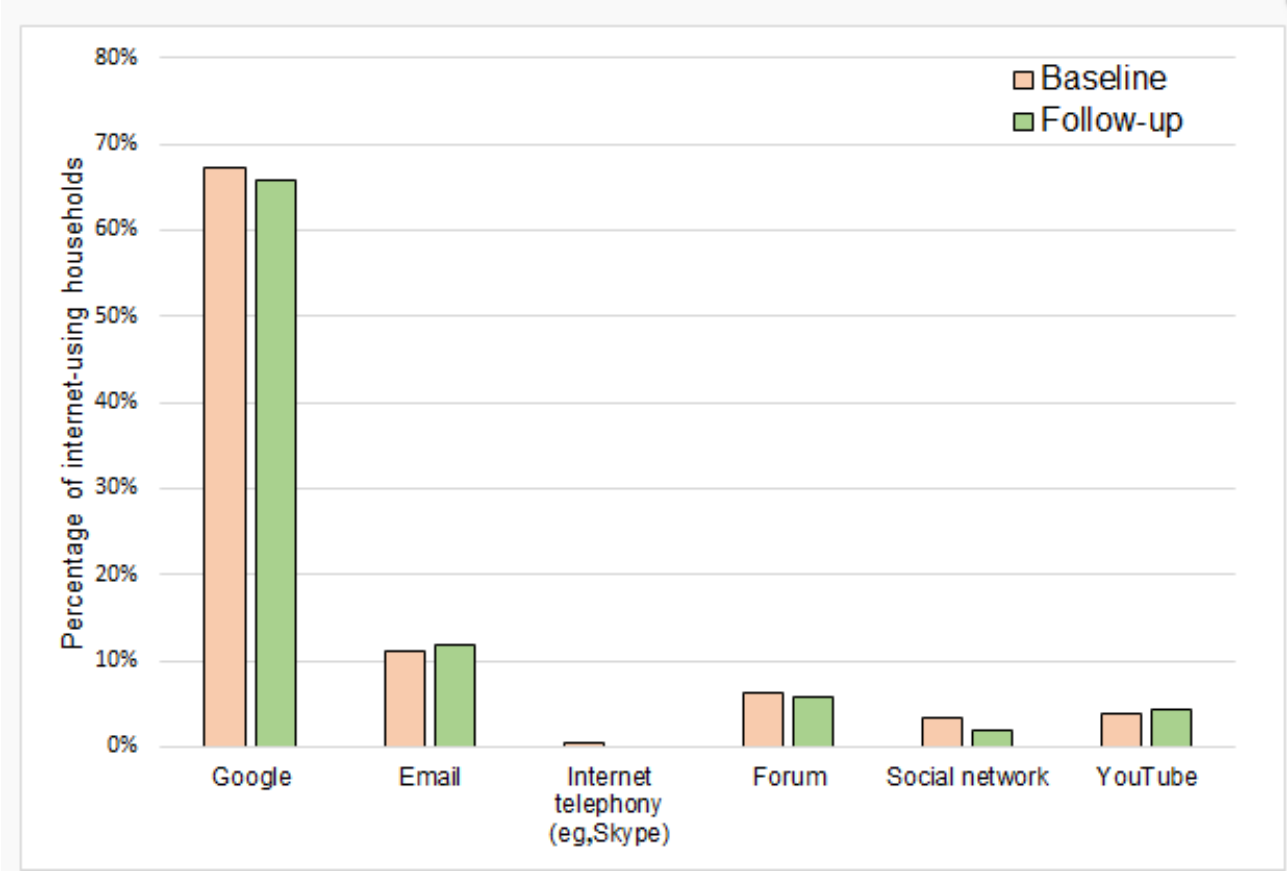


Figure 4. How internet-using households had used the internet for health-related activities.



XSL•FO
RenderX

Table 1. Summary of electronic health readiness and the 4 subvariables on matched households.

Variable	Baseline		Follow-up		Change	
	Mean	SD	Mean	SD	Mean	SD
Electronic health readiness	4.36	1.72	4.59	1.78	*0.23* [a]	0.06
Personal	5.49	2.92	5.77	2.94	*0.28* [a]	0.02
Provision	4.06	1.70	4.26	1.78	*0.20* [a]	0.08
Support	1.89	1.87	1.88	1.80	−0.01	−0.07
Economic	1.57	0.90	1.67	0.88	0.10	−0.02

[a]Values in italics have $P<.01$.

Electronic Health Readiness

Of the 236 households with complete data, half (51.3%, 121/236) showed no change in their eHealth readiness score, a third (33.5%, 79/236) showed an increase in their eHealth readiness score (maximum increase 3), and 36 (15.3%, 36/236) showed a decrease in their eHealth readiness score (maximum decrease of −3).

Overall eHealth readiness scores increased significantly from baseline to follow-up for these 236 households (mean=4.36 vs mean=4.59; t_{235}=4.18; $P<.001$; CI=0.13 to 0.35). The SD of readiness (eHealth inequalities) among responders remained similar (1.72 vs 1.78). Analyses of the 4 subvariables that contribute to the calculation of eHealth readiness scores indicated that both personal and provision subvariables increased over the 18 months (t_{255}=3.191; P=.002 and t_{258}=3.410; P=.001), whereas economic (P=.12) and support (P=.97) subvariables showed no significant change (Table 1).

As might be expected, new internet users had higher increases in their readiness scores compared with continued users (1.56 vs 0.26; t_{197}=−4.76; $P \leq .001$).

Previous users who had stopped using the internet (new nonusers) showed the biggest decreases in their readiness score, with an average reduction of 1.75. These 5 households had significantly lower readiness scores at baseline (3.00 vs 5.04; t_{197}=3.78; $P<.001$) than the 195 who were *continuedusers*.

As new adopters of the internet showed the largest increase in readiness scores, potentially these households alone may have been responsible for the sample increase in readiness scores. To investigate this, further analysis was conducted on *continuedusers*, excluding new internet users; this showed significant increases in readiness scores (mean=5.04 vs mean=5.30; t_{189}=4.57; $P<.001$; CI=0.15 to 0.38).

Discussion

Overall Impact

No one has previously examined the impact of concurrent improvements in internet infrastructure alongside person-based interventions. We assessed the impact of such interventions on personal eHealth readiness via a cluster quasi-randomized factorial controlled trial. Although eHealth readiness increased over the course of the study, this change could not be explained by the interventions, either singly or in combination. This could be because there really was no improvement or that the questionnaire approach we used was not sensitive to the change.

Our single booklet posted to a house and short limited discussions with a few practices was very “low dose.” The implementation of *Superfast* could potentially have more impact, but it relies on uptake. As a pragmatic randomized controlled trial, our analysis was *intention-to-treat*; thus, low uptake in all 3 arms could *swamp* a possible improvement that might be seen in an *as treated* analysis.

The use of eHealth relies on 3 separate but supporting conditions: (1) the personal ability to use it, (2) the presence of systems to provide it, and (3) the infrastructure available to support it. For this reason, despite finding no positive impact of any of the interventions, we argue that this study demonstrated a possible method to explore the impact of infrastructure improvements alongside complementary interventions.

Changes Across Cornwall in Electronic Health Readiness and Electronic Health Inequalities

A third of our household respondents improved their eHealth readiness over 18 months of study, and overall, the mean eHealth readiness had improved without any increase in eHealth inequalities. There was no evidence that it was the already “eHealth ready” becoming further advantaged over the unconnected; people from across the whole “scale” had shown improvements.

The increased eHealth readiness reflected the increase in the proportion of internet users (79% vs 82%) in line with reported change for the United Kingdom [72]. Only 1 in 5 nonusers at baseline had started to use the internet, whereas only 5 people (<3%) had stopped using the internet at follow-up. But the overall increase in eHealth readiness was not solely due to new internet users. When new users were excluded from the analysis, the increase in readiness was still significant. This suggests that existing internet users became more ready to use the internet for health.

Despite the increase in the level of readiness to use eHealth, the types of use remained the same. Using a search engine to find health information was the most frequent activity. Although others have proposed that social media could be used more for health [73], few people in this study used it to obtain health information or contact health care professionals or organizations.

XSL•FO
RenderX

Implementation of Superfast

Our *intention-to-treat* analysis examined the impact of *Superfast* regardless of whether a household adopted the service by upgrading their internet supply. On this basis, there was no improvement in readiness scores between early and late receivers of *Superfast* during the 18 months of study. However, at our study's follow-up date, BT-estimated uptake of *Superfast* across Cornwall was still quite low (28%) despite *Superfast* being available for most of Cornwall. It was not possible to obtain data on which households in our sample had adopted *Superfast;* therefore, it was not possible to conduct an *as treated* analysis, that is, we cannot tell if households who adopted *Superfast* had higher readiness scores.

Analysis of categorical responses suggests more subtle changes. More late receivers than early receivers were happy with their internet speed at follow-up. As early receivers had become accustomed to the greater speed or bandwidth, both their internet use and expectations of broadband provision are likely to have increased.

Slow uptake of *Superfast* is also seen nationally; Ofcom [74] reported that by the end of 2015, only 42% of UK households had taken up offered *Superfast* services. This uptake was higher than Cornwall's uptake, but *Superfast* had been available in other parts of the United Kingdom for longer. A total of 38% thought cost was a concern, which is quite a high proportion of people might be therefore put off by switching to superfast. We do not know if these people remained with slower speed access rather than take up superfast because of cost.

Only 3 respondents said they would use eHealth more if they could get a good internet connection, suggesting that very few perceived their current internet speed as a barrier to eHealth use. However, at the time of this study, no Web-based health services available in Cornwall required a S*uperfast* connection. The most common eHealth activity was using Google, and most of the Web-based health information was in simple text and picture format not requiring a high-speed connection. At the time of our data collection, very few (<5%) reported using health services requiring higher speeds, such as YouTube. Nationally, in 2016, the NHS YouTube channel had less than 26,000 subscribers. Moreover, even video streaming only required modest speed connections of around 500 Kbps [75].

Superfast rollout seemed to have had a measurable impact on people's perception of internet speed, but this did not translate to measurable increases of eHealth readiness. On the other hand, Cornwall is now structurally more ready to adopt eHealth services such as video calling.

Tailored Household Booklet

Our study demonstrated how TBs could be produced based on questionnaire and geography. For example, for internet nonusers, it provided guides on how to get started. For internet users, it provided information on how and where they could find and use Web-based health information.

Booklets included information about Web-based services provided by recipients' local GP.

However, this 1 booklet delivered by post and then asked about some months later was not effective. The recollection of receiving a booklet was extremely low, with only 5 households reporting they had received a booklet. The most likely explanation is that the booklet was perceived as junk mail and never read. The booklet design may be of use for other organizations such as Healthwatch or if given directly by a GP or practice nurse, but as a low-intensity intervention delivered by hand to the house, it was ineffective.

General Practice Intervention

The GPI was ineffective. Few GPs agreed to meet the lead author, and the discussions suggested that these GPs were knowledgeable about eHealth. For example, the researcher raised the potential use of a phone triage system, championed by some GPs [76], but many GPs responded mentioning a recent article in the *Lancet* [77] that had shown increased workloads. The likelihood, therefore, was that the researcher was no better informed than the GPs and, thus, was unlikely to raise awareness of new digital possibilities.

On the other hand, where the researcher attended larger practice meetings, it was apparent that the views and opinions of GPs differed drastically within the same surgery. In 1 meeting of note, the topic of information prescription sparked a huge debate over its usefulness. Several GPs within the practice were very positive toward information prescription, often directing patients to specific URLs with information on their condition and even printing out Web-based information for those who had limited access. Other GPs had very strong views against using information prescription, preferring that the patient spoke to them only and not use the internet. This discrepancy in GPs' attitude has been well documented [78] and highlights the continued inequalities of service provision.

The difficulty in recruiting GPs has been demonstrated in previous research [79]. The GPI was designed to prevent this, by being short in length with minimal requirements for GP participation, although this did not seem effective. The time of year may have prevented a higher participation rate, with many GPs citing a busy flu season impacting their availability. However, it is likely that GPs will always be busy [80] and reluctant to participate, without a keen indication of potential benefits.

Addressing barriers to the implementation of eHealth technology is a complex process that requires support from health services. It is important for policy makers and hospital or practice managers to understand the specific barriers that challenge the practicing GPs and design appropriate interventions to address barriers and promote facilitating factors [81]. Some barriers such as cost associated with the adoption and maintenance of eHealth technology may require incentive [82].

Limitations

The study suffered from a low response rate; responders were disproportionally female, older, and came from areas with higher estimated house values. As the *Superfast* rollout was outside the control of the researchers, it was not possible to randomize. Early receivers of *Superfast* were likely from areas in Cornwall that had existing internet infrastructure and were less rural.

XSL•FO
RenderX

Therefore, it is possible that late receivers contained a higher proportion of isolated households. The sampling method, designed to reduce the potential of contamination between arms of the study, removed postcodes with shared GP practices following selection. As the selection of postcodes was ordered by population, a limited measure or rurality, postcodes that shared a GP practice with another more populated postcode had a much lower chance of being selected. Although this approach was vital to reduce the high risk of contamination between intervention conditions, it again meant that highly rural postcodes may have less chance of being included. It is possible, therefore, that the eHealth readiness from this sample is overoptimistic compared with Cornwall as a whole.

Further Research

This study has provided 2 measurements of eHealth readiness within Cornwall over an 18-month period. There is potential to continue this study to provide a longitudinal view of the change in eHealth readiness over the coming years. A continued longitudinal study will provide insight into the change over time and allow for the impact of the *Superfast* rollout to be further assessed. As discussed, the actual uptake figures of *Superfast* are low, estimated at 28%, but these are expected to increase over the coming years. Continued measurement may show a continued increase in eHealth readiness as the uptake rates increase. Importantly, it will also allow for the inequalities in readiness to be monitored.

With the implementation of *Superfast* across the county, Cornwall has the potential to be a prime location for research into eHealth. The infrastructure improvement has made it possible for Cornwall to support highly demanding eHealth services such as video consultations or live streaming of health clinics. Presently, the county does not provide such systems, but now has the structural groundwork for research in this area. There is the potential for randomized trials of such services to be organized and conducted. Small trial projects might have to be conducted at the hospital level to show feasibility. This study will help show the potential benefits of such services, which may encourage innovations to be adopted more widely, such as at the GP level. In addition, future research in the area will provide further insight into the significant barriers toward eHealth use; with the *physical* speed barrier removed, other personal and organizational barriers are likely to be further highlighted. This will help researchers examine how to address those barriers and design effective interventions.

Conclusions

Over the 18-month study, households in Cornwall became more *eHealth ready*. It is possible that the rollout of *Superfast* contributed to this, but we were unable to show that definitively. It is unlikely that our other 2 interventions had any effect. The study illustrates the complexity of trying to assess such interventions by a randomized trial, and our methods for a cluster quasi-randomized factorial controlled trial may be of use for others.

Acknowledgments

This study was funded with the aid of a studentship from the Superfast Cornwall project, the EU, and Plymouth University. Superfast Cornwall, with the help of BT, provided rollout data for use within this study. The authors would like to thank the initial steering group, which included Dr Ellen Helsper, Professor Grant Blank, Katherine Stewart, Dr Andy Neville, and Peter Farrell. The authors would also like to thank Paul Hewson for statistical advice and Katie Edwards for her comments on the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The modified Personal eHealth Readiness Questionnaire used to measure eHealth Readiness, which was delivered to participants at baseline and follow-up.

[PDF File (Adobe PDF File), 241KB - jmir_v21i3e11386_app1.pdf]

Multimedia Appendix 2

Example of the page database used to construct tailored booklets, which were delivered to all households within the tailored booklet intervention arm of the study.

[PDF File (Adobe PDF File), 2MB - jmir_v21i3e11386_app2.pdf]

References

1. Office for National Statistics. 2011 Census: Aggregate data (England and Wales) URL: https://www.ons.gov.uk/census/2011census/2011censusdata [accessed 2018-06-19] [WebCite Cache ID 70I74Mixu]
2. Cornwall Council. 2011. Cornwall and Isles of Scilly Strategy to Reduce Health Inequalities Good Health and Well Being for Everyone 2011-2016 URL: https://www.cornwall.gov.uk/media/19802423/health-inequalities-2011-16-v2.pdf [accessed 2018-06-19] [WebCite Cache ID 70I7Dpg1L]

XSL•FO
RenderX

3. Superfast Cornwall. Superfast Cornwall: Our Story URL: https://www.superfastcornwall.org/wp-content/uploads/2018/04/Superfast-Cornwall-Our-Story.pdf [accessed 2018-06-19] [WebCite Cache ID 70I7yrNPw]
4. BDUK. GOV.UK. London: Department for culture, media, and sport; 2011 Sep. Broadband Delivery Programme: Delivery Model URL: https://www.gov.uk/government/publications/broadband-delivery-programme-delivery-model-september-2011 [accessed 2019-01-28] [WebCite Cache ID 70I8EZfJG]
5. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
6. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
7. Jaana M, Paré G. Home telemonitoring of patients with diabetes: a systematic assessment of observed effects. J Eval Clin Pract 2007 Apr;13(2):242-253. [doi: 10.1111/j.1365-2753.2006.00686.x] [Medline: 17378871]
8. Jackson CL, Bolen S, Brancati FL, Batts-Turner ML, Gary TL. A systematic review of interactive computer-assisted technology in diabetes care. Interactive information technology in diabetes care. J Gen Intern Med 2006 Feb;21(2):105-110 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.00310.x] [Medline: 16390512]
9. Azar M, Gabbay R. Web-based management of diabetes through glucose uploads: has the time come for telemedicine? Diabetes Res Clin Pract 2009 Jan;83(1):9-17. [doi: 10.1016/j.diabres.2008.09.055] [Medline: 19056140]
10. Linn A, Vervloet M, van Dikj L, Smit E, van Veert JC. Effects of eHealth interventions on medication adherence: a systematic review of the literature. J Med Internet Res 2011 Dec 5;13(4):e103 [FREE Full text] [doi: 10.2196/jmir.1738] [Medline: 22138112]
11. Eland-de Klok P, van Os-Medendorp H, Vergouwe-Meijer A, Bruijnzeel-Koomen C, Ros W. A systematic review of the effects of e-health on chronically ill patients. J Clin Nurs 2011 Nov;20(21-22):2997-3010. [doi: 10.1111/j.1365-2702.2011.03743.x] [Medline: 21707807]
12. Reger MA, Gahm GA. A meta-analysis of the effects of internet- and computer-based cognitive-behavioral treatments for anxiety. J Clin Psychol 2009 Jan;65(1):53-75. [doi: 10.1002/jclp.20536] [Medline: 19051274]
13. 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]
14. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. PLoS One 2017;12(12):e0189904 [FREE Full text] [doi: 10.1371/journal.pone.0189904] [Medline: 29267334]
15. 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]
16. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
17. Krukowski RA, Tilford JM, Harvey-Berino J, West DS. Comparing behavioral weight loss modalities: incremental cost-effectiveness of an internet-based versus an in-person condition. Obesity (Silver Spring) 2011 Aug;19(8):1629-1635 [FREE Full text] [doi: 10.1038/oby.2010.341] [Medline: 21253001]
18. Sjöström M, Umefjord G, Lindholm L, Samuelsson E. Cost-effectiveness of an internet-based treatment program for stress urinary incontinence. Neurourol Urodyn 2015 Mar;34(3):244-250 [FREE Full text] [doi: 10.1002/nau.22540] [Medline: 24347521]
19. van der Meer V, van den Hout WB, Bakker MJ, Rabe KF, Sterk PJ, Assendelft WJ, SMASHING (Self-Management in Asthma Supported by Hospitals, ICT, Nurses and General Practitioners) Study Group. Cost-effectiveness of internet-based self-management compared with usual care in asthma. PLoS One 2011;6(11):e27108 [FREE Full text] [doi: 10.1371/journal.pone.0027108] [Medline: 22096523]
20. Ljótsson B, Andersson G, Andersson E, Hedman E, Lindfors P, Andréewitch S, et al. Acceptability, effectiveness, and cost-effectiveness of internet-based exposure treatment for irritable bowel syndrome in a clinical sample: a randomized controlled trial. BMC Gastroenterol 2011 Oct 12;11:110 [FREE Full text] [doi: 10.1186/1471-230X-11-110] [Medline: 21992655]
21. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res 2010;12(5):e53 [FREE Full text] [doi: 10.2196/jmir.1436] [Medline: 21169166]
22. van Os-Medendorp H, Koffijberg H, Eland-de Kok PC, van der Zalm A, de Bruin-Weller MS, Pasmans SG, et al. E-health in caring for patients with atopic dermatitis: a randomized controlled cost-effectiveness study of internet-guided monitoring and online self-management training. Br J Dermatol 2012 May;166(5):1060-1068. [doi: 10.1111/j.1365-2133.2012.10829.x] [Medline: 22268960]
23. Dutton WH, Blank G. Oxford Internet Surveys. 2011. Next Generation Users: The Internet in Britain URL: http://oxis.oii.ox.ac.uk/wp-content/uploads/sites/43/2014/11/oxis2011-report.pdf [accessed 2018-06-19] [WebCite Cache ID 70I9QfXAJ]

XSL•FO
RenderX

24. Young C, Wilkins A. ResearchGate. 2013. Digital First - 1 Digital First The delivery choice for England's population URL: https://www.researchgate.net/publication/235664010_Digital_First_-_1_Digital_First_The_delivery_choice_for_England's_population [accessed 2018-06-19] [WebCite Cache ID 70I9o7xKL]
25. van Dyk JA. The evolution of the digital divide - the digital divide turns to inequality of skills and usage. In: Bus M, Crompton M, Hildebrandt M, Metakides G, editors. Digital Enlightenment Yearbook 2012. Amsterdam, The Netherlands: IOS Press; 2012:57-75.
26. Pearce KE, Rice RE. Digital divides from access to activities: comparing mobile and personal computer internet users. J Commun 2013;28(2):721-744. [doi: 10.1111/jcom.12045]
27. Sundquist M, Markendahl J. EconStor.: Econstor; 2015 Jun 27. A case study cost modelling of regulatory alternatives to mitigate the mobile network coverage and capacity problems in rural areas URL: http://hdl.handle.net/10419/127181 [accessed 2019-01-28] [WebCite Cache ID 75lWvDxFx]
28. Gatto SL, Tak SH. Computer, internet, and e-mail use among older adults: benefits and barriers. Educational Gerontology 2008 Aug 22;34(9):800-811. [doi: 10.1080/03601270802243697]
29. Xie B. Older adults, computers, and the internet: future directions. Gerontechnology 2003 Jun:289-305 [FREE Full text] [doi: 10.4017/gt.2003.02.04.002.00]
30. Farrington J, Philip LJ, Cottrill CD, Abbott PA, Blank G, Dutton W. Two-Speed Britain: Rural Internet Use. Aberdeen: Aberdeen University Press; 2015:1-56.
31. Rains SA. Health at high speed: broadband internet access, health communication, and the digital divide. Commun Res 2008;35(3):283-297 [FREE Full text] [doi: 10.1177/0093650208315958]
32. Lemaire E, Boudrias Y, Greene G. Low-bandwidth, internet-based videoconferencing for physical rehabilitation consultations. J Telemed Telecare 2001;7(2):82-89. [doi: 10.1258/1357633011936200] [Medline: 11331045]
33. Bartlett C, Simpson K, Turner AN. Patient access to complex chronic disease records on the internet. BMC Med Inform Decis Mak 2012 Aug 6;12(4):87-50 [FREE Full text] [doi: 10.1186/1472-6947-12-87] [Medline: 22867441]
34. Jones R. Development of a questionnaire and cross-sectional survey of patient eHealth readiness and eHealth inequalities. Med 2 0 2013;2(2):e9 [FREE Full text] [doi: 10.2196/med20.2559] [Medline: 25075244]
35. Smith AC, Kimble RM, O'brien A, Mill J, Wootton R. A telepaediatric burns service and the potential travel savings for families living in regional Australia. J Telemed Telecare 2016 Dec 2;13(3_suppl):76-79. [doi: 10.1258/135763307783247130]
36. Kruse RL, Koopman RJ, Wakefield BJ, Wakefield DS, Keplinger LE, Canfield SM, et al. Internet use by primary care patients: where is the digital divide? Fam Med 2012 May;44(5):342-347 [FREE Full text] [Medline: 23027117]
37. Wagner TH, Bundorf MK, Singer SJ, Baker LC. Free internet access, the digital divide, and health information. Med Care 2005 Apr;43(4):415-420. [doi: 10.1097/01.mlr.0000156857.14152.6e] [Medline: 15778645]
38. Bowen D, Meischke H, Bush N, Wooldridge J, Robbins R, Ludwig A, et al. Predictors of women's internet access and internet health seeking. Health Care Women Int 2003 Dec;24(10):940-951. [doi: 10.1080/07399330390244130] [Medline: 14742131]
39. Blackburn C, Read J. Using the internet? The experiences of parents of disabled children. Child Care Health Dev 2005 Sep;31(5):507-515. [doi: 10.1111/j.1365-2214.2005.00541.x] [Medline: 16101645]
40. Lenhart A, Horrigan J, Allen K, Boyce A, Madden M, O'Grady E. Pew Research Center. 2003. The Ever-Shifting Internet Population: A new look at Internet access and the digital divide URL: http://www.pewinternet.org/2003/04/16/the-ever-shifting-internet-population-a-new-look-at-internet-access-and-the-digital-divide/ [accessed 2018-06-19] [WebCite Cache ID 70IBjcmoU]
41. Mancini J, Noguès C, Adenis C, Berthet P, Bonadona V, Chompret A, et al. Patients' characteristics and rate of Internet use to obtain cancer information. J Public Health (Oxf) 2006 Sep;28(3):235-237. [doi: 10.1093/pubmed/fdl019] [Medline: 16809788]
42. Greenhalgh T, Wood GW, Bratan T, Stramer K, Hinder S. Patients' attitudes to the summary care record and HealthSpace: qualitative study. Br Med J 2008 Jun 7;336(7656):1290-1295 [FREE Full text] [doi: 10.1136/bmj.a114] [Medline: 18511764]
43. Khechine H, Pascot D, Prémont P. Use of health-related information from the Internet by English-speaking patients. Health Informatics J 2008 Mar;14(1):17-28. [doi: 10.1177/1460458207086331] [Medline: 18258672]
44. Williams P, Nicholas D, Huntington P. Non use of health information kiosks examined in an information needs context. Health Info Libr J 2003 Jun;20(2):95-103 [FREE Full text] [doi: 10.1046/j.1471-1842.2003.00428.x] [Medline: 12786908]
45. 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]
46. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. Int J Med Inform 2006 Jan;75(1):8-28. [doi: 10.1016/j.ijmedinf.2005.07.032] [Medline: 16125453]
47. Houston TK, Allison JJ. Users of internet health information: differences by health status. J Med Internet Res 2002 Dec;4(2):E7 [FREE Full text] [doi: 10.2196/jmir.4.2.e7] [Medline: 12554554]

XSL•FO
RenderX

48. Renahy E, Parizot I, Chauvin P. Health information seeking on the internet: a double divide? Results from a representative survey in the Paris metropolitan area, France, 2005-2006. BMC Public Health 2008;8:69 [FREE Full text] [doi: 10.1186/1471-2458-8-69] [Medline: 18291024]
49. Birru MS, Monaco VM, Charles L, Drew H, Njie V, Bierria T, et al. Internet usage by low-literacy adults seeking health information: an observational analysis. J Med Internet Res 2004 Sep 3;6(3):e25 [FREE Full text] [doi: 10.2196/jmir.6.3.e25] [Medline: 15471751]
50. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. J Med Internet Res 2011 Nov;13(4):e94 [FREE Full text] [doi: 10.2196/jmir.1750] [Medline: 22094891]
51. Blank G. Oxford Internet Surveys. Proxy users in the UK: an overview URL: http://oxis.oii.ox.ac.uk/blog/proxy-users-uk-overview/ [accessed 2018-06-20] [WebCite Cache ID 70JhlkrZh]
52. Adams N, Stubbs D, Woods V. Psychological barriers to Internet usage among older adults in the UK. Med Inform Internet Med 2005 Mar;30(1):3-17. [doi: 10.1080/14639230500066876] [Medline: 16036626]
53. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
54. Trocchia PJ, Janda S. A phenomenological investigation of Internet usage among older individuals. J Consum Mark 2000;17(7):605-616. [doi: 10.1108/07363760010357804]
55. Hardiker NR, Grant MJ. Barriers and facilitators that affect public engagement with eHealth services. Stud Health Technol Inform 2010;17:13-17. [Medline: 20841641]
56. Age UK. 2013. Digital inclusion evidence review URL: https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/reports-and-briefings/active-communities/rb_sept13_age_uk_digital_inclusion_evidence_review.pdf [accessed 2018-06-20] [WebCite Cache ID 70JiIRWsS]
57. Fogel J, Albert SM, Schnabel F, Ditkoff BA, Neugut AI. Use of the Internet by women with breast cancer. J Med Internet Res 2002;4(2):E9 [FREE Full text] [doi: 10.2196/jmir.4.2.e9] [Medline: 12554556]
58. Peterson MW, Fretz PC. Patient use of the internet for information in a lung cancer clinic. Chest 2003 Feb;123(2):452-457. [doi: 10.1378/chest.123.2.452] [Medline: 12576365]
59. Bukachi F, Pakenham-Walsh N. Information technology for health in developing countries. Chest 2007 Nov;132(5):1624-1630. [doi: 10.1378/chest.07-1760] [Medline: 17998362]
60. Légaré E, Vincent C, Lehoux P, Anderson D, Kairy D, Gagnon MP, et al. Telehealth readiness assessment tools. J Telemed Telecare 2010;16(3):107-109. [doi: 10.1258/jtt.2009.009004] [Medline: 20386031]
61. Snyder-Halpern R. Indicators of organizational readiness for clinical information technology/systems innovation: a Delphi study. Int J Med Inform 2001 Oct;63(3):179-204. [doi: 10.1016/S1386-5056(01)00179-4] [Medline: 11502432]
62. Lehman WE, Greener JM, Simpson DD. Assessing organizational readiness for change. J Subst Abuse Treat 2002 Jun;22(4):197-209. [doi: 10.1016/S0740-5472(02)00233-7] [Medline: 12072164]
63. Oliver DR, Demiris G. An assessment of the readiness of hospice organizations to accept technological innovation. J Telemed Telecare 2004;10(3):170-174. [doi: 10.1258/135763304323070832] [Medline: 15165444]
64. Demiris G, Patrick T, Khatri N. Assessing home care agencies' readiness for telehealth. AMIA Annu Symp Proc 2003:825 [FREE Full text] [Medline: 14728330]
65. Khoja S, Scott RE, Casebeer AL, Mohsin M, Ishaq AF, Gilani S. e-Health readiness assessment tools for healthcare institutions in developing countries. Telemed J E Health 2007 Aug;13(4):425-431. [doi: 10.1089/tmj.2006.0064] [Medline: 17848110]
66. Jennett P, Jackson A, Ho K, Healy T, Kazanjian A, Woollard R, et al. The essence of telehealth readiness in rural communities: an organizational perspective. Telemed J E Health 2005 Apr;11(2):137-145. [doi: 10.1089/tmj.2005.11.137] [Medline: 15857254]
67. 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]
68. Zoopla. 2012. Zoopla Full Postcode Finder URL: https://www.zoopla.co.uk/postcode-finder/ [accessed 2019-01-28] [WebCite Cache ID 75lZEq7Bm]
69. Broos A, Roe K. The digital divide in the playstation generation: self-efficacy, locus of control and ICT adoption among adolescents. Poetics 2006;34(4):306-317 [FREE Full text] [doi: 10.1016/j.poetic.2006.05.002]
70. National Health Service. 2015. Find GP Services URL: https://www.nhs.uk/Service-Search/GP/LocationSearch/4 [accessed 2018-06-20] [WebCite Cache ID 70Jj3HouU]
71. Learn My Way. 2014. Learn how to use the internet URL: https://www.learnmyway.com/ [accessed 2019-01-28] [WebCite Cache ID 75lZ9jm9X]
72. Office for National Statistics. 2016 May 20. Internet users in the UK Internet users in the UK: 2016 URL: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2016 [accessed 2018-01-28] [WebCite Cache ID 70JjHjosb]
73. Thackeray R, Neiger BL, Smith AK, van Wagenen SB. Adoption and use of social media among public health departments. BMC Public Health 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]

XSL•FO
RenderX

74. Ofcom. 2016 Mar 24. UK Home broadband performance: The performance of fixed-line broadband delivered to UK residential consumers URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0030/78267/fixed-bb-speeds-nov15-consumer-summary.pdf [accessed 2019-01-28] [WebCite Cache ID 70JjU8shl]
75. YouTube Help. System requirements URL: https://support.google.com/youtube/answer/78358?hl=en-GB [accessed 2018-06-20] [WebCite Cache ID 70JjZht54]
76. St Levan Surgery. URL: http://www.stlevansurgery.nhs.uk [accessed 2018-06-20] [WebCite Cache ID 70JjdRyOM]
77. Campbell JL, Fletcher E, Britten N, Green C, Holt TA, Lattimer V, et al. Telephone triage for management of same-day consultation requests in general practice (the ESTEEM trial): a cluster-randomised controlled trial and cost-consequence analysis. Lancet 2014 Nov 22;384(9957):1859-1868 [FREE Full text] [doi: 10.1016/S0140-6736(14)61058-8] [Medline: 25098487]
78. Gavgani VZ, Shiramin AR. Physician directed information prescription service (IPs): barriers and drivers. Aslib Proc 2013 Mar;65(3):224-241. [doi: 10.1108/00012531311330629]
79. Jones RB, O'Connor A, Brelsford J, Parsons N, Skirton H. Costs and difficulties of recruiting patients to provide e-health support: pilot study in one primary care trust. BMC Med Inform Decis Mak 2012 Mar 29;12:25 [FREE Full text] [doi: 10.1186/1472-6947-12-25] [Medline: 22458706]
80. VanGeest JB, Johnson TP, Welch VL. Methodologies for improving response rates in surveys of physicians: a systematic review. Eval Health Prof 2007 Dec;30(4):303-321. [doi: 10.1177/0163278707307899] [Medline: 17986667]
81. de Grood C, Raissi A, Kwon Y, Santana MJ. Adoption of e-health technology by physicians: a scoping review. J Multidiscip Healthc 2016;9:335-344 [FREE Full text] [doi: 10.2147/JMDH.S103881] [Medline: 27536128]
82. Anderson JG. Social, ethical and legal barriers to e-health. Int J Med Inform 2007;76(5-6):480-483. [doi: 10.1016/j.ijmedinf.2006.09.016] [Medline: 17064955]

Abbreviations

BT: British Telecom
eHEALS: eHealth literacy scale
eHealth: electronic health
GP: general practitioner
GPI: general practice intervention
ICT: information and communication technology
NHS: National Health Service
NS: did not have Superfast
PERQ: Personal eHealth Readiness Questionnaire
S: Superfast
TB: tailored booklet
TBI: tailored booklet intervention

Edited by G Eysenbach; submitted 26.06.18; peer-reviewed by J Powell, S Perez; comments to author 07.10.18; revised version received 11.11.18; accepted 09.12.18; published 11.03.19.

Please cite as:
Abbott-Garner P, Richardson J, Jones RB
The Impact of Superfast Broadband, Tailored Booklets for Households, and Discussions With General Practitioners on Personal Electronic Health Readiness: Cluster Factorial Quasi-Randomized Control Trial
J Med Internet Res 2019;21(3):e11386
URL: http://www.jmir.org/2019/3/e11386/
doi:10.2196/11386
PMID:30855234

©Philip Abbott-Garner, Janet Richardson, Ray B Jones. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 11.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Cancer-Related Internet Use and Its Association With Patient Decision Making and Trust in Physicians Among Patients in an Early Drug Development Clinic: A Questionnaire-Based Cross-Sectional Observational Study

Goldy C George[1], PhD; Eucharia C Iwuanyanwu[1], PHS, MPAS, PAC; Adrianna S Buford[1], MPH; Sarina A Piha-Paul[1], MD; Vivek Subbiah[1], MD; Siqing Fu[1], MD; Daniel D Karp[1], MD; Shubham Pant[1], MBBS; Christina O Hinojosa[1], MS, CCRP; Kenneth R Hess[2], MS, PhD; Charles S Cleeland[3], PhD; Elmer V Bernstam[4], MD, MSE; Funda Meric-Bernstam[1], MD; David S Hong[1], MD

[1]Department of Investigational Cancer Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[2]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[3]Department of Symptom Research, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[4]School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, United States

Corresponding Author:
Goldy C George, PhD
Department of Investigational Cancer Therapeutics
The University of Texas MD Anderson Cancer Center
Houston, TX,
United States
Phone: 1 7137941347
Email: ggeorge2@mdanderson.org

Abstract

Background: The role of cancer-related internet use on the patient-physician relationship has not been adequately explored among patients who are cancer-related internet users (CIUs) in early-phase clinical trial clinics.

Objective: We examined the association between cancer-related internet use and the patient-physician relationship and decision making among CIUs in an early drug development clinic.

Methods: Of 291 Phase I clinic patients who completed a questionnaire on internet use, 179 were CIUs. Generations were defined by the year of patient's birth: "millennials" (after 1990) and "Generation X/Y" (1965-1990) grouped as "Millennials or Generation X/Y"; "Baby Boomers" (1946-1964); and "Greatest or Silent Generation" (1945 and earlier). Statistical analyses included the Wilcoxon matched-pairs signed-rank test and the Mann-Whitney U test.

Results: CIUs were 52% (94/179) female, 44% (78/179) were older than 60 years, and 60% (108/179) had household incomes exceeding US $60,000. The sources of information on cancer and clinical trials included physicians (171/179, 96%), the internet (159/179, 89%), and other clinical trial personnel (121/179, 68%). For the overall sample and each generation, the median values for trust in referring and Phase I clinical trial physicians among early drug development clinic CIUs were 5 on a 0-5 scale, with 5 indicating "complete trust." CIUs' trust in their referring (5) and phase 1 (5) physicians was higher than CIUs' trust in Web-based cancer-related information (3; $P<.001$ for both). CIUs who reported visiting the National Cancer Institute (NCI) website, NCI.org, to learn about cancer reported higher levels of trust in Web-based cancer-related information than CIUs who did not use the NCI website ($P=.02$). Approximately half of CIUs discussed internet information with their doctor. Only 14% (23/165) of CIUs had asked their physician to recommend cancer-related websites, and 24% (35/144) of CIUs reported at least occasional conflict between their physician's advice and Web-based information.

Conclusions: Despite the plethora of websites related to cancer and cancer clinical trials, patients in early-phase clinical trial settings trust their physicians more than Web-based information. Cancer-related organizations should provide regularly updated links to trustworthy websites with cancer and clinical trial information for patients and providers and educate providers on reliable cancer websites so that they can better direct their patients to appropriate internet content.

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e10348) doi:10.2196/10348

KEYWORDS

cancer; clinical trial; doctor; early-phase clinical trial; internet; patient; physician; patient-physician relationship; symptoms

Introduction

Early-phase clinical trials form a critical link between preclinical testing of novel therapies and Food and Drug Administration approval of these agents [1]. Patients in early-phase oncology clinical trials tend to have advanced cancer that is refractory to standard therapies. Given the complexities of early-phase clinical trials [1] and the advanced nature of the disease being treated, patients in early drug development clinics need accurate information when they consider enrollment or continued treatment in clinical trials. Information for early-phase clinical trial patients is especially important in the light of the suggestion that some patients may join early-phase clinical trials without fully understanding all of the goals of Phase I trials (particularly goals related to the research aspects of an early-phase clinical trial, such as dose escalation and estimation of maximum tolerated dose) [2]. Also, patients may sometimes join clinical trials with unrealistic expectations of personal disease-related benefits from the trial, under a condition termed as the therapeutic misconception [2].

The internet is often used as a source of cancer-related information by patients with cancer [3], including early-phase clinical trial clinic patients [4]. The information found on the internet may be a source of empowerment for some patients, but it may also present conflicting information, leaving some patients anxious or confused [4,5]. Also, with the mainstreaming of social networking sites, patients are able to form a circle of faceless friends in chat rooms, through blogs, and in Web-based communities dedicated to people like themselves. For example, the use of internet cancer support groups has increased [6]. Thus, it has been stated that the internet has the potential to redefine patient-physician relationship due to the vast amounts of information on the internet that are accessible to patients [7,8]. Also, patients may bring internet-derived cancer- and clinical trial-related information to consultations with their oncologist [3] to supplement, confirm, or refute information provided by their physician. Alternatively, patients may seek out vast amounts of cancer- and treatment-related information to promote a trusting and therapeutic doctor-patient relationship [9] (relating to the concept of psychological autonomy). While several models (including paternalistic, informative, interpretive, and deliberative models) [10-12] have been proposed to understand or explain the doctor-patient relationship and the level of patient involvement in the treatment decision-making process, the patient-physician relationship and related aspects have not been adequately explored among patients in early drug development clinic settings who use the internet for cancer-related purposes. The purpose of this study was to examine cancer-related internet users (CIUs) in an early drug development clinic, factors related to the patient-physician relationship (eg, patients' trust in referring and clinical trial physicians), and decision making (eg, patients' decisions, such as selecting the treating hospital and joining a clinical trial). This study ties in with the goals of the National Cancer Moonshot, which seeks to increase patient access to quality cancer care and increase clinical trial participation.

Methods

Procedure

Patients were recruited from the Clinical Center for Targeted Therapy (CCTT), the Phase I clinical trials clinic at The University of Texas MD Anderson Cancer Center. CCTT is one of the largest early-phase clinical trial centers in the world. CCTT patients have advanced solid tumors or lymphomas and are self-referred or referred by their primary oncologist. Phase I or CCTT patients were approached with a standardized script and requested to complete an anonymous questionnaire as they waited for their scheduled clinical appointment. Patients were asked to deposit their completed questionnaire in a locked box, rather than to return it to a person, in order to ensure the anonymity of the patient completing the survey.

Patients

Patients who had a diagnosis of advanced cancer, were able to read and understand English, and were aged ≥18 years were eligible for the study. Patients who had severe visual or cognitive impairment that would interfere with the ability to complete the questionnaire were excluded. The study was approved by the Institutional Review Board of MD Anderson, and all participants provided informed consent.

Materials

The questionnaire was developed by an investigative team with expertise in early-phase clinical trials and research questionnaire development methodology [4,13]. Questions were based on extant literature relating to internet use in patients with cancer [14-17] or developed de novo for our study, taking into account the unique nature of the phase I clinic's patient population [4]. The questionnaire was pilot-tested among 18 patients from the early-phase clinical trials clinic, and feedback on the ease of understanding, readability, and applicability of the context to patients with advanced cancer was elicited [4]. Based on patients' feedback during the pilot test, the survey instrument was further revised. The final survey instrument was thus tailored to the early-phase clinical trials context. The questionnaire included questions on demographics, cancer-related internet use, trust in physicians, and patient decision making related to whether they use the internet to make decisions about where to be treated and whether to join a trial. For example, trust in a referring physician was assessed using the question, "How much do you trust your referring physician?" Trust in Web-based cancer information was assessed with "How much do you trust Web-based cancer-related information?" Trust in referring and Phase I physicians and trust in Web-based cancer-related information were measured using six-point Likert-type scales where 0=none and 5=complete trust.

XSL•FO
RenderX

Statistical Analysis

SPSS v23 (IBM Corp) was used for analyses. Generational differences have been shown to be associated with certain aspects of cancer internet use patterns [18]. Thus, patients were grouped by generation based on year of birth: "Millennials" (after 1990) and "Generation X/Y" (1965-1990) combined as "Millennials or Generation X/Y"; "Baby Boomers" (1946-1964); and "Greatest or Silent Generation" (1945 and earlier). Box plots were used to graphically represent the distributions (including medians and interquartile ranges) of the variables of trust in referring and Phase I physicians and trust in Web-based cancer-related information for the overall sample and by generation. The Wilcoxon matched-pairs signed-rank test was used to compare CIUs' trust in physicians with their trust in Web-based cancer-related information. The Mann-Whitney U test was used to compare CIUs' trust in their referring or Phase I physician and trust in Web-based cancer-related information based on whether they used selected websites (eg, National Cancer Institute [NCI] website) to learn about cancer or clinical trials. Significance was set at an alpha level of.05, and all tests were two-tailed.

Results

Patient Characteristics

The questionnaire was completed by 291 patients in the phase I clinic (CCTT), of whom 179 (179/291, 62%) were CIUs; these CIUs are the focus of this manuscript. CIUs were defined as those who had indicated a response of "Yes" to the question "Do you access the internet for cancer-related purposes (for example, for cancer- or clinical trial-related information or for emotional or social support for your cancer)?" Characteristics of CIUs are included in Table 1.

Sources of Information on Cancer and Clinical Trials for Patients

Among CIUs in the Phase I clinical trials clinic, 96% (171/179) relied on physicians as a source of information on cancer and clinical trials and 89% (159/179) used the internet for cancer-related and clinical trial-related information. Other sources of information that CIUs used to learn about cancer and cancer treatment or clinical trials included other clinical trial personnel, such as nurses or physician assistants (121/179, 68%); information pamphlets from the treating hospital, the NCI, the American Cancer Society, the American Society of Clinical Oncology, or other cancer-related organizations (83/179, 46%); other patients (64/179, 36%); books (47/179, 26%) and magazines (35/179, 20%); scientific journals (33/179, 18%); and smartphone apps (8/179, 5%). Multimedia Appendix 1 shows the sources of information on cancer and cancer treatment (including clinical trials) for the overall sample of patients in the early drug development (Phase I clinic; N=291).

Internet and Patient Decision Making Related to the Selection of Treating Hospital and Clinical Trials Clinic

The major drivers of where CIUs decided to receive their cancer care were the reputation of the hospital and referring physician recommendations (Table 2). Nearly 58% (99/171) of CIUs reported that the internet had not influenced their decision to come to the treating hospital in any way; the remaining 42% (72/171) of CIUs indicated that the internet had influenced their decision to come to the treating hospital. Most CIUs indicated that the internet had not influenced their decision to visit the Phase I clinic at the treating hospital (134/165, 81%) or their decision to enroll in a Phase I clinical trial (129/160, 81%). Nearly 90% (149/168, 89%) of CIUs stated that the major determinant of their choice of clinical trial physician was based on their referring physician's recommendation.

XSL•FO
RenderX

Table 1. Characteristics of cancer-related internet users in an early drug development clinic (N=179).

Characteristics	Internet users, n (%)
Generation	
Millennials or Generation X/Y	49 (24)
Baby Boomers	89 (47)
Greatest or Silent	41 (29)
Gender	
Male	85 (48)
Female	94 (53)
Race or ethnicity	
Nonminority	152 (85)
Minority	27 (15)
Tumor type	
Lung	22 (12)
Colorectal	20 (11)
Melanoma	19 (11)
Head and neck	16 (9)
Ovarian	16 (9)
Breast	15 (8)
Sarcoma	10 (6)
Prostate	9 (5)
Endometrial	9 (5)
Pancreatic	6 (3)
Others[a]	37 (21)

[a]Included thyroid (n=6), lymphoma (n=3), hepatocellular (n=3), cervical (n=3), kidney (n=2), brain (n=2), and other tumor types (n=18).

XSL•FO
RenderX

Table 2. Decision making related to cancer treating hospital and early-phase clinical trial enrollment among cancer-related internet users in an early drug development clinic.

Questionnaire item	Overall (N=179), n (%)	Millennials or Generation X/Y (n=49), n (%)	Baby Boomers (n=89), n (%)	Greatest or Silent (n=41), n (%)
How did you decide where to get your cancer care?				
Reputation of cancer care organization/hospital	112/174 (64)	35/48 (73)	47/85 (55)	30/41 (73)
Physician recommendation	111/174 (64)	30/48 (63)	57/85 (67)	24/41 (59)
Family or friend recommendation	42/174 (24)	12/48 (25)	19/85 (22)	11/41 (27)
Did the information on the Internet influence your decision to come to the treating hospital?				
No, not at all	99/171 (58)	21/48 (44)	53/83 (64)	25/40 (63)
Yes, absolutely	45/171 (26)	14/48 (29)	22/83 (27)	9/40 (23)
Yes, somewhat	27/171 (16)	13/48 (27)	8/83 (10)	6/40 (15)
Did the information on the Internet influence your decision to come to the Phase I clinic at the treating hospital?				
No, not at all	134/165 (81)	38/47 (81)	63/81 (78)	33/37 (89)
Yes, somewhat	15/165 (9)	4/47 (9)	8/81 (10)	1/37 (3)
Yes, absolutely	16/165 (10)	5/47 (11)	10/81 (12)	3/37 (8)
Did the information on the Internet influence your decision to enroll on a Phase I clinical trial?				
No, not at all	129/160 (81)	35/45 (78)	62/79 (79)	32/36 (89)
Yes, somewhat	14/160 (9)	4/45 (9)	9/79 (11)	2/36 (6)
Yes, absolutely	17/160 (11)	6/45 (13)	8/79 (10)	2/36 (6)
How did you decide which clinical trial physician to use?				
Referring physician recommendation	149/168 (89)	44/47 (94)	71/83 (86)	34/38 (90)
Assigned by the treating hospital	15/168 (9)	1/47 (2)	10/83 (12)	4/38 (11)
Family or friend recommendation	6/168 (4)	3/47 (6)	1/83 (1)	2/38 (5)
Internet	4/164 (2)	1/47 (2)	4/83 (5)	1/38 (3)

Patient Discussions of Internet Information With Their Physician

Approximately half of the CIUs indicated that they had discussed or planned to discuss internet information with their doctor (Table 3). Among CIUs who answered “Yes” to the question “Have you discussed/do you plan to discuss Internet information with your doctor?”, reasons cited for discussing internet information with their physician included wanting to educate themselves (69/86, 80%), to be proactive to improve their health (65/86, 76%), to obtain their physician’s expertise (62/86, 72%), to educate their caregiver or family member (37/86, 43%), to meet an emotional need (10/86, 12%), or other reasons (to “clarify confusing information”, “discuss options”, “make sure important information is correct”, 3/86, 4%). For CIUs who answered “No” to the question “Have you discussed/do you plan to discuss internet information with your doctor?”, reasons for not doing so included trusting their physician to make the best choice for them (31/68, 46%), not feeling a need to clarify anything with their physician (22/68, 32%), believing that the information was not relevant to their situation or disease (5/68, 7%), forgetting to mention the information (5/68, 7%), not having enough time with physician (4/68, 6%), not wanting to bother the doctor (4/68, 6%), or other reasons (including “they already know”, 2/68, 3%).

Comparison of Patients’ Trust in Referring and Phase I Physicians and Their Trust in Web-based Cancer Information

In the overall sample and in each generation, median values for trust in the referring and Phase I clinical trial physicians among CIUs were 5 on a 0-5 scale, with 5 indicating “complete trust” (Figure 1). Trust in referring or Phase I physician did not differ significantly between CIUs who reported accessing the internet at least sometimes through smartphone or cell phone for cancer or clinical trial information and CIUs who reported that they never or rarely did so (P>.05). Also, CIUs’ trust in their referring or phase I physician did not vary by CIUs self-reported rating of their symptoms of feeling depressed or anxious, pain, tiredness or fatigue, or of their experience of difficulty moving around (P>.05).

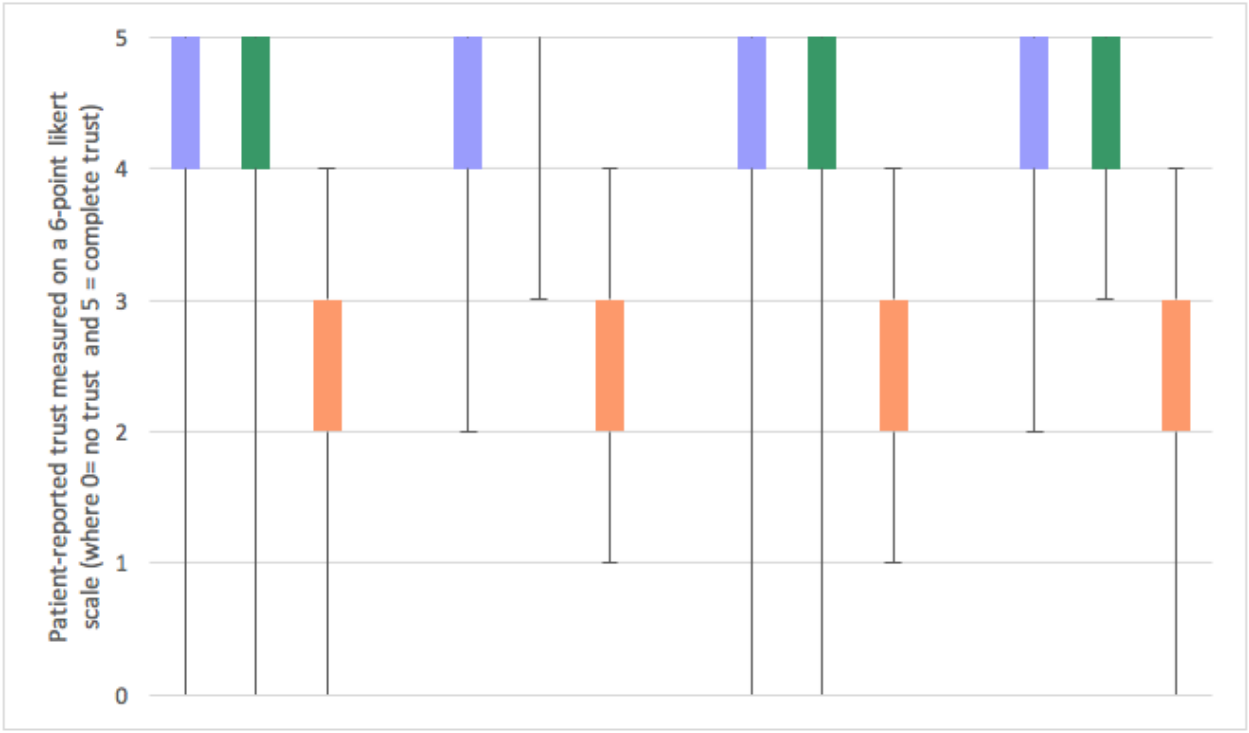
XSL•FO
RenderX

Table 3. Impact of the internet on the patient-physician relationship among cancer-related internet users in an early drug development clinic.

Characteristics	Overall (N=179), n (%)	Millennials or Generation X/Y (n=49), n (%)	Baby Boomers (n=89), n (%)	Greatest or Silent (n=41), n (%)
Have you discussed/do you plan to discuss Internet information with your doctor?				
Yes	87/167 (52)	26/46 (56)	49/82 (60)	12/39 (31)
No	80/167 (48)	20/46 (44)	33/82 (40)	27/39 (69)
How do you verify that information from the Internet is correct?				
I ask my doctor	111/156 (71)	33/46 (72)	52/73 (71)	26/37 (70)
I ask the PA, APN, or other members of the clinical team	64/156 (41)	14/46 (30)	36/73 (49)	14/37 (38)
If I see it in multiple places I believe it is true	34/156 (22)	14/46 (30)	13/73 (18)	7/37 (19)
I don't verify that things I read online are true	14/156 (9)	3/46 (7)	6/73 (8)	5/37 (14)
Other (source credibility, I check other sources for data and evidence, I visit only reliable sites, sometimes I just take a chance)	8/156 (5)	2/46 (4)	6/73 (8)	0/37 (0)
Do you trust your physician's advice over the information found on the Internet?				
Yes	158/163 (97)	43/44 (98)	79/82 (96)	36/37 (97)
No	5/163 (3)	1/44 (2)	3/82 (4)	1/37 (3)
If you trust your physician's advice more, what are the reasons for that?				
Your physician has had extensive training and will make the best decisions for you	121/150 (81)	37/42 (88)	58/73 (80)	26/35 (74)
You have a good relationship with your physician	84/150 (56)	22/42 (52)	44/73 (60)	18/35 (51)
Your physician can explain information better	67/150 (45)	15/42 (36)	36/73 (49)	16/35 (46)
If you trust advice from the Internet more, what are the reasons for that?				
More comprehensive	1/3 (33)	0/1 (0)	0/1 (0)	1/1 (100)
More up-to-date	1/3 (33)	0/1 (0)	0/1 (0)	1/1 (100)
Did your physician's recommendation conflict with what you found on the Internet?				
No	105/144 (73)	30/40 (75)	48/69 (70)	27/35 (77)
Sometimes yes, sometimes no	35/144 (24)	9/40 (23)	18/69 (26)	8/35 (23)
Yes	4/144 (3)	1/40 (3)	3/69 (4)	0/35 (0)
Did you ask your physician for recommendations for cancer-related websites?				
No	142/165 (86)	41/47 (87)	66/81 (82)	35/37 (95)
Yes	23/165 (14)	6/47 (13)	15/81(19)	2/37 (5)

XSL•FO
RenderX

Figure 1. Box plots showing the distribution of trust in referring physicians, early-phase clinical trial physicians, and Web-based cancer-related information for cancer-related internet users (CIUs) in an early drug development clinic. Purple boxes indicate trust in referring physician, green boxes indicate trust in Phase I clinical trial physician, and orange boxes indicate trust in Web-based cancer-related information. In the overall sample of CIUs and within each generation, CIUs reported significantly higher trust in their referring or Phase I physician than trust in Web-based cancer-related information.



Based on the median scores of trust measured on the six-point Likert scales ranging from 0 (“no trust”) to 5 (“complete trust”) in this anonymous survey, CIUs trusted referring and Phase I physicians more than Web-based cancer-related information (5.0 vs 3.0, *P*<.001 for both; Figure 1). In general, CIUs’ trust in their referring or Phase I physician and CIUs’ trust in Web-based cancer information did not vary by the specific websites that CIUs reported using to learn about cancer or clinical trials (*P*>.05). However, CIUs who reported visiting the NCI website to learn about cancer reported higher levels of trust in Web-based cancer-related information than CIUs who did not use the NCI website (median=4.0 for both, but a significant difference between groups based on the Mann-Whitney U test: Mann-Whitney U=2576; *P*=.02). Also, CIUs’ trust in their referring or phase I physician did not vary by whether the internet had made them aware of information on their cancer, clinical trials, side effects of treatment, management of symptoms, prognosis for their cancer, or new or alternative treatments (*P*>.05).

Nearly 97% (158/163) of CIUs indicated that they trusted their physician’s advice over the information on the internet (Table 3). Almost all CIUs (150/153, 98% of CIUs) reported that internet-derived information did not cause conflict between them and their treating physician. Also, almost all CIUs (157/158, 99%) indicated that information on the internet did not cause them to ignore their doctor’s advice. Approximately 73% (105/144) of CIUs indicated that their physician’s recommendation did not conflict with what they found on the internet, whereas 24% (35/144) reported at least occasional differences between their physician’s advice and Web-based information. Only 14% (23/165) of CIUs had asked their physician for recommendations for cancer-related websites.

Factors Patients Use to Determine the Trustworthiness of Websites

CIUs indicated that their means for determining the trustworthiness of websites providing medical information related to their disease were the reputation of the website (90/152, 59%), recommendation of physician (77/152, 51%), and credentials of website authors (69/152, 45%). When asked about factors that they valued most on cancer-related websites, Phase I clinic CIUs cited trustworthiness (96/157, 61%), ease of understanding (86/157, 55%), ease of navigation (69/157, 44%), and physician recommendation (59/157, 38%). When asked “Compared to websites with advertisements, how much would you trust websites without advertisements?”, 50% (80/160) of CIUs indicated “about the same”, 39% (66/160) indicated “more”, and 9% (14/160) indicated “less”.

Discussion

Principal Findings

To our knowledge, this is the first analysis of the impact of cancer-related internet use on the patient-physician relationship among patients with advanced cancer in an early drug development clinic. A very high percentage of early-phase CIU patients used the internet for information. We found that, despite the prevalence of the internet, social media, and smartphone

XSL•FO
RenderX

apps, early drug development clinic patients relied most on physicians as a source of information on cancer and clinical trials. We also found that CIUs reported high trust in their physicians. Approximately half of the CIUs discussed or planned to discuss internet information with their physicians. Few CIUs asked physicians to recommend websites.

Despite the ubiquity of the internet, CIUs in the early drug development clinic still relied on physicians for information about clinical trials and cancer. Early drug development clinic patients also reported a high level of trust in their referring and Phase I clinical trial physicians. This is congruent with findings in cancer survivors [19] the general population of adults [8,20], patients with breast cancer [21], and patients with early-stage papillary thyroid cancer [22], suggesting that patients have a deep level of trust in their physicians and that physicians are the preferred and most trusted source of health information.

The finding that 42% (72/171) of CIUs used the internet, in part, to select the institution in which they would investigate or pursue cancer care options suggests the potential of internet use and cancer center marketing in influencing health care utilization. By contrast, few patients (23/165, 14%) sought physician recommendations for cancer-related websites. We may be missing an opportunity to educate patients by recommending additional informative websites with clinical trial information, especially in an era of ever-shorter clinic visits. This is also reflected in our finding that physician recommendation was a common reason for trusting a website—if physicians did recommend specific websites, the recommendation(s) would likely influence patient behavior.

Although most patients reported that their physician's recommendation did not conflict with the information they found on the internet, nearly 1 in 4 patients reported at least occasional conflict. This is particularly important given that the quality of internet-based oncology and clinical trial information may vary greatly [23,24]. As some internet sources may be based on hearsay and may not be well-resourced, it might be helpful to encourage physicians or members of the clinical team to discuss what the patient has seen on the internet and help clarify any discrepancies. It may also be helpful for providers to point patients to trustworthy cancer-related websites. Cancer-related organizations, such as the American Society for Clinical Oncology, the American Association for Cancer Research, and governmental institutes such as the NCI, should consider augmenting regularly updated links to recommended websites with trustworthy cancer and clinical trial information and educating physicians so that they can better guide their patients.

Strengths and Limitations

This study offered several insights into the impact of cancer-related internet use on the doctor-patient relationship among early drug development clinic patients. Its strengths are that the questionnaire incorporated input from both oncologists and patients in the early drug development clinic setting and that it was pilot-tested among early drug development clinic patients. Limitations include that data on the number of patients who were approached and the number of patients who declined to participate were not collected, and therefore, the response rate cannot be calculated. A possible limitation is that this is a single-center study and that our Phase I clinical trials unit is well-recognized for its early-phase trials. Thus, our study population may be quite unique in that patients may come to MD Anderson's Phase I unit because of the availability of clinical trials. Whether similar results would be observed in other institutions conducting early-phase trials should be investigated in future studies.

Conclusion

In conclusion, the results suggest that despite the large number of websites related to cancer and cancer clinical trials, patients in early-phase clinical trial settings still trust their physicians for cancer or clinical trial information over information found on the internet. Cancer-related organizations should consider augmenting regularly updated links to trustworthy, verified websites with cancer and clinical trial information and should educate physicians so that they can better direct their patients to appropriate internet content.

Acknowledgments

The authors thank Jeanie F Woodruff for the scientific editing of the manuscript.

The research reported in this manuscript was presented in part at the NCI-EORTC-AACR Molecular Targets and Cancer Therapeutics International Symposium in Munich, Germany, and the abstract was published as part of conference proceedings in the European Journal of Cancer 2016;68:S65.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sources of information on cancer and cancer treatment (including clinical trials) for early drug development clinic patients (N=291).

[PDF File (Adobe PDF File), 34KB - jmir_v21i3e10348_app1.pdf]

XSL•FO
RenderX

References

1. Wong KM, Capasso A, Eckhardt SG. The changing landscape of phase I trials in oncology. Nat Rev Clin Oncol 2016 Feb;13(2):106-117. [doi: 10.1038/nrclinonc.2015.194] [Medline: 26552953]
2. Godskesen T, Nygren P, Nordin K, Hansson M, Kihlbom U. Phase 1 clinical trials in end-stage cancer: patient understanding of trial premises and motives for participation. Support Care Cancer 2013 Nov;21(11):3137-3142. [doi: 10.1007/s00520-013-1891-7] [Medline: 23864288]
3. Hesse BW, Greenberg AJ, Rutten LJF. The role of Internet resources in clinical oncology: promises and challenges. Nat Rev Clin Oncol 2016 Dec;13(12):767-776. [doi: 10.1038/nrclinonc.2016.78] [Medline: 27273045]
4. George GC, Buford A, Hess K, Piha-Paul SA, Zinner R, Subbiah V, et al. Cancer-Related Internet Use and Online Social Networking Among Patients in an Early-Phase Clinical Trials Clinic at a Comprehensive Cancer Center. JCO Clinical Cancer Informatics 2018 Nov 01;24(2):1-14. [doi: 10.1200/CCI.17.00030]
5. Groen WG, Kuijpers W, Oldenburg HS, Wouters MW, Aaronson NK, van HWH. Empowerment of Cancer Survivors Through Information Technology: An Integrative Review. J Med Internet Res 2015;17(11):e270 [FREE Full text] [doi: 10.2196/jmir.4818] [Medline: 26614438]
6. Im E, Lee B, Chee W. Shielded from the real world: perspectives on Internet cancer support groups by Asian Americans. Cancer Nurs 2010;33(3):E10-E20 [FREE Full text] [doi: 10.1097/NCC.0b013e3181c8e5d5] [Medline: 20357657]
7. Hartzband P, Groopman J. Untangling the Web--patients, doctors, and the Internet. N Engl J Med 2010 Mar 25;362(12):1063-1066. [doi: 10.1056/NEJMp0911938] [Medline: 20335581]
8. Hesse BW, Moser RP, Rutten LJ. Surveys of physicians and electronic health information. N Engl J Med 2010 Mar 4;362(9):859-860. [doi: 10.1056/NEJMc0909595] [Medline: 20200398]
9. Lee ST, Lin J. A Self-Determination Perspective on Online Health Information Seeking: The Internet vs. Face-to-Face Office Visits With Physicians. J Health Commun 2016 Jun;21(6):714-722. [doi: 10.1080/10810730.2016.1157651] [Medline: 27186966]
10. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. JAMA 1992;267(16):2221-2226. [Medline: 1556799]
11. Emanuel EJ, Pearson SD. Physician autonomy and health care reform. JAMA 2012 Jan 25;307(4):367-368. [doi: 10.1001/jama.2012.19] [Medline: 22274681]
12. Day TA. The ultimate constant of head and neck oncology—the physician-patient relationship. JAMA Otolaryngol Head Neck Surg 2014 Dec;140(12):1218-1224. [doi: 10.1001/jamaoto.2014.1567] [Medline: 25068431]
13. Hong DS, George GC, Iwuanyanwu EC, Tavana B, Falchook GS, Piha-Paul SA, et al. Factors related to biopsy willingness in patients with advanced cancer in a phase 1 clinic for molecularly targeted therapy. J Cancer Res Clin Oncol 2013 Jun;139(6):963-970. [doi: 10.1007/s00432-013-1404-6] [Medline: 23455881]
14. Castleton K, Fong T, Wang-Gillam A, Waqar MA, Jeffe DB, Kehlenbrink L, et al. A survey of Internet utilization among patients with cancer. Support Care Cancer 2011 Aug;19(8):1183-1190. [doi: 10.1007/s00520-010-0935-5] [Medline: 20556435]
15. James N, Daniels H, Rahman R, McConkey C, Derry J, Young A. A study of information seeking by cancer patients and their carers. Clin Oncol (R Coll Radiol) 2007 Jun;19(5):356-362. [doi: 10.1016/j.clon.2007.02.005] [Medline: 17399963]
16. Bylund CL, Gueguen JA, D'Agostino TA, Imes RS, Sonet E. Cancer patients' decisions about discussing Internet information with their doctors. Psychooncology 2009 Nov;18(11):1139-1146. [doi: 10.1002/pon.1511] [Medline: 19137507]
17. van de Poll-Franse LV, van Eenbergen MCHJ. Internet use by cancer survivors: current use and future wishes. Support Care Cancer 2008 Oct;16(10):1189-1195. [doi: 10.1007/s00520-008-0419-z] [Medline: 18293014]
18. Girault A, Ferrua M, Lalloué B, Sicotte C, Fourcade A, Yatim F, et al. Internet-based technologies to improve cancer care coordination: current use and attitudes among cancer patients. Eur J Cancer 2015 Mar;51(4):551-557. [doi: 10.1016/j.ejca.2014.12.001] [Medline: 25661828]
19. Finney Rutten LJ, Agunwamba AA, Wilson P, Chawla N, Vieux S, Blanch-Hartigan D, et al. Cancer-Related Information Seeking Among Cancer Survivors: Trends Over a Decade (2003-2013). J Cancer Educ 2016 Jun;31(2):348-357. [doi: 10.1007/s13187-015-0802-7] [Medline: 25712202]
20. 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]
21. 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]
22. D'Agostino TA, Shuk E, Maloney EK, Zeuren R, Tuttle RM, Bylund CL. Treatment decision making in early-stage papillary thyroid cancer. Psychooncology 2018 Jan;27(1):61-68. [doi: 10.1002/pon.4383] [Medline: 28124394]
23. Abel GA, Cronin AM, Earles K, Gray SW. Accessibility and Quality of Online Cancer-Related Clinical Trial Information for Naïve Searchers. Cancer Epidemiol Biomarkers Prev 2015 Oct;24(10):1629-1631 [FREE Full text] [doi: 10.1158/1055-9965.EPI-15-0274] [Medline: 26265204]

XSL•FO
RenderX

24. Storino A, Castillo-Angeles M, Watkins AA, Vargas C, Mancias JD, Bullock A, et al. Assessing the Accuracy and Readability of Online Health Information for Patients With Pancreatic Cancer. JAMA Surg 2016 Sep 01;151(9):831-837. [doi: 10.1001/jamasurg.2016.0730] [Medline: 27144966]

Abbreviations

CCTT: Clinical Center for Targeted Therapy
CIU: cancer-related internet users
NCI: National Cancer Institute

Edited by G Eysenbach; submitted 08.03.18; peer-reviewed by M Tkáč, E Neter; comments to author 25.04.18; revised version received 26.05.18; accepted 29.05.18; published 14.03.19.

Please cite as:
George GC, Iwuanyanwu EC, Buford AS, Piha-Paul SA, Subbiah V, Fu S, Karp DD, Pant S, Hinojosa CO, Hess KR, Cleeland CS, Bernstam EV, Meric-Bernstam F, Hong DS
Cancer-Related Internet Use and Its Association With Patient Decision Making and Trust in Physicians Among Patients in an Early Drug Development Clinic: A Questionnaire-Based Cross-Sectional Observational Study
J Med Internet Res 2019;21(3):e10348
URL: https://www.jmir.org/2019/3/e10348/
doi:10.2196/10348
PMID:30869638

©Goldy C George, Eucharia C Iwuanyanwu, Adrianna S Buford, Sarina A Piha-Paul, Vivek Subbiah, Siqing Fu, Daniel D Karp, Shubham Pant, Christina O Hinojosa, Kenneth R Hess, Charles S Cleeland, Elmer V Bernstam, Funda Meric-Bernstam, David S Hong. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Peer Adoption and Development of Health Innovations by Patients: National Representative Study of 6204 Citizens

Pedro Oliveira[1,2], PhD; Leid Zejnilovic[3], PhD; Salomé Azevedo[2], MSc; Ana Maria Rodrigues[4,5,6], MD, PhD; Helena Canhão[4,5,7], MD, PhD

[1]Copenhagen Business School, Copenhagen, Denmark

[2]Católica-Lisbon School of Business and Economics, Lisbon, Portugal

[3]Nova School of Business and Economics, Nova University, Lisbon, Portugal

[4]CEDOC, EpiDoC Unit, Nova Medical School, Nova University, Lisbon, Portugal

[5]Associação EpiSaúde, Évora, Portugal

[6]Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal

[7]National School of Public Health, Nova University, Lisbon, Portugal, Lisbon, Portugal

Corresponding Author:
Pedro Oliveira, PhD
Copenhagen Business School
Department of Strategy and Innovation
Kilen, Kilevej 14, DK-2000 Frederiksberg
Copenhagen, DK-2000
Denmark
Phone: 45 38152578
Email: po.si@cbs.dk

Abstract

Background: There is growing evidence that many patients and caregivers innovate by developing new solutions to cope with their health disorders. Given the easy access to vast internet resources and peers globally, it is increasingly important to understand what may influence user innovation and its adoption in health for improving individual well-being and ensuring their safety, in particular, how interactions with peers and physicians or search behavior, along with sociodemographics, may influence the decision to develop a solution or adopt one developed by a peer.

Objective: The aim of this paper was to study the development and peer-to-peer adoption of user innovations in health care and identify individual-level factors associated with these processes.

Methods: Data were collected via computer-assisted phone survey from a large, random, and representative sample of adult residents in Portugal (N=6204). User innovation questions were added to 1 wave of an ongoing observational, longitudinal, population-based epidemiological study. By asking about individual innovation activity, the sample was split into 3 groups: (1) the developers of health-related solutions for own use (developers), (2) the adopters of solutions developed by other patients or caregivers (peer-to-peer adopters), and (3) the rest of the population. Within the last group, *intention to adopt* was measured and used as a proxy of future behavior. Regression analysis is used to test the associations.

Results: In the population considered in this paper, an estimated 1.3% (75/6008) reported having developed a solution for own use and 3.3% reported to have adopted a solution developed by peers. The 3 groups (developers, adopters, and remaining population) have distinctive characteristics. Gender plays an important role in the solution development, as women are less likely to develop one (odds ratio [OR] 0.4, 95% CI 0.20-0.81; $P<.05$). Education is positively associated with the development activity (OR 1.13, 95% CI 1.03-1.24; $P<.05$) but also with the intentions to adopt a peer-developed solution. Search for health-related information is positively associated with the development, adoption, and the intentions to adopt a solution. Interactions with peers over the internet are rare, but in-person interactions are frequent and have a positive association with the dependent variables in all 3 groups. The results also suggest that trust in doctors represents an important dimension that shapes the attitudes of the population toward peer-developed solutions.

Conclusions: This paper demonstrates the importance of the peer community, doctor-patient relationship, citizen's search for information on innovation, and individual attitudes toward peer-to-peer adoption in health care. It stresses the need for a reliable

XSL•FO
RenderX

Web-based health-related information and the necessity to deeper understand complex relationships between the need to improve health and fulfill the need and the perception of the health care system.

(J Med Internet Res 2019;21(3):e11726) doi:10.2196/11726

KEYWORDS

citizen; patient; innovation; therapeutics; health; development; adoption; physician-patient relations; social interactions; search engine

Introduction

Background

User innovators are the ones who have developed a new good or service or modified an existing good or service for own use; they differ from producer innovators for whom profit is the dominant motivation to innovate [1]. Innovation scholars have demonstrated the existence of this empirical phenomenon in numerous industries, including health care [2]. Survey evidence from measurement studies of user innovation at a national level estimated that up to 0.5% of citizens in the United States, Japan, Finland, and the United Kingdom modify or create new products and services for personal health care–related use [3,4].

The largest group of user innovators in health care are *patient innovators*, patients or their nonprofessional caregivers (eg, parents and family members), who modify or develop a treatment, a technical aid product, or a medical device to cope with a health condition [5]. Besides developing, they may also share or adopt solutions developed by other peers, they organize themselves in communities, and either individually and or jointly solve problems, and even do limited trials with solutions they develop [6-9]. A study conducted in a population of rare disease patients and their nonprofessional caregivers showed that the frequency of user innovation might be higher among those afflicted with rare diseases than in the general population. The authors reported that 36% of interviewed survey respondents had developed a solution that was new to them, and 8% of the interviewed survey respondents had developed solutions that may be novel to the medical practice [6].

When the value of patient-developed solutions is considered, there is evidence of patients reporting significant improvements in the quality of life after using their self-developed solutions [6]. Furthermore, the study by Oliveira and Canhão [10] identified successful cases of novel patient-developed solutions that made a significant impact on medical practice. Health benefits and social value of innovations can only be achieved when innovations diffuse—when they are adopted and used by other people. In national-level surveys, the fraction of diffused user innovations observed varied from 5% to 17%, with the most common diffusion pathway being peer-to-peer exchange [3,4,11]. In the rare disease study, 32% of solutions reported by the patients and caregivers were shared with others—almost double the highest diffusion incidence observed in the general population—but only 5% shared the information with their doctor [6]. The existing evidence strongly suggests that the innovation and diffusion activity by patient innovators is significant but mostly hidden from the traditional health care system. The adoption side of the peer-to-peer innovations in health has received little or no attention from academia.

Rational and Aim

Innovation and adoption activity by patients and caregivers in health care may be strongly influenced by health care–related and sociotechnological contextual factors. However, no prior work systematically explored such relationships. For example, we know that people invest significant efforts to search for health-related information (online and offline) [12] and may have well-developed strategies for evaluating the credibility of the information [13]. However, we do not know the relationship of such a search with the innovation or peer-to-peer adoption activity. Other sociotechnical and health care contextual factors of interest may also include characteristics of the peer networks among patients [14,15], doctor-patient relationships [15-17], personal responsibility for health, or the trust in the availability of scientific breakthroughs for their health disorder. This work is, to the best of our knowledge, the first to systematically explore the relationships between these contextual factors and health-related peer-to-peer innovation and adoption activity among citizens.

Methods

Research Design

The data used in this paper are survey responses from a random and representative sample of adult residents in Portugal (N=6204), collected via computer-assisted phone survey conducted by NOVA Medical School. Professional interviewers were additionally trained by a psychologist to communicate the innovation questions, and 2 of the authors trained them to fill in the survey responses in a computer program during the conversation.

The innovation section of the survey, which is the focus of this paper, was integrated into a larger project, the second wave of a longitudinal, prospective, observational, population-based study named Epidemiology of Chronic Diseases (EpiDoC) (Figure 1) [18]. The first wave, entitled EpiDoC 1 Epidemiology of Rheumatic Diseases Study (EpiReumaPt), was rigorously designed to gather a representative random sample of residents in Portugal. In this phase, data were collected by face-to-face interviews of 10,661 individuals in the period from 2011 to 2013 [19]. Inclusion criteria were as follows: (1) Portuguese speaking individuals, (2) aged 18 years or older, (3) noninstitutionalized (excluding hospital or nursing homes, military barracks, and prisons), (4) for whom cognitive and physical impairments did not prevent completion of the survey, and (5) who were living in a private household in the country.

Participants were selected through a process of multistage random sampling. The sample was stratified according to the Portuguese statistic regions in the 2001 Census and the size of

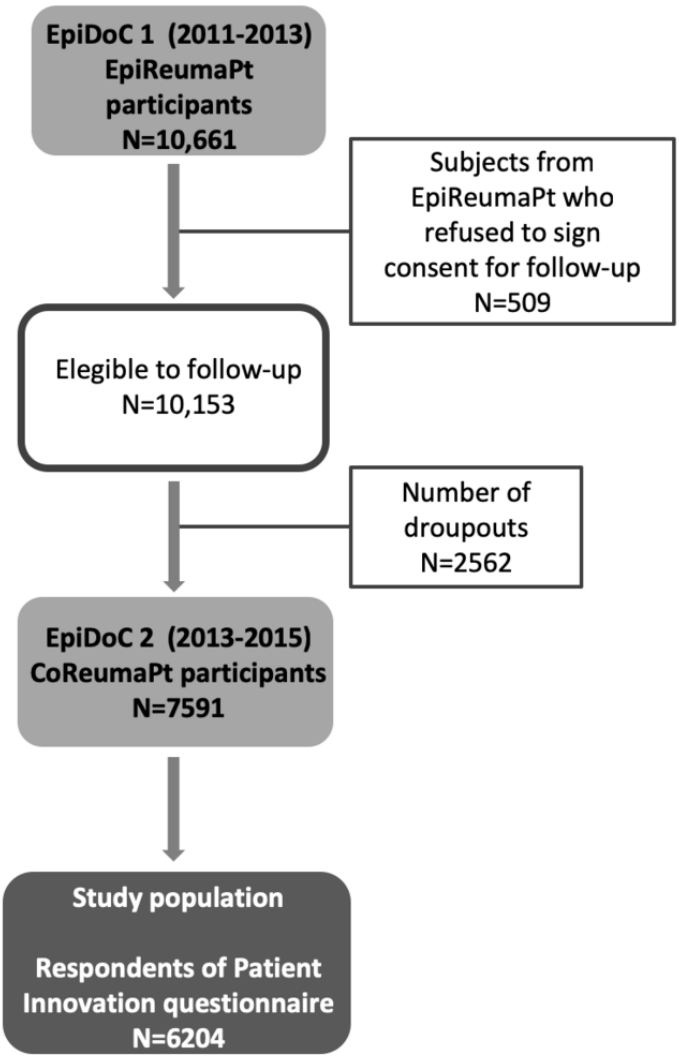
XSL•FO
RenderX

the population (<2000; 2000-9999; 10,000-19,999; 20,000-99,999; and ≥100,000 inhabitants). The number of participants of each stratum was proportional to the actual distribution of the population. In Madeira and the Azores, the sample size was increased (oversampling) to allow separate analyses in these regions. Candidate households were selected through a random route process; sampling points were randomly selected on the maps of each locality, where the interviewer began a systematic step count (defined for each locality according to its size), granting each household and everyone an equal probability of being chosen [19].

Most of the EpiDoc wave participants (10,153) also agreed to integrate into a prospective cohort and be contacted in the next round of surveying (EpiDoc 2)—the cohort of rheumatic diseases (CoReumaPt) wave (2013-2015). The Portuguese National Commission for Data Protection and the NOVA Medical School Ethics Committee have approved both EpiReumaPt and CoReumaPt [19]. The participants provided informed consent to participate in all phases of the study, and the study was conducted in accordance with the Declaration of Helsinki. Professional interviewers conducted all the interviews. In EpiDoC 2 (CoReumaPt), the follow-up phase, 7591 (out of 10,153) individuals completed a computer-assisted telephone survey. Of these, 6204 individuals were asked for and answered the innovation activity part of the survey. The questions about innovation activity were introduced 2 months after the launch of the follow-up study, which explains the difference in the number of respondents for this paper.

To guarantee the representativeness of the sample in relation to the Portuguese population (Mainland and *Madeira* and *Azores* islands), extrapolation weights were computed and used in statistical analysis. The weights were obtained by calibrating the extrapolation weights originally designed for the EpiDoC 1 (EpiReumaPt) sample. Participants and nonparticipants of the EpiDoC 2 (CoReumaPt) study were compared regarding their sociodemographic, socioeconomic, and health status characteristics. Weights were then adjusted based on this comparison and the stratification by statistical regions in Portugal, sex, and age groups [18].

Figure 1. Flowchart of the population-based study named Epidemiology of Chronic Diseases (EpiDoC). The first wave (2011-2013) is entitled EpiDoC 1 - Epidemiology of rheumatic diseases study (EpiReumaPt). The second wage (2013-2015) is entitled EpiDoc 2 - Cohort of rheumatic diseases (CoReumaPt).



XSL•FO
RenderX

Table 1. Scales, sources, and factor loadings.

Construct	Items	Factor loadings (N=6204)
Self-responsibility for health (2 items, adapted from a scale by Hibbard et al [22])	"It is me, more than any other person, who is responsible for my health and well-being." (5-point Likert scale: 1=totally disagree, 5=totally agree)	0.9
Self-responsibility for health (2 items, adapted from a scale by Hibbard et al [22])	"The most important factor that influences my well-being and health is my active role and responsibility for my health." (5-point Likert scale: 1=totally disagree, 5=totally agree)	0.9
Trust in medical doctor (Reduced scale proposed by Anderson et al [23])	"I trust my doctor so much that I always try to follow his/her advice." (5-point Likert scale: 1=totally disagree, 5=totally agree)	0.9
Trust in medical doctor (Reduced scale proposed by Anderson et al [23])	"If my doctor tells me something is so, then it must be true." (5-point Likert scale: 1=totally disagree, 5=totally agree)	0.9
Trust in medical doctor (Reduced scale proposed by Anderson et al [23])	"I feel my doctor does everything he/she should for my medical care." (5-point Likert scale: 1=totally disagree, 5=totally agree)	0.8
Perceptions of medical science frontiers (new)	Do you believe that the medical science can treat your disease? (5-point Likert scale: 1=not at all, 5=completely trust)	0.9
Perceptions of medical science frontiers (new)	How likely is it that the medical science can successfully treat you for your disease? (5-point Likert scale: 1=not at all, 5=it certainly can)	0.9
Intention to adopt a patient-developed solution [20]	How likely is it that you would use a solution developed by another patient to help you cope with your ailment? (5-point Likert scale: 1=very unlikely, 5=I would definitely)	0.9
Intention to adopt a patient-developed solution [20]	Do you intend to use a solution developed by another patient to help you cope with your ailment? (5-point Likert scale: 1=I do not intend to use, 5=I definitely intend to use one)	0.9

Survey Flow and 3 Groups of the Respondents

The first group of questions in the innovation section of the survey measured contextual factors, self-responsibility for health management, search for health information, the frequency of online and face-to-face interactions with peers, and trust in medical doctors and medical science. Next, respondents answered a question that split the sample into 3 groups: (1) the developers of solutions to cope with their health disorders, (2) adopters of health-related solutions developed by other patients or caregivers, and (3) the remaining population.

The question asked if an individual had developed a health-related solution or adopted a health-related solution developed by other patients. In the case of an affirmative response, the survey continued with sections that focused on the details about solution development or adoption, dividing the population into the developer or adopter groups. The third group, those who neither innovated nor adopted a solution, were asked whether they have ideas about potential solutions for health-related problems they so far encountered. Those who neither developed nor adopted a patient-developed solution were also asked about their intentions to adopt a patient-developed solution. For all the respondents, the survey started with sociodemographic questions and ended with life habits, functional, and quality of life questionnaires.

Creative Activity: Solution Development or Adoption of a Patient-Developed Solution

Questions regarding user innovation were built upon a questionnaire used in user innovation measurement surveys [3], adapting it to the health care context. As the conversation was phone-based, the calling party explicitly introduced the purpose of the innovation-related questions. The interviewers asked the respondents if they had, in their free time, done anything that would help them or someone close to them to cope with their health disorders. To ease the interpretation, we provided mental cues of what the potential solutions may be. The cues suggested to the respondents were medical aid instrument, medical dispositive, behavioral strategy (eg, a diet or an exercise plan), tools for everyday life at home or work, solutions related to one's appearance, medication or a combination of drugs, and natural products.

The question about creative activity is formulated to ask about newly developed or modified solutions and about the adoption of a patient-developed solution for personal use or for someone close to the respondent. From the survey responses on this question, 2 dependent (binary) variables were created, a solution development variable (developer) and a solution adoption variable (adopter). Furthermore, the origin of the advice for the adopter was asked for, to ensure that the source of the solution is a patient/caregiver. The objective was to identify the characteristics of those who had engaged in creative activity, regardless of the artifact's quality that is developed or adopted.

A large share of the population is likely to be neither solution developers nor adopters of peer-developed solutions, as not everyone has a need for a solution to cope with health-related issues. For this group, and to study the drivers of the attitude toward peer-developed solutions in a general population, the theory of planned behavior [20] is used. According to the theory, intentions are a relatively good proxy of future behaviors [21]. The behavioral intentions to adopt a patient-developed solution in this paper are measured using a 2-item scale (Table 1). Note that because of the survey complexity, the groups are exclusive,

XSL•FO
RenderX

which means that the developers were neither asked if they had also adopted nor about their intentions to adopt a solution. Furthermore, the adopters were not asked about their intentions to adopt a solution.

Health Care Contextual Factors

Considering the earlier stated goal of this paper, learning about individual creative activity in the health care context, a set of questions was added to the survey. To learn about individual search efforts, the survey measured the depth of search for health information and health-related solutions as the average weekly time spent searching. As social interactions among patients may influence adoption or intentions to adopt a patient-developed solution, the interviewees were asked about the frequency of their interactions with individuals who are afflicted with the same health disorder or who share interests in the disorder; 5-point Likert scales were used to represent different levels of frequency of interactions.

To measure the perceptions of the scientific frontier, trust in medical doctors, and the attitude toward personal health management, where possible, existing scales were used (reported in Table 1).

The medical part of the survey, pertinent to the epidemiological study, included standard measurement instruments that assessed health and quality of life. In this paper, the EuroQoL-5D (EQ-5D) score [24] was used, as it is a validated, reliable, and short standard health state instrument that suits the context well. As we had 2 measures of EQ-5D score, 1 from EpiReumaPt (2011-2013) and the other from CoReumaPt (2013-2015), a variable that represents the difference between the 2 EQ-5D scores has been generated. Note that this variable is used only in the context of future activities and is included only in the model that predicts the intentions to adopt a solution. To assess who is afflicted with a health disorder, a binary variable is generated from self-reported data, indicating which individual has a clinically diagnosed chronic noncommunicable disease. The list included the following groups of diseases/health disorders: diabetes, pulmonary disease, cardiac disease, gastrointestinal disease, neurologic disease, mental disease, neoplastic disease, thyroid and parathyroid disease, and rheumatic disease. A person is considered ill if there was a report of having at least one disorder from the list (Multimedia Appendix 1).

Statistical Analysis

Given that measurement instruments were used, the survey was pretested on 106 randomly selected interviewees. In this step, exploratory factor analysis is conducted to test whether theoretically constructed 4 factors could be identified and if there is a sufficient level of internal consistency.

Descriptive statistics are reported for the full sample (N=6204) after applying probability weights to obtain the population estimates [19,25]. Statistical software, StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC, was used to conduct multivariate regression on survey data with the probability weights. The objective of the analysis is to explore the existence of statistically significant associations between the outcomes and independent variables. A total of 3 sets of analysis are conducted for the 3 groups. For the first 2 groups, developers and adopters, logistic regression is used to study the group differences. For the third group, the ones who neither developed nor adopted a peer-developed solution, the sample was divided in 2 subpopulations: (1) patients with at least one chronic noncommunicable disease and (2) healthy population. For both subsamples, ordinary least square models were used to explore the associations of the independent variables with the intentions to adopt. In all models, the threshold of statistical significance is set to $P<.05$.

The scales for self-responsibility for health, trust in doctors, and perceptions of medical science frontiers are included as standardized values, and the interpretation of the coefficients should be in terms of the SDs from the population mean.

Results

Sociodemographic Characteristics, Health Status, and Peer Interactions

The results of exploratory factor analysis on the initial sample of 106 individuals suggest that the items load well on the 4 factors, and all the factors have high internal consistency (alpha≥.7) [22]. In Table 1, factor loadings are reported for the entire sample.

Descriptive statistics are reported in Table 2 for the 3 groups (developers, adopters, and remaining population). In total, 6008 responses were included after removing responses of the individuals who could not answer the creative activity–related question. For categorical variables, the absolute count is provided, together with population estimates percentages in the brackets. For continuous variables, means and SDs are population estimates.

The results show that 1.3% of the population reported being developers and 3.3% peer-to-peer adopters. The respondents have on average 9 years of formal education (SD 4 years), and 49% reported being diagnosed with at least one noncommunicable chronic disease.

There are notable differences among the 3 groups along several dimensions. Within the developers' group, males represent the majority (66%). Furthermore, unemployment or temporal disability/retirement among the developers (54%) is higher than that among adopters (39%) or the remaining population (35%). The developers have, on average, 1 more year of education than the adopters. For all 3 groups, interacting with peers (patients/caregivers) via the internet is rare, and the remaining population (neither developers nor adopters) are more active in that regard, with 2% more active people than that in the other 2 groups. Majority of the developers and adopters have frequent in-person interactions, 65% and 53%, respectively. The adopters have a higher number of comorbidities, 2.1 compared with 1.9 for developers and 1.5 for the rest of the population, on average. Although all 3 groups have left-skewed self-responsibility for health (4.9 out of 5), developers are more active than others, as 64% exercise regularly compared with around 40% in the other 2 groups. All 3 groups have high trust in doctors, with a marginally higher value for the remaining population, 4.5 compared with 4.3 out of 5. Perception of medical science

XSL•FO
RenderX

frontier is also left-skewed, with average values of 3.5 for the adopters and 3.8 for the other 2 groups.

Absolute values of correlations between independent variables (correlation matrix available upon request) were below .35, with 4 exceptions. These exceptions were (1) age and education (r=−.6), (2) age and having at least one chronic disease (r=.41), (3) having at least one chronic disease and the quality of life score (r=−.38), and (4) quality of life score and education (r=.38). However, at these values, the correlation listed above are not considered problematic regarding multicollinearity.

Table 2. Descriptive statistics for the 3 groups (Innovator, Adopter, Remaining Population).

Population characteristics	Innovator (n=75)	Adopter (n=210)	Remaining population (n=5723)
Gender (female), n (%)	40 (34.0)	172 (58.4)	3146 (47.6)
Age (years), mean (SD)	44.62 (14.40)	49.53 (17.09)	46.41 (17.85)
Years of education, mean (SD)	9.36 (3.28)	8.33 (3.93)	8.92 (3.81)
Employment status, n (%)			
Employed full-time, part-time, or domestic worker	38 (45.8)	116 (60.9)	3193 (65.2)
Temporally work disabled/retired	24 (22.4)	70 (30)	1834 (24.9)
Unemployed	8 (31.7)	23 (9.1)	479 (9.9)
Portuguese Nomenclature of Territorial Units for Statistics, n (%)			
Norte	17 (33.1)	65 (32.5)	1883 (37.8)
Centro	23 (36.4)	49 (25.4)	1210 (23.7)
Lisboa	15 (17.2)	46 (28.2)	1135 (24.3)
Alentejo	5 (8.2)	12 (7.4)	264 (6.3)
Algarve	1 (2.2)	4 (3)	170 (3.6)
Azores	7 (1.4)	17 (1.5)	521 (2.0)
Madeira	7 (1.6)	17 (2)	540 (2.3)
Physical exercise at least once per week, n (%)	37 (63.9)	85 (40.0)	2390 (44.6)
Quality of life, EQ-5D[a] score—CoReumaPt[b], mean (SD)	0.71 (0.23)	0.73 (0.27)	0.80 (0.26)
EQ-5D score difference CoReumaPt[b]-EpiReumaPt[c], mean (SD)	−0.05 (0.26)	−0.09 (0.27)	−0.05 (0.24)
Number of chronic diseases, mean (SD)	1.81 (1.64)	2.07 (2.55)	1.51 (1.71)
Frequency of interaction with other patients or caregivers (face-to-face), n (%)			
Never	30 (36.2)	88 (48.8)	3181 (58.3)
Less than once a week	12 (10.1)	61 (26.9)	1298 (22.3)
At least once a week	33 (53.7)	60 (24.3)	1207 (19.4)
Interaction with other patients or caregivers (internet), n (%)			
No	69 (92)	248 (92)	5450 (96.1)
Yes	6 (8)	16 (8)	179 (3.9)
Depth of search: search time on health (hours per week), mean (SD)	1.97 (4.2)	1.09 (2.08)	0.47 (1.64)
Self-responsibility for health, mean (SD)	4.82 (0.28)	4.90 (0.30)	4.86 (0.41)
Trust in doctors scale, mean (SD)	4.31 (0.73)	4.35 (0.86)	4.47 (0.79)
Intentions to adopt, mean (SD)	—[d]	—	2.63 (0.3)
Perceptions of medical science frontier, mean (SD)	3.81 (0.85)	3.53 (1.12)	3.85 (0.94)

[a]EQ-5D: EuroQoL-5D.

[b]CoReumaPt: Cohort of rheumatic diseases (EpiDoC 2).

[c]EpiReumaPt: Epidemiology of rheumatic diseases study (EpiDoC 1).

[d]Not applicable.

XSL•FO
RenderX

Results of the Multivariable Analysis

Results of the multivariable analysis are shown in Table 3. A total of 3 dependent variables (developer, adopter, and intentions to adopt) correspond to the 3 groups of interest.

Developer

Considering the solution development for own use (model 1), the results show that women are less likely to develop a solution for own use (OR 0.40, 95% CI 0.20-0.81; *P*<.05). Education is positively associated with the development activity (OR 1.13, 95% CI 1.03-1.24; *P*<.05), and the developers are more likely to be unemployed than employed (OR 6.45, 95% CI 2.40-17.29; *P*<.01). Considering health care contextual factors, the developers are more likely to have face-to-face interactions with other patients or caregivers (once a month or more) than no interaction (OR 4.92, 95% CI 2.20-10.99; *P*<.01). Furthermore, they are more likely to have at least one chronic, noncommunicable disease (OR 2.85, 95% CI 1.30-6.27; *P*<.01) and to search for health information more intensely than the rest of the population (OR 1.15, 95% CI 1.04-1.26).

Adopter

The population of adopters is significantly different from the population of developers. Adoption (model 2) is weakly positively associated with female gender (OR 1.54, 95% CI 0.94-2.52; *P*<.10) and online interactions with other patients (OR 2.12, 95% CI 0.95-4.74; *P*<.1). Adopters are, like developers, more likely to invest time to search for health information than the remaining population (OR 1.11, 95% CI 1.03-1.20; *P*<.01).

Remaining Population—Intentions to Adopt

In models 3 and 4, the dependent variable is the intention to adopt a solution developed by a patient or a nonprofessional caregiver. Intentions, according to the theory of planned behavior, are a proxy for actual behavior. In this paper, they are interpreted as attitudes toward peer-developed solutions.

The results for the subsample of individuals with at least one chronic noncommunicable disease (model 3) suggest a distinct combination of statistically significant associations. Intentions to adopt are negatively associated with age (beta=−.01; 95% CI −0.02 to −0.01; *P*<.01) and positively associated with education (beta=.02; 95% CI 0.00-0.03; *P*<.05). Like developers and adopters, in the remaining population, those with a chronic disease with higher intentions to adopt are more likely to invest their time to search for health-related information (beta=.07; 95% CI 0.02-0.13; *P*<.05) and more likely to have frequent in-person interactions with other patients/caregivers. Unlike the other 2 groups, developers and adopters, doctor-patient relationship plays an important role; the lower the trust in doctor, the higher is the intention to adopt a peer-developed solution (beta=−.08; 95% CI −0.14 to −0.02; *P*<.01).

Within the remaining population, the subsample of individuals without a chronic disease is very similar to the subsample of those with a chronic disease. Distinctive characteristic of the former subgroup is a negative association between retirement/temporary work disability and the intentions to adopt a peer-developed solution (beta=−.23; 95% CI −0.41 to −0.04; *P*<.05; model 4).

XSL•FO
RenderX

Table 3. Multivariable analysis with population estimates.

Population characteristics	Developer (versus all the others); Model 1, odds ratio (95% CI)	Adopter (versus all the others); Model 2, odds ratio (95% CI)	Intentions to adopt (remaining population with a chronic disease); Model 3, beta estimates (95% CI)	Intentions to adopt (remaining population without a chronic disease); Model 4, beta estimates (95% CI)
Gender: female versus male	*0.40*[a] (0.20 to 0.81)	1.54[b] (0.94 to 2.52)	*− .12*[a] (−0.23 to −0.00)	0.01 (−0.10 to 0.12)
Age (years)	0.99 (0.96 to 1.02)	1.00 (0.99 to 1.02)	*− .01*[c] (−0.02 to −0.01)	*− 0.01*[c] (−0.01 to −0.00)
Education (in years)	*1.13*[a] (1.03 to 1.24)	0.95 (0.87 to 1.03)	*.02*[a] (0.00 to 0.03)	*0.03*[c] (0.01 to 0.04)
Employment: temporarily work disabled/retired versus employed (full- or part-time)	1.41 (0.53 to 3.79)	0.95 (0.53 to 1.72)	−.02 (−0.16 to 0.13)	*− 0.23*[a] (−0.41 to −0.04)
Employment: unemployed versus employed (full- or part-time)	*6.45*[c] (2.40 to 17.29)	1.02 (0.55 to 1.90)	.04 (−0.13 to 0.21)	−0.14 (−0.34 to 0.06)
Marital status (married or union versus single or widow or divorced)	2.46[b] (0.98 to 6.14)	1.20 (0.66 to 2.19)	−.00 (−0.13 to 0.12)	−0.02 (−0.14 to 0.09)
Regular physical exercise	*1.87*[a] (1.03 to 3.40)	0.85 (0.52 to 1.38)	.08 (−0.06 to 0.21)	0.03 (−0.08 to 0.14)
Health state, EQ-5D[d] score—CoReumaPt[e]	0.63 (0.25 to 1.62)	1.09 (0.30 to 4.03)	−.02 (−0.26 to 0.23)	−0.23 (−0.61 to 0.15)
Score Difference EQ-5D score: CoReumaPt[e]-EpiReumaPt[f]	—[g]	—	−.09 (−0.31 to 0.14)	−0.03 (−0.37 to 0.32)
Face-to-face Interaction with other patients/caregivers: less than once a month versus no interactions	0.68 (0.27 to 1.71)	1.24 (0.74 to 2.08)	*.27*[c] (0.13 to 0.40)	—
F2F Interaction with other patients/caregivers: once a month or more versus no interactions	*4.92*[c] (2.20 to 10.99)	1.21 (0.72 to 2.02)	*.27*[c] (0.13 to 0.42)	—
Online interactions with other patients	0.62 (0.12 to 3.14)	2.12[b] (0.95 to 4.74)	−.09 (−0.38 to 0.19)	—
Health information search depth (hours per week)	*1.15*[c] (1.04 to 1.26)	*1.11*[c] (1.03 to 1.20)	*.07*[a] (0.02 to 0.13)	−0.00 (−0.04 to 0.04)
Personal responsibility for health (standardized)	1.06 (0.73 to 1.56)	1.14 (0.90 to 1.44)	−.04[b] (−0.09 to 0.00)	0.04 (−0.04 to 0.11)
Trust in physician (standardized)	0.92 (0.72 to 1.18)	0.94 (0.80 to 1.09)	*− .08*[c] (−0.14 to −0.02)	*− 0.14*[c] (−0.19 to −0.09)
Perceptions of medical science frontier (standardized)	0.99 (0.71 to 1.37)	0.81 (0.63 to 1.05)	−.01 (−0.07 to 0.04)	−0.02 (−0.08 to 0.04)
With at least one disease versus no disease	*2.85*[c] (1.30 to 6.27)	1.02 (0.60 to 1.73)	—	—

[a]*P*<.05.

[b]*P*<.10.

[c]*P*<.01.

[d]EQ-5D: EuroQoL-5D.

[e]CoReumaPt - Cohort of rheumatic diseases (EpiDoC 2).

[f]EpiReumaPt – Epidemiology of rheumatic diseases study (EpiDoC 1).

[g]Not applicable.

Discussion

Principal Findings

The analysis suggests that solution development and adoption of peer-developed solutions are relatively infrequent but significant phenomena and that the 3 groups have distinctive characteristics. Population estimates of the share of solution developers, 1.3%, is over 2 times higher than the estimate of the share of health care–related innovation by citizens in the United Kingdom [4]. As there was no restriction on the novelty of the reported solutions, as it could not be established, the reported estimates in this paper are regarding solution developers and not user innovators. In other words, the comparison is not applicable. Regarding the population estimates of the adopters, this is the first time for such an estimate to be taken.

A series of results are aligned with the extant academic literature in user innovation. Regression results suggest that solution developers are more often men and educated individuals, confirming the findings from the study of user innovations by consumers in the United Kingdom [4]. Furthermore, our paper

XSL•FO
RenderX

showed that active interactions with peers are positively associated with solution development, which corroborates the findings of Hienerth and Lettl [26] who studied the influence of peer communities on user innovation. From a public policy perspective, if the goal is to stimulate solution development by patients and caregivers in health care, a meaningful investment could be in the development of communities of peers. An example of a successful investment is the *Enabling the Future* community where patients, caregivers, and community members come together to develop open-source models of 3-dimensional-printed hands [7]. Albeit, integrating knowledge and experience of users to improve health care is complex, and the effects often fail short of the expectations [27]. Although intuition may suggest internet as a great platform for communities, the results of this paper suggest that most of the respondents, developers, adopters, and even ill people from the remaining population who have higher intentions to adopt a peer-developed solution prefer in-person interactions among peers over contacts through the internet.

Our study showed that developers are more likely to be unemployed than employed. A plausible explanation may be that they have more time to reflect upon needs and solutions or that unemployment is associated to a higher likelihood of suffering from health disorders and to having financial difficulties [23], which implies the higher need to solve problems.

Group-mean comparison suggests that the average time spent searching for health-related information by the developers (2 hours/week) is almost twice the time spent by the adopters (1.1 hours/week) and 4 times higher than that for the rest of the population (0.5 hours/week). This result emphasizes the importance of the provision of curated and accurate information, especially when, following an advice of a peer without consulting a health professional may be quite dangerous. For example, applying a plant extract without understanding side effects or permitted dosages may provoke serious health issues.

Considering adoption of solutions developed by patients or caregivers, the regression results do not suggest any stark characteristic of the group of adopters. However, the application of the intentions to adopt, a concept from the theory of planned behavior, reveals an important association. The attitude of those who did not engage in neither developing a solution nor adopting one may be influenced by the doctor-patient relationship. In light of the safety concerns regarding the diffusion of (self-made) health solutions in informal communities of patients and caregivers, doctors are a vital element of the health care system that helps patients to establish safety and efficacy of the available solutions. A negative association between the intentions to adopt and age possibly reflects the generational change in the perception of the role of the conventional health care system. In particular, older individuals may be used to the paternalistic doctor-patient relationship, and they may put a higher value on the official source of health-related solutions. Education is positively associated with the intentions to adopt a solution, which is also potentially linked to the paradigm shift in health care from paternalistic to more egalitarian relationships between patients and health professionals.

Limitations, Strengths, and Further Research

In this paper, data have been collected from a prospective cohort; as we worked with cross-sectional data, only associations may be claimed. Recollection and interpretation bias may be present in the data. Although some people may have developed or adopted a patient-developed solution without being aware of it, the focus of this work was to explore the characteristics of those who are aware and have had chosen to develop or adopt a peer-developed solution. Hence, these biases are likely not to influence the results significantly.

A set of preemptive steps were taken before administering the survey to control for item-related (common method) bias, as suggested by Podsakoff et al [28]. These measures include ease of cognitive load on the individuals and the design of the questions and their order to avoid the item-related bias.

The advantage of the study is the size of the sample and the sampling design. As this study is conducted in 1 country, it casts doubt whether the results are generalizable to other cultural and health care policy settings.

Conclusions

This paper is the first-of-type exploratory analysis of creative activities in the general population that focuses on health care and takes into consideration health care contextual factors. It demonstrates distinctive characteristics of: (1) the patients and caregivers who are developers of solutions, (2) the adopters of peer-developed solutions, and (3) the attitudes of the remaining population. Two actionable takeaways from the study are the importance of supplying reliable health-related information to patients who are searching and of the investment in good doctor-patient relationships.

Treating patients as equals is becoming the new mantra in organized health care systems [29], and we need to consider carefully what it means regarding their creative work, knowledge contribution, and organization and delivery of medical care. Only when we understand and support the creative contributions of the patients, we will have a system that truly integrates them.

Acknowledgments

The authors are grateful to the Public Health Initiatives Programme (PT06), funded by EEA Grants Financial Mechanism 2009-2014. In addition, they are grateful for the funding provided by the Fundação para a Ciência e a Tecnologia (FCT) and the Carnegie Mellon Portugal Program through project Technology, Entrepreneurship and Innovation Policy Lab, CMUP-ERI/TPE/0028/2013. They are also thankful for help of the EpiReumaPt study group and the invaluable comments and wisdom of Professor Eric von Hippel (Massachusetts Institute of Technology).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Data description.

[XLSX File (Microsoft Excel File), 10KB - jmir_v21i3e11726_app1.xlsx]

References

1. von Hippel E. Economics of product development by users: the impact of “sticky” local information. Manage Sci 1998 May;44(5):629-644. [doi: 10.1287/mnsc.44.5.629]
2. von Hippel E. Open user innovation. In: Hall BH, Rosenberg N, editors. Handbook of the Economics of Innovation. Amsterdam: North Holland; Jan 2010:411-427.
3. de Jong JP, von Hippel E, Gault F, Kuusisto J, Raasch C. Market failure in the diffusion of consumer-developed innovations: patterns in Finland. Res Policy 2015 Dec;44(10):1856-1865. [doi: 10.1016/j.respol.2015.06.015]
4. von Hippel E, de Jong JP, Flowers S. Comparing business and household sector innovation in consumer products: findings from a representative study in the United Kingdom. Manage Sci 2012;58:1669-1681 [FREE Full text] [doi: 10.1287/mnsc.1110.1508]
5. Habicht H, Oliveira P, Shcherbatiuk V. User innovators: when patients set out to help themselves and end up helping many. Die Unternehmung 2012;66(3):277-294 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2144325.
6. Oliveira P, Zejnilovic L, Canhão H, von Hippel E. Innovation by patients with rare diseases and chronic needs. Orphanet J Rare Dis 2015 Apr 9;10:41 [FREE Full text] [doi: 10.1186/s13023-015-0257-2] [Medline: 25887544]
7. Zejnilovic L, Oliveira P, Canhao H. Innovations by and for patients, their place in the future health care system. In: Albach H, Meffert H, Pinkwart A, Reichwald R, von Eiff W, editors. Boundaryless Hospital. Berlin: Springer; 2016:341-357.
8. Canhão H, Oliveira P, Zejnilovic L. Patient innovation–empowering patients, sharing solutions, improving lives. New Eng J Med Catalyst 2016 Nov [FREE Full text]
9. Canhão H, Zejnilovic L, Oliveira P. Revolutionizing healthcare by empowering patients to innovate. Euro Med J Innov 2017;1:31-34 [FREE Full text]
10. Oliveira P, Canhho H. Users as service innovators: evidence across healthcare and financial services. In: Lakhani KR, Harhoff D, editors. Revolutionizing Innovation: Users, Communities, and Open Innovation. Cambridge: MIT Press; 2016.
11. Ogawa S, Pongtanalert K. Visualizing invisible innovation continent: evidence from global consumer innovation surveys. SSRN Journal 2011. [doi: 10.2139/ssrn.1876186]
12. 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]
13. 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]
14. Earp J, French EA, Gilkey MB. e-patients: how they can help us heal healthcare. In: Patient Advocacy For Health Care Quality: Strategies For Achieving Patient-centered Care. Burlington, Massachusetts: Jones & Bartlett Learning; 2007.
15. DeBronkart ED. Let patients help! A patient engagement handbook. United States: Createspace Independent Publishing Platform; 2013.
16. Dwyer AA, Quinton R, Morin D, Pitteloud N. Identifying the unmet health needs of patients with congenital hypogonadotropic hypogonadism using a web-based needs assessment: implications for online interventions and peer-to-peer support. Orphanet J Rare Dis 2014 Jun 11;9:83 [FREE Full text] [doi: 10.1186/1750-1172-9-83] [Medline: 24915927]
17. Harvey K. The evidence for complementary medicines. Health Voices 2012 Nov;11:6 [FREE Full text]
18. Dias SS, Rodrigues AM, Gregório MJ, de Sousa RD, Branco J, Canhão H. Cohort profile: the epidemiology of chronic diseases cohort (EpiDoC). Int J Epidemiol 2018 Dec 1;47(6):1741-172j [FREE Full text] [doi: 10.1093/ije/dyy185] [Medline: 30212889]
19. Rodrigues A, Gouveia N, da Costa LP, Eusébio M, Ramiro S, Machado P, et al. EpiReumaPt- the study of rheumatic and musculoskeletal diseases in Portugal: a detailed view of the methodology. Acta Reumatol Port 2015;40(2):110-124 [FREE Full text] [Medline: 26219965]
20. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
21. Sheeran P, Milne S, Webb TL, Gollwitzer PM. Semantic Scholar. 2005. Implementation Intentions and Health Behaviour URL: https://pdfs.semanticscholar.org/14d6/f90615d3506f92f23c50a8d1b57fc8289c87.pdf [accessed 2019-02-26] [WebCite Cache ID 76TaTkd3i]
22. Hibbard J, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]

XSL•FO
RenderX

23. Andersen I, Brønnum-Hansen H, Kriegbaum M, Hougaard C, Hansen FK, Diderichsen F. Increasing illness among people out of labor market - a Danish register-based study. Soc Sci Med 2016 May;156:21-28. [doi: 10.1016/j.socscimed.2016.03.003] [Medline: 27017087]
24. Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10158943]
25. Gregório MJ, Rodrigues A, Eusébio M, Sousa R, Dias S, André B, et al. Dietary patterns characterized by high meat consumption are associated with other unhealthy life styles and depression symptoms. Front Nutr 2017;4:25 [FREE Full text] [doi: 10.3389/fnut.2017.00025] [Medline: 28660194]
26. Hienerth C, Lettl C. Exploring how peer communities enable lead user innovations to become standard equipment in the industry: community pull effects. J Prod Innov Manag 2011 Oct 13;28(s1):175-195. [doi: 10.1111/j.1540-5885.2011.00869.x]
27. Serapioni M, Matos AR. Citizen participation and discontent in three Southern European health systems. Soc Sci Med 2014 Dec;123:226-233. [doi: 10.1016/j.socscimed.2014.06.006] [Medline: 24947551]
28. Podsakoff PM, MacKenzie SB, Podsakoff NP. Sources of method bias in social science research and recommendations on how to control it. Annu Rev Psychol 2012 Jan;63:539-569. [doi: 10.1146/annurev-psych-120710-100452] [Medline: 21838546]
29. Barnett S, Katz A. Patients as partners in innovation. Semin Pediatr Surg 2015 Jun;24(3):141-144. [doi: 10.1053/j.sempedsurg.2015.02.014] [Medline: 25976152]

Abbreviations

CoReumaPt: Cohort of Rheumatic Diseases
EpiDoC: Epidemiology of Chronic Diseases
EpiReumaPt: Epidemiology of Rheumatic Diseases Study
EQ-5D: EuroQoL-5D
OR: Odds ratio

Edited by G Eysenbach; submitted 30.07.18; peer-reviewed by A Dwyer, E Lettieri, S Chen; comments to author 08.10.18; revised version received 23.12.18; accepted 23.01.19; published 26.03.19.

Please cite as:
Oliveira P, Zejnilovic L, Azevedo S, Rodrigues AM, Canhão H
Peer Adoption and Development of Health Innovations by Patients: National Representative Study of 6204 Citizens
J Med Internet Res 2019;21(3):e11726
URL: http://www.jmir.org/2019/3/e11726/
doi:10.2196/11726
PMID:30912748

©Pedro Oliveira, Leid Zejnilovic, Salomé Azevedo, Ana Maria Rodrigues, Helena Canhão. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design and Preliminary Findings From a New Electronic Cohort Embedded in the Framingham Heart Study

David D McManus[1,2,3], MD, ScM; Ludovic Trinquart[4,5], PhD, MPH; Emelia J Benjamin[4,6,7], MD, ScM; Emily S Manders[4], BS; Kelsey Fusco[4], BS; Lindsey S Jung[5], MS; Nicole L Spartano[4,8], PhD; Vik Kheterpal[9], MD; Christopher Nowak[9], BA; Mayank Sardana[10], MBBS; Joanne M Murabito[4,11], MD, ScM

[1]Cardiology Division, Department of Medicine, University of Massachusetts Medical School, Worcester, MA, United States

[2]Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

[3]Department of Medicine, University of Massachusetts Medical School, Worcester, MA, United States

[4]Boston University's and National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, United States

[5]Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States

[6]Section of Preventive Medicine and Epidemiology and Section of Cardiovascular Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, United States

[7]Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States

[8]Section of Endocrinology, Diabetes, Nutrition, and Weight Management, Boston University School of Medicine, Boston, MA, United States

[9]Care Evolution, Ann Arbor, MI, United States

[10]Cardiology Division, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[11]Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, United States

Corresponding Author:
Joanne M Murabito, MD, ScM
Boston University's and National Heart, Lung, and Blood Institute's Framingham Heart Study
73 Mount Wayte Ave
Framingham, MA, 01701
United States
Phone: 1 508 935 3400
Email: murabito@bu.edu

Abstract

Background: New models of scalable population-based data collection that integrate digital and mobile health (mHealth) data are necessary.

Objective: The aim of this study was to describe a cardiovascular digital and mHealth electronic cohort (e-cohort) embedded in a traditional longitudinal cohort study, the Framingham Heart Study (FHS).

Methods: We invited eligible and consenting FHS Generation 3 and Omni participants to download the electronic Framingham Heart Study (eFHS) app onto their mobile phones and co-deployed a digital blood pressure (BP) cuff. Thereafter, participants were also offered a smartwatch (Apple Watch). Participants are invited to complete surveys through the eFHS app, to perform weekly BP measurements, and to wear the smartwatch daily.

Results: Up to July 2017, we enrolled 790 eFHS participants, representing 76% (790/1044) of potentially eligible FHS participants. eFHS participants were, on average, 53±8 years of age and 57% were women. A total of 85% (675/790) of eFHS participants completed all of the baseline survey and 59% (470/790) completed the 3-month survey. A total of 42% (241/573) and 76% (306/405) of eFHS participants adhered to weekly digital BP and heart rate (HR) uploads, respectively, over 12 weeks.

Conclusions: We have designed an e-cohort focused on identifying novel cardiovascular disease risk factors using a new smartphone app, a digital BP cuff, and a smartwatch. Despite minimal training and support, preliminary findings over a 3-month follow-up period show that uptake is high and adherence to periodic app-based surveys, weekly digital BP assessments, and smartwatch HR measures is acceptable.

(J Med Internet Res 2019;21(3):e12143) doi:10.2196/12143

XSL•FO
RenderX

KEYWORDS

smartphone; tele-medicine; blood pressure monitoring; ambulatory; cohort studies

Introduction

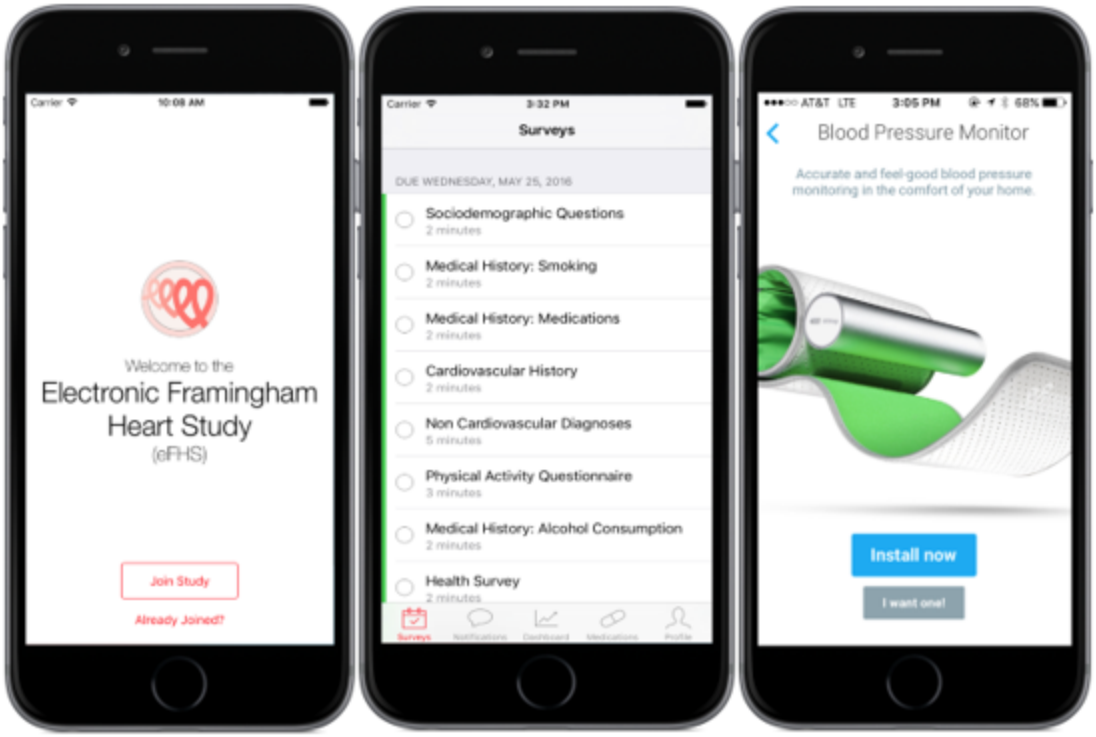
Background

Despite its great contributions to our understanding of cardiovascular disease (CVD) and its risk factors, traditional cohort epidemiology can be costly and time consuming [1-3], and it has been recently criticized for a lack of integration of "*real world* health information, such as activity level, heart rate (HR) , or home blood pressure (BP) [4]. Recommendations for launching digital epidemiology have been put forth by a working group from the National Heart, Lung, and Blood Institute on Epidemiology and Population Science with the need for evaluation of scalability and reliability of digital and mobile health (mHealth) data [5]. The rapid evolution of commercial mHealth technologies presents a transformational opportunity for CVD phenotyping outside of the clinical and Research Center settings [6-8]. Nevertheless, to date, electronic cohorts (e-cohorts) have largely enrolled middle-aged and younger volunteers with a low prevalence of CVD and its risk factors [9]. With the National Institutes of Health's *All of Us* Program, plans to engage substantial numbers of participants electronically are evolving [10], but, to date, a few ongoing studies have successfully enrolled participants in mHealth research [9,11]. Furthermore, little is known about how mHealth and digital CVD phenotypes, such as activity level, relate to other CVD risk factors and outcomes in middle-aged and older adults [12].

Objectives

In this paper, we describe the design of the electronic Framingham Heart Study (eFHS). The broad objective of the eFHS is to add new mobile and digital phenotypes into the rich Framingham Heart Study (FHS) and leverage its deep CVD risk factor phenotyping, longitudinal follow-up, and established relationships with its participants. We plan to compare novel measures obtained from commonly used commercial mHealth and digital devices with other gold standard CVD phenotypes obtained in the FHS Research Center as well as CVD outcomes. We have developed a custom smartphone app, the *eFHS* app, (see Figure 1), that is capable of messaging participants, administering surveys to collect health history and behavior updates, and pulling data from co-deployed wearables, including a digital BP cuff and smartwatch. The research center staff are deploying the eFHS app and devices and training participants on system components in the context of their scheduled and ongoing FHS Research Center examinations. In addition to describing our approach and methodology, we report the characteristics of the eFHS participants, rates of adherence to the eFHS survey, and digital BP and smartwatch research protocols over the 3 months since deployment.

Figure 1. Screenshots of the electronic Framingham Heart Study app loading screen, questionnaire, and interface with digital health device (blood pressure monitor).



XSL•FO
RenderX

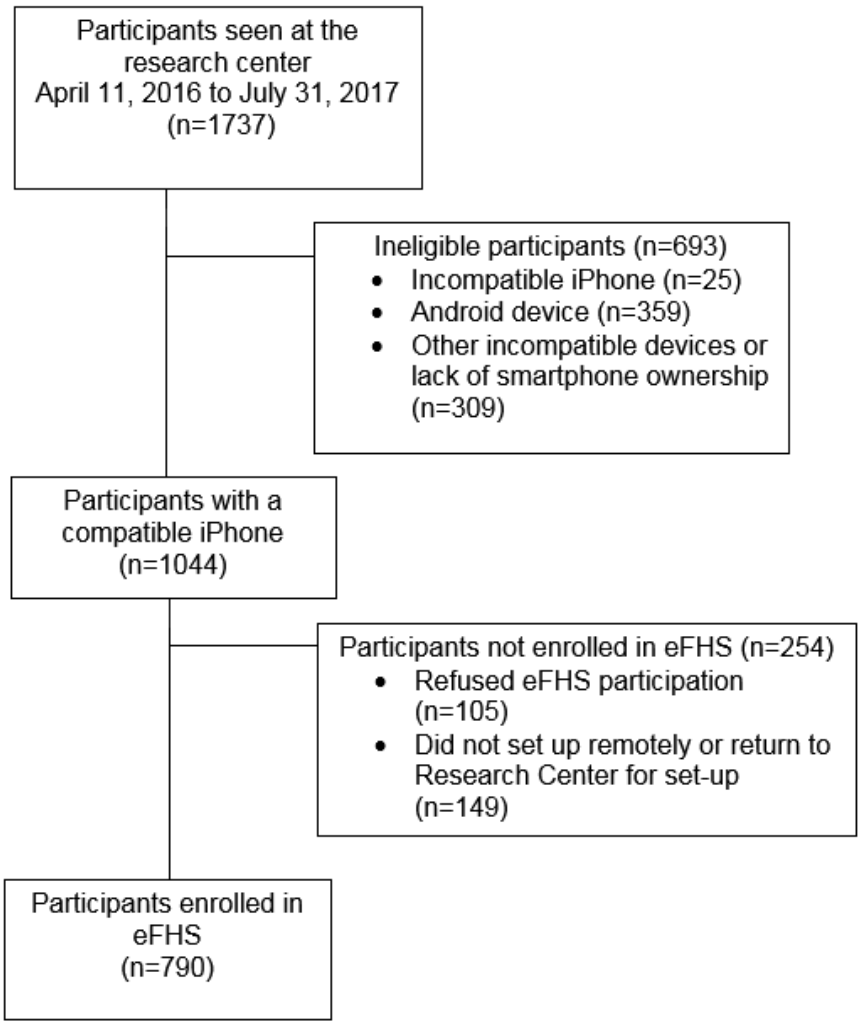
Methods

Study Cohort

The eFHS is enrolling participants from the Framingham Heart Study (FHS) Third Generation Cohort (Gen 3, total n=4095), multiethnic Omni Group 2 Cohort (Omni 2, total n=410), and New Offspring Spouse Cohort (NOS, n=103) enrolled from 2002 to 2005, who are examined every 6 to 8 years, with details previously reported [13]. eFHS is leveraging the in-person NOS Examination 3 in the Omni 2 in the FHS Gen 3 that began in April 2016 in the FHS Research Center [14]. eFHS began enrollment on June 20, 2016, with the smartphone app and a digital BP cuff for weekly home BP monitoring. A smartwatch for daily activity and HR monitoring was added on November 17, 2016. The eligible participants were invited to participate in eFHS as 1 component of the routine research center examination, and they were allowed to decline the use of both digital devices. To be eligible for eFHS, English speaking participants were required to own an iPhone with a compatible iOS (version 9 or higher), reside in the United States, and be willing to provide permissions for notifications and data sharing with the Research Center. All participants provided informed consent to participate in the eFHS, both as part of the overall consent for Examination 3 and an electronic consent within the eFHS app itself (2-level consent, the consent forms are available on the FHS website [15,16]. The eFHS protocol was reviewed and approved by the FHS Executive Committee and the Institutional Review Board at the Boston University Medical Center.

As of June 2018, enrollment in eFHS is still ongoing, it will continue until all FHS participants attend Examination 3. From June 2016 to July 31, 2017, we have enrolled 790 participants in eFHS. There were 1737 FHS participants examined at the Research Center. Of these, 1044 FHS participants (60%) owned a compatible iPhone and were considered eligible for enrollment in the eFHS (Figure 2). Of those who were ineligible, 359 (21%) were Android operating system (OS) smartphone owners and 25 (1%) used an iPhone running an iOS version too old to be compatible with the eFHS app. In addition, 105 eligible FHS participants (10%) declined participation in eFHS and 149 who initially agreed did not complete the setup of the app. Of the enrolled participants, 689 (64%) downloaded the app and paired the devices in the Research Center with study staff on the day of their scheduled study examination. A total of 21 eFHS participants out of the 790 (2%) returned after their examination for a dedicated eFHS training session and device set up in the Research Center, and the remainder were provided with study devices and instructions by study staff and they set up remotely.

Figure 2. Flow diagram depicting the enrollment of participants in electronic Framingham Heart Study and reasons for nonenrollment.



XSL•FO
RenderX

Study Smartphone Application

We developed a custom smartphone app (Figure 1,Multimedia Appendix 1) by using Apple's ResearchKit with the assistance of an industry partner (CareEvolution). The app was designed to electronically distribute and collect survey data and communicate with participants outside of the FHS Research Center. The app can be downloaded from the App Store. The study staff assist participants to download the eFHS app, either via written instruction or via direct assistance in the Research Center. The first-time the participants downloads the app, they are prompted to log in and register as an eFHS participant. They are then asked to provide an electronic informed consent and provide notification permissions. Health surveys are administered through the app at baseline and then at 3-month intervals (Multimedia Appendix 2). The eFHS app also prompts health access permissions that enable pairing with mobile and digital health devices, including the smartwatch and BP cuff distributed to participants. All participants receive a welcome to the eFHS message, reminders when surveys are due, notification of new activities, and reminder messages to comply with study protocols, including weekly home BP measurements and daily smartwatch use, as appropriate.

Once participants download the eFHS app, they are prompted to complete surveys that include digital versions of several standardized assessments performed in the Research Center. There are 9 baseline surveys (Multimedia Appendix 2) that request demographic information, as well as information about level of physical activity (FHS physical activity index), health-related behaviors (ie, alcohol consumption, smoking), medication use, mood (Center for Epidemiologic Studies Depression Scale, CES-D) [17], CVD risk factors, and baseline CVD and other health conditions. Additional abbreviated surveys are being distributed at 3, 6, 9, and 12 months after enrollment. At 3 months, 1 survey is deployed to assess the level of self-reported physical activity. Additional information, including a health history update, mood and function assessments (CES-D), health survey, and a repeat assessment of physical activity and mobility are planned for the 6-month digital surveys (Multimedia Appendix 2).

Study Digital Devices

In addition to the eFHS app, eFHS participants are offered 2 digital devices as part of the eFHS: (1) the Nokia-Withings digital BP cuff model BP-801 (beginning June 2016, Figure 1) and (2) the Apple Watch Generation 0 (beginning November 2016). The Nokia-Withings BP device was selected because it is Food and Drug Administration–approved for home BP monitoring, has been validated (accuracy ±3 mmHg) [18], and has been used in other digital device studies [9]. The Apple Watch was selected because of its compatibility with the Apple iOS, because it is one of the most commonly used wrist activity monitors capable of both activity and HR monitoring and because it has been shown to be the most accurate commercial wrist-worn HR and a highly accurate step monitor [19].

All eFHS participants are instructed on the proper use of the digital BP cuff and instructed to perform weekly home BP measurements. When feasible, the research staff ask participants to perform their first digital BP in the Center to ensure understanding of the BP protocol and functions of the digital BP cuff. Participants are provided with written instructions on the use of the digital cuff as well as proper home BP monitoring techniques [20]. Participants are asked to take their digital BP once per week on the same day at about the same time, if possible. Participants are provided with written BP guidelines for use to contact their health care provider if readings are elevated. In 2016, during the study period, Withings was purchased by Nokia. A new smartphone app (Nokia Health Mate) was required to sync the digital BP cuff with the eFHS app. The research staff contacted participants and helped them download the new app to ensure continued adherence with study protocols. Additional BP recordings obtained by participants are also collected and transmitted to the Research Center.

All eFHS participants are instructed on the proper use of the Apple Watch. When possible, the research staff synchronize the participant's smartphone with the eFHS app and Apple Watch in the research center. When this is not possible, the participant synchronizes the device at home. Beginning in 2017, the Apple Watch Generation 0 required an update be installed on it before it could be paired with any iPhone. Initially, participants were asked to take the smartwatch home after this change occurred because of time constraints and were provided with written instructions. The study staff then obtained dedicated study iPhones to download this Apple Watch upgrade onto all study smartwatches before they were deployed. As part of the eFHS research protocol, participants are asked to wear the smartwatch daily and are sent home with instructions on proper smartwatch use and charging.

Study Reminders

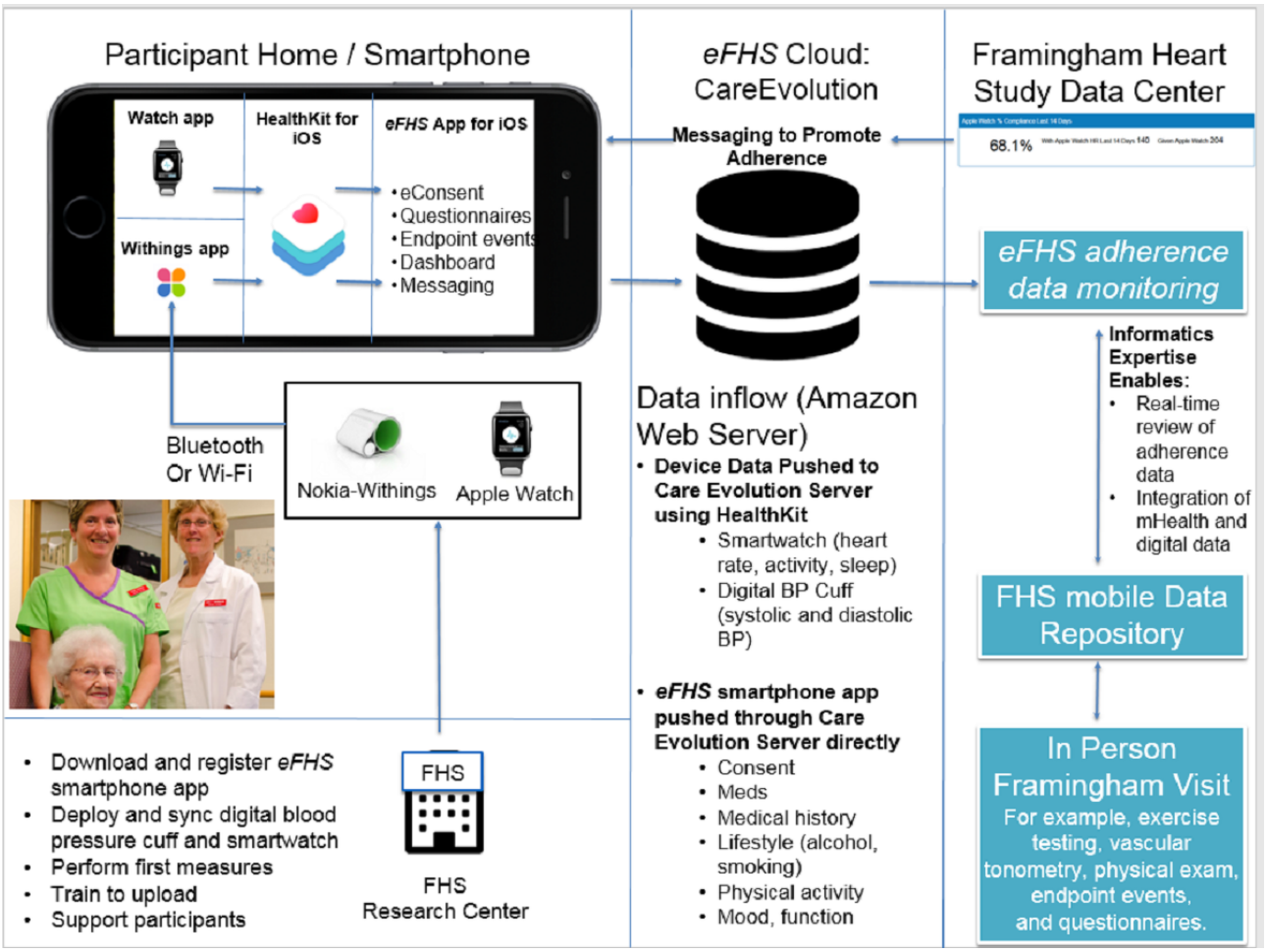
Participants receive scripted reminders (Multimedia Appendix 3) through the app at critical points. For example, when joining the eFHS, a notification message is sent as follows: "Welcome to the Electronic Framingham Heart Study. During the next two weeks, please complete your initial surveys available on the surveys tab below." Participants also receive messages when new surveys become available, "There are new surveys available to complete. Please open the eFHS app and complete them." Reminder notifications are sent when surveys come due, for example, "Reminder: You have surveys to complete" or "You have surveys due today. Please open the eFHS app and complete them." For participants who had some incomplete surveys and then completed all surveys the following message is sent, "Thank you for completing all your surveys. Your contribution is a vital part of our ongoing research efforts! Emily and the FHS team."

Data Management

All data acquired as part of the eFHS are stored on the participant's mobile device until the participant has a stable data connection. Then, study data are passively pushed to a secure cloud server in the United States in a Health Insurance Portability and Accountability Act compliant fashion (Figure 3). Adherence to study protocols is monitored by the study team. Data labeled with the participant's eFHS ID are batched and transmitted to the FHS Research Center servers. FHS statisticians perform quality control activities and further prepare data for analyses.

XSL•FO
RenderX

Figure 3. Electronic Framingham Heart Study (eFHS) components (eFHS smartphone app, digital blood pressure cuff, and smartwatch) and data infrastructure. BP: blood pressure; eFHS: electronic Framingham Heart Study.



Study Analyses

Primary eFHS outcome variables include the following: survey, BP, wearable device, and overall adherence over 12 months. In prespecified analyses, we will identify predictors of successful eFHS protocol adherence to inform future mHealth research. We also plan to explore relations among validated measures of cardiovascular health obtained in the FHS Research Center during examination 3, including measures of vascular stiffness and exercise capacity from formal cardiopulmonary exercise testing, with home BP, and average daily HR and steps, respectively.

For the present descriptive analysis, we described the characteristics of all eFHS participants enrolled as of July 31, 2017 and compared them with exam 3 participants who did not participate (either ineligible or declined). Baseline characteristics of all eFHS and other exam 3 participants are presented as means ± standard deviation for continuous variables and as numbers and percentages for nominal variables.

We examined completion rates of the baseline and 3-month surveys and use of digital devices at baseline and 3 months. We restricted our analysis of survey completion to eligible participants (baseline, n=790) and 3-month surveys (n=790). We report the proportion with successfully uploaded survey data (some survey completed, entire survey completed, and no survey completion) for the digital baseline survey (comprising 9 questionnaires, Multimedia Appendix 2) and survey completion at 3 months (comprising 1 questionnaire, Multimedia Appendix 2). Similarly, we examined adherence to weekly home BP monitoring only among participants who consented to take the Nokia Withings BP cuff (n=573) as part of the eFHS. Similarly, we assessed smartwatch use only among participants who took an Apple Watch and consented to wear it daily as part of the eFHS (n=405). We examined the weekly proportion of participants with at least one BP and HR recording, respectively.

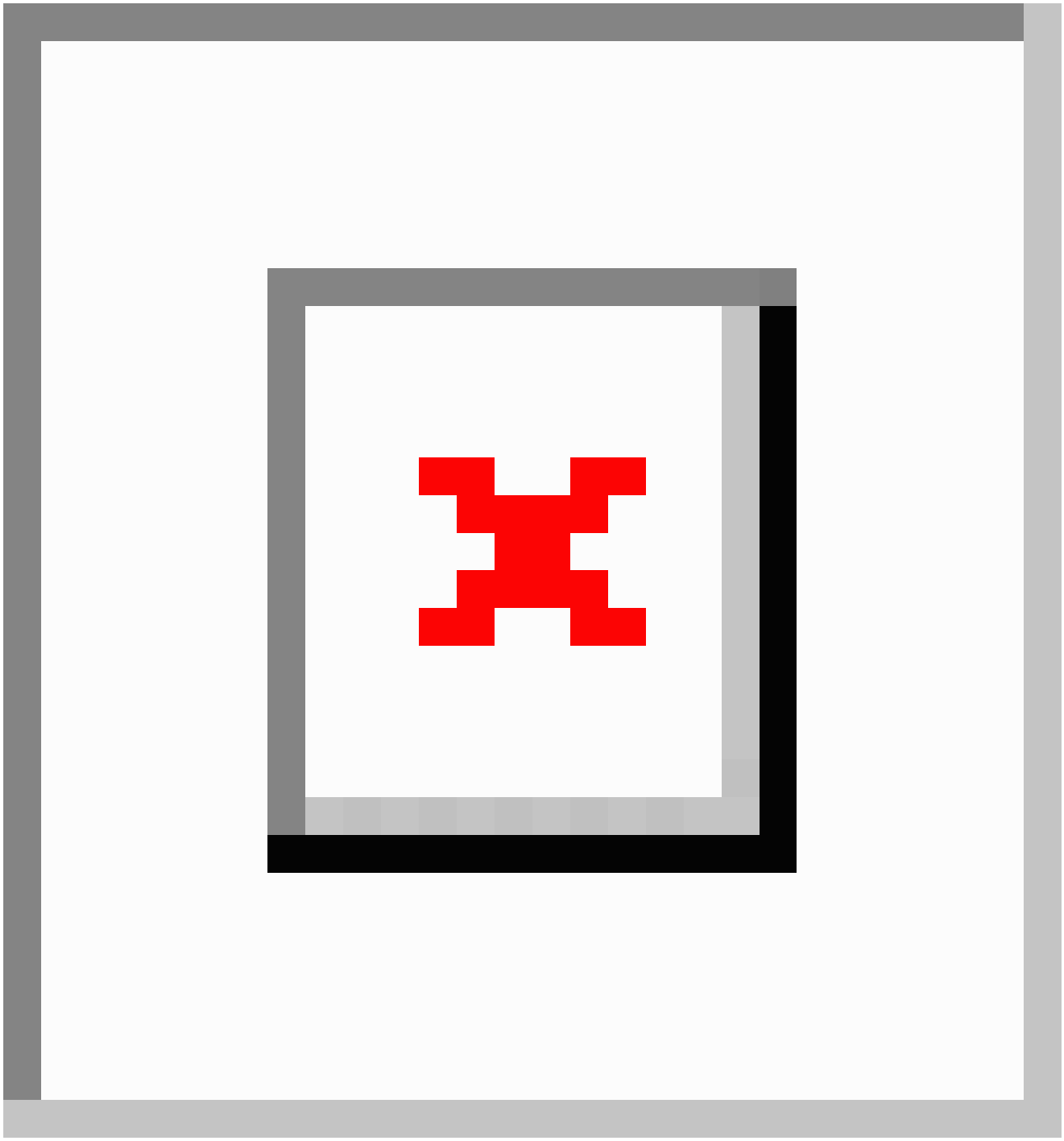
Results

Characteristics of Electronic Framingham Heart Study Participants

Between June 2016 and July 2017, we enrolled 790 participants in eFHS. Due primarily to a lack of compatible smartphone, 947 additional FHS participants were not enrolled (Figure 2). eFHS participants were similar to FHS participants who were ineligible (ie, incompatible smartphone) or declined to participate in the eFHS with respect to their race. However, eFHS participants were slightly younger, more likely to be female, attained a higher level of education, more likely to report being in excellent health, and were less likely to be affected by CVD or its risk factors than those who were ineligible or did not participate (Multimedia Appendix 4).

XSL•FO
RenderX

Figure 4. Proportion of eFHS participants uploading at least A) one digital BP weekly or B) one heart rate weekly over the 12 weeks following enrollment.



App and Device Deployment and Adherence

As shown in Multimedia Appendix 5, out of the 790 (85%) eFHS participants completed all the baseline surveys and an additional 38 out of the 790 participants (5%) completed some of the baseline surveys. A total of 470 out of the 790 (59%) eligible participants completed the 3-month survey. The average time to complete all of the baseline and 3-month digital surveys was modest (18 min for the baseline surveys and 2 min for the 3-month survey). Relating to the staggered rollout of study devices and delayed smartwatch availability, a greater number of eFHS participants received a digital BP cuff (n=573) than those who received a study Apple Watch (n=405). We did not collect data on the reasons why some participants opted to take 1, both, or no smart devices, because of logistical constraints. However, based on the communications with the FHS Research Center staff, we believe that choices were partly influenced by the perceived time burden imposed by home BP or wearable device use.

Some digital BP or HR data were obtained from 79% (n=626) of the 790 eFHS participants who elected to take at least one study device. Just under half (44%) of the participants (n=156/353, received the BP cuff, smartwatch and app) transmitted both BP and HR recordings during the 12th study week. Adherence to the use of the study smartwatch (Figure 4, as defined for study purposes as uploading at least one HR value during the week, was appreciably (about 30% higher) and

XSL•FO
RenderX

consistently higher over the 3 months following enrollment than the adherence to the use of the digital BP cuff (Figure 4), as defined as uploading at least one BP value during the week. We observed slightly higher rates of adherence to BP monitoring but no significant difference with respect to rates of adherence to HR monitoring among eFHS participants with a previous history of hypertension or cardiovascular disease as compared with those free from hypertension or CVD in our sample (Multimedia Appendix 6).

Discussion

Study Background and Design

The eFHS is a novel e-cohort embedded in the ongoing prospective FHS of middle-aged Gen 3 and Omni-group 2 cohort participants attending an in-person study examination. We have described the methodology for recruitment of the eFHS cohort, the characteristics of this cohort after 1 year of enrollment, as well as survey completion and adherence to a digital BP cuff and a smartwatch (for once weekly home BP and HR measurement, respectively) over 3 months of follow-up.

Our decision to target the FHS Gen 3 and Omni 2 cohorts for participation in the creation of an e-cohort was substantially informed by 2 previous studies. First, a digital connectedness survey was conducted in 2014 to quantify internet access and device use among FHS participants, including computers, internet, smartphones, wireless devices, and apps [14]. We received information from 6503 FHS participants: a total of 87% (n=5678) of our participants had internet access and 60% (n=3907) used smartphones. We observed that 69% (n=2689) of smartphone users owned iOS (Apple) compatible devices, 22% (n=855) had Androids, and most smartphone users were FHS Gen 3 participants [14]. Second, in 2016, we conducted a randomized controlled trial (NCT02531763) among FHS families using a smartphone app (Moves or Fitbit) and a wrist-based activity monitor (Fitbit) in the *B* ehavioral *E* conomics *F* ramingham *I* ncentive *T* rial, *BE FIT*, study [21]. FHS participants randomly assigned to the intervention arm entered a 12-week game to achieve physical activity step goals. Among the 200 trial participants, 97% completed the study. The BE FIT study enabled us to (1) gain experience deploying a smartphone app and wearable devices, (2) engage FHS participants in mHealth activities (89% of the intervention arm and 83% of the control arm agreed or strongly agreed to the statement “I enjoyed participating”), and (3) determine that participants were highly likely to adhere to device use over at least 3 months.

The enrollment and retention rates in epidemiology studies have declined over the last 2 decades [22,23]. Furthermore, the financial and logistical burdens associated with traditional cohort-based epidemiology are substantial [24]. Finally, cohort participants spend just 4 hours in the research center every 6 to 8 years or a few hours per year with their designed health care provider, but they spend 5000 waking hours each year making decisions that can profoundly affect health (eg, to stop smoking, to engage in regular physical activity, and to take prescribed medications to treat BP) [25]. Real-world data obtained from participants over extended periods can potentially enhance cardiovascular risk factor identification and prediction of incident CVD [10]. To that end, development of electronic epidemiological cohorts, or e-cohorts, whereby information is collected electronically by leveraging Web-based platforms, smart devices, or smartphone-compatible applications, offer potentially sustainable alternatives in this era of digitization [9,26-29].

In contrast to several other e-cohorts that primarily use Web-based platforms for enrollment and interaction [9,26-29], we elected to develop a novel app for survey administration and deploy it for use in the context of a scheduled study exam conducted at the FHS research center. Electronic surveys administered using Web-based platforms are fast to administer and easier to code than smartphone apps [30,31]. In a recent analysis conducted by the Pew Research Center, general response rates for study participants using smartphone app-based surveys were lower than the Web-based surveys (any survey completed by 76% vs 94%, in app vs Web-based group, respectively; all surveys completed by 12% vs 17%, in app vs Web-based group, respectively) [32]. However, smartphone app-based surveys offer several advantages, including ease of administration, integration of data from other smart devices, and collection of *real-time* responses that may occur outside the home or office and thereby suffer from less recall bias than Web-based surveys [32].

Study Enrollment

Just under half (n=790, 48%) of the 1737 FHS participants enrolled in the eFHS, a rate lower than has been observed in other e-cohorts (ie, Health eHeart, 78% of registered participants completed the consent, 86% of consented completed at least one survey, 37% of consented participants completed all surveys) [9]. Nevertheless, only 105 participants (10%) of the 1044 FHS participants with a compatible smartphone declined eFHS participation and the vast majority of FHS Gen 3 and Omni 2 participants who were eligible agreed to enroll, with ownership of an Android OS smartphone being the major reason for not being enrolled (Figure 2). To address this problem, we have recently developed an eFHS app for the Android OS and are now actively enrolling these participants into the eFHS. Furthermore, once enrolled and consented, survey completion rates are higher than some *de novo* e-cohorts [33]. We hypothesize that embedding enrollment in the Research Center of an ongoing cohort was instrumental in promoting initial enrollment and facilitating high rates of survey completion through in-person training.

Study Adherence

Among the 790 eFHS participants, we observed that the high rates of completion for our eFHS app-administered baseline survey (90%; 713 participants completed at least some of the baseline survey) and 3-month survey were 59% (n=470), with little reinforcement. The high completion rate is notable in light of the feature that participants needed to keep the eFHS app open to receive any eFHS notifications, including a reminder to complete the surveys. The high rates of survey completion likely resulted from app download in the research center for most participants as well as through optimization of our app survey methodology to avoid respondent fatigue. We optimized

XSL•FO
RenderX

the digital surveys through the use of progress indicators, by limiting the number of answer options and allowing participants to complete the questionnaire across multiple sittings and by restricting the number of questions per survey to 10 to 20 [34]. Compared with other contemporary e-cohorts with ongoing data collection, our rates of survey completion were higher than that those in previous studies, including eCARDIA (52% survey completion) [29] and the Health eHeart Study (37% completed all surveys and 29% of myHeart participants completed all surveys) [9]. Our findings suggest that linking e-cohort enrollment to a traditional research center study visit offers distinct advantages with regard to higher rates of adherence to digital health survey completion. We are developing interactive messaging functionalities as part of an eFHS dashboard in an effort to boost survey completion and promote long-term adherence to survey completion.

Digital Device Deployment and Adherence

The majority, 573 (73%), of the 790 eFHS participants agreed to use a digital BP cuff weekly and over half, 405 out of the 790 participants (51%), agreed to wear a smartwatch daily. The difference in device acceptance is related to the delayed start for the smartwatch. Once the smartwatch deployment began, acceptance of the 2 devices was similar. Adherence to device use was initially high, likely secondary to the significant technical support provided to eFHS participants, particularly those who had their digital devices set up and synchronized with the eFHS app in the FHS research center. However, digital BP use waned significantly over the 3 months following enrollment, with fewer than half of participants uploading a BP value in the 12th study week. Digital BP adherence rates in our study were slightly lower than those seen in another recent study, in which the rate of adherence to twice-monthly home BP monitoring (iVitality) was 64% over a 6-month period [35]. We required participants to upload a BP more frequently and our eFHS sample of 790 participants had a relatively low prevalence of diagnosed hypertension (24%, n=186), which may have contributed to the lower rate of study protocol adherence as compared with this smaller e-cohort study involving participants enrolled via the Web. This hypothesis appears to be supported by the slightly higher rate of adherence among participants with a history of hypertension in our cohort (Multimedia Appendix 5). However, our findings should be interpreted with caution in light of the limited number of participants with a history of CVD in our sample.

The rates of smartwatch use were significantly higher than the rates of digital BP use throughout the first 12 weeks of the study despite the release of an iOS update during this period that interfered with uploading HR values from some Apple Watches. Differences in rates of adherence to the digital BP cuff and smartwatch are likely attributable to the passive nature of data acquisition from the smartwatch and the value to participants from having a functional watch [36].

Study Challenges

Our team encountered several technical challenges as the eFHS progressed. Participants *forced closed* the eFHS app unknowingly, preventing transmission of BP and HR data. We updated our in-person training and written materials to advise participants to keep the eFHS app open. The BP cuff manufacturer Withings changed to Nokia and we encountered a brief period of difficulty downloading and configuring the BP app with the device. A new version of the Nokia Health Mate app was released, and that resolved the problem we encountered. Finally, there was an iOS update to the Apple Watch that significantly added to the time to pair the watch and iPhone in the Research Center with staff support, leading many participants to opt to take the device and instructions home to set up independently. The Research Center staff have obtained dedicated study iPhones to update the watches before the in-person examinations to speed the synchronization.

Differences between Electronic Framingham Heart Study and Framingham Heart Study Participants

In a manner consistent with other digital or e-cohorts, as compared with other Gen 3 and Omni 2 FHS participants who were ineligible or declined participation, eFHS participants were younger, more likely to be women, and had more favorable socioeconomic determinants of health and better self-reported overall health status and lower prevalence of cardiovascular risk factors [9,11,26-28]. Our findings mirror Pew data on iOS device ownership and have implications for recruitment of participants for the National Institutes of Health's *AllofUS* initiative, as methods to address sampling biases might be needed to obtain representative data from e-cohorts that are applicable to the overall US population [10]. We are attempting to create a more generalizable e-cohort by developing enhancements to our app, and we will reexamine differences among eFHS versus other FHS participants who are ineligible or decline participation after we complete the enrollment of participants with Android OS–based devices.

Future Directions

The FHS represents a useful setting to define adherence to novel mobile methods for cardiovascular phenotyping using mobile and digital devices among older adults at risk for CVD. Once device deployment and follow-up are complete, we plan to analyze long-term rates of survey, BP, and wearable device adherence over 12 months. In subsamples of eFHS participants, we intend to study the effectiveness of various mHealth supportive interventions, such as personalized short message service text messaging, on protocol adherence. Finally, we plan to leverage the richness of the FHS through in-person examination completed by all study participants on the day of their enrollment. We will analyze relations among validated measures of cardiovascular health, particularly markers of vascular stiffness and exercise capacity, home BP, as well as HR and activity. The overarching goal of the eFHS will be to inform clinical practice and research of how mHealth and digital health measures relate to cardiovascular health and CVD risk.

Study Strengths and Limitations

Our study findings should be considered in the context of their strengths and limitations. We are developing an eFHS cohort from the Gen 3 and Omni 2 (multiethnic) participants who are presenting to the research center for an in-person scheduled examination. In so doing, we have enhanced our ability to enroll and train participants on app and device use, as well as to

XSL•FO
RenderX

correlate findings from digital devices with the data collected during the in-person examinations. Furthermore, we were able to attain high rates of digital survey completion using a novel smartphone app and acceptable rates of home device use over 3 months, all with limited post enrollment technical support and minimal use of push notifications or reminders. However, our study has certain limitations. First, the eFHS cohort largely comprises middle-aged, white participants—mean age 53 (8) years—and, enrollment rates, survey completion rates, and digital device home use adherence rates observed in our study might not be generalizable to the samples with older cohort participants or cohorts involving a larger proportion of individuals of other racial/ethnic backgrounds. Second, the app was designed for use with the iPhone, the CareEvolution team has extended the app for Android users who are currently being enrolled in the study. Third, we restricted our outcomes to survey completion and adherence to mobile devices at 3-month intervals for our current analyses. We plan to examine long-term adherence rates at 6 and 12 months as the study proceeds.

Conclusions

In this paper, we present our ongoing work to develop a new e-cohort derived from the FHS participants, as well as characteristics of the study population. We report acceptable rates of initial enrollment, high rates of home digital survey completion using a novel eFHS smartphone app, high rates of home HR monitoring using a smartwatch, and modest rates of adherence to weekly BP measurement using a digital BP cuff. Our work and initial findings provide data and show that it is feasible to embed an e-cohort within an ongoing traditional epidemiology study and that middle-aged participants are capable of completing home-based digital surveys and cardiovascular phenotyping using digital devices in the short term. Our initial experience is promising, and we look forward to linking new digital cardiovascular phenotypic data, including daily activity, sedentary time, average HR, and home BP collected in eFHS to validated cardiovascular risk factors and incident cardiovascular outcomes collected in the FHS.

Acknowledgments

This study was supported by an award from the Robert Wood Johnson Foundation (number 74624) and a grant from the National Heart Lung and Blood Institute (R01HL141434); investigator time from the following grants: R01HL126911, 2R01 HL092577, 1P50HL120163 (EJB), R01HL126911 (DDM), R01HL137734 (DDM), R01HL137794 (DDM), R01HL13660 (DDM).The Apple Watches were provided to Boston University by the Apple Corporation at no cost to the study.

Conflicts of Interest

DDM discloses equity stakes or consulting relationships with Flexcon, Bristol-Myers Squibb, Boston Biomedical Associates, Pfizer, Samsung, and research support from Sanofi Aventis, Otsuka Pharmaceuticals, Philips Healthcare, Biotronik, Bristol Myers Squibb, Boeringher Ingelheim, and Pfizer. DDM has inventor stake in Mobile Sense Technologies, LLC. VK is principal, and CN is an employee of Care Evolution, Inc, a health care technology company CareEvolution and a leading provider of secure interoperability solutions. Apple was not involved in the study design, analysis, interpretation, or reporting of the study results. Other authors have no relevant disclosures.

Multimedia Appendix 1

Screenshots of the Welcome, Registration, Survey, and Task Completion screens.

[PNG File, 170KB - jmir_v21i3e12143_app1.PNG]

Multimedia Appendix 2

Electronic Framingham Heart Study baseline and follow-up survey components.

[PNG File, 98KB - jmir_v21i3e12143_app2.png]

Multimedia Appendix 3

Scripted notifications, condition triggers, and notification content deployed through the electronic Framingham Heart Study smartphone application. eFHS: electronic Framingham Heart Study; FHS: Framingham Heart Study.

[PNG File, 69KB - jmir_v21i3e12143_app3.PNG]

Multimedia Appendix 4

Characteristics of participants in electronic Framingham Heart Study vs those ineligible or not consenting.

[PNG File, 167KB - jmir_v21i3e12143_app4.PNG]

XSL•FO
RenderX

Multimedia Appendix 5

Number and proportion of eligible electronic Framingham Heart Study participants completing all, some, or none of the baseline and 3-month surveys.

[PNG File, 16KB - jmir_v21i3e12143_app5.PNG]

Multimedia Appendix 6

Proportion of eFHS participants uploading at least A) one digital Blood Pressure (BP) weekly or B) one heart rate (HR) weekly over the 12 weeks following enrollment, stratified by prevalent hypertension (HTN) and cardiovascular disease (CVD).

[PNG File, 29KB - jmir_v21i3e12143_app6.PNG]

References

1. Lauer MS. Time for a creative transformation of epidemiology in the United States. J Am Med Assoc 2012 Nov 7;308(17):1804-1805. [doi: 10.1001/jama.2012.14838] [Medline: 23117782]
2. Khoury MJ, Lam TK, Ioannidis JP, Hartge P, Spitz MR, Buring JE, et al. Transforming epidemiology for 21st century medicine and public health. Cancer Epidemiol Biomarkers Prev 2013 Apr;22(4):508-516 [FREE Full text] [doi: 10.1158/1055-9965.EPI-13-0146] [Medline: 23462917]
3. Sorlie PD, Bild DE, Lauer MS. Cardiovascular epidemiology in a changing world--challenges to investigators and the National Heart, Lung, and Blood Institute. Am J Epidemiol 2012 Apr 1;175(7):597-601 [FREE Full text] [doi: 10.1093/aje/kws138] [Medline: 22415032]
4. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. Am J Physiol Regul Integr Comp Physiol 2017 Dec 1;312(3):R358-R367 [FREE Full text] [doi: 10.1152/ajpregu.00349.2016] [Medline: 28052867]
5. Roger VL, Boerwinkle E, Crapo JD, Douglas PS, Epstein JA, Granger CB, et al. Strategic transformation of population studies: recommendations of the working group on epidemiology and population sciences from the National Heart, Lung, and Blood Advisory Council and Board of External Experts. Am J Epidemiol 2015 Mar 15;181(6):363-368 [FREE Full text] [doi: 10.1093/aje/kwv011] [Medline: 25743324]
6. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. J Am Med Assoc 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
7. Powell AC, Landman AB, Bates DW. In search of a few good apps. J Am Med Assoc 2014 May 14;311(18):1851-1852. [doi: 10.1001/jama.2014.2564] [Medline: 24664278]
8. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? J Am Med Assoc 2013 Dec 11;310(22):2395-2396. [doi: 10.1001/jama.2013.281078] [Medline: 24158428]
9. Guo X, Vittinghoff E, Olgin JE, Marcus GM, Pletcher MJ. Volunteer participation in the health eHeart study: a comparison with the US population. Sci Rep 2017 Dec 16;7(1):1956 [FREE Full text] [doi: 10.1038/s41598-017-02232-y] [Medline: 28512303]
10. National Institutes of Health: All of Us Research Program. URL: https://allofus.nih.gov/ [accessed 2018-05-13] [WebCite Cache ID 6zOGpqUe7]
11. McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts cardiovascular health study. JAMA Cardiol 2017 Jan 1;2(1):67-76. [doi: 10.1001/jamacardio.2016.4395] [Medline: 27973671]
12. Burke LE, Ma J, Azar KMJ, Bennett GG, Peterson ED, Zheng Y, American Heart Association Publications Committee of the Council on Epidemiology and Prevention, Behavior Change Committee of the Council on Cardiometabolic Health, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, Council on Quality of Care and Outcomes Research, and Stroke Council. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from the American Heart Association. Circulation 2015 Sep 22;132(12):1157-1213. [doi: 10.1161/CIR.0000000000000232] [Medline: 26271892]
13. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am J Epidemiol 2007 Jun 1;165(11):1328-1335. [doi: 10.1093/aje/kwm021] [Medline: 17372189]
14. Fox CS, Hwang S, Nieto K, Valentino M, Mutalik K, Massaro JM, et al. Digital connectedness in the Framingham Heart Study. J Am Heart Assoc 2016 Apr 13;5(4):e003193 [FREE Full text] [doi: 10.1161/JAHA.116.003193] [Medline: 27076568]
15. Framingham Heart Study. Group 3 Exam 3 Research Consent Form URL: https://www.framinghamheartstudy.org/wp-content/uploads/2017/08/Group-3-Exam-3-Informed-Consent-IRB.pdf [accessed 2018-05-13] [WebCite Cache ID 6zOFruLTX]
16. Framingham Heart Study. Consent to Participate in the eFHS Research Study URL: https://www.framinghamheartstudy.org/wp-content/uploads/2017/08/H-36586-eFHS-Approved-Consent.pdf [accessed 2018-05-13] [WebCite Cache ID 6zOFgmREb]
17. Mohebbi M, Nguyen V, McNeil JJ, Woods RL, Nelson MR, Shah RC, ASPREE Investigator Group. Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms

XSL•FO
RenderX

in healthy community dwelling older adults. Gen Hosp Psychiatry 2018;51:118-125 [FREE Full text] [doi: 10.1016/j.genhosppsych.2017.08.002] [Medline: 28890280]
18. Topouchian J, Agnoletti D, Blacher J, Youssef A, Chahine MN, Ibanez I, et al. Validation of four devices: Omron M6 Comfort, Omron HEM-7420, Withings BP-800, and Polygreen KP-7670 for home blood pressure measurement according to the European Society of Hypertension International Protocol. Vasc Health Risk Manag 2014;10:33-44 [FREE Full text] [doi: 10.2147/VHRM.S53968] [Medline: 24476688]
19. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017 May 24;7(2):1-12 [FREE Full text] [doi: 10.3390/jpm7020003] [Medline: 28538708]
20. Parati G, Stergiou G, O'Brien E, Asmar R, Beilin L, Bilo G, European Society of Hypertension Working Group on Blood Pressure MonitoringCardiovascular Variability. European Society of Hypertension practice guidelines for ambulatory blood pressure monitoring. J Hypertens 2014 Jul;32(7):1359-1366. [doi: 10.1097/HJH.0000000000000221] [Medline: 24886823]
21. Patel MS, Benjamin EJ, Volpp KG, Fox CS, Small DS, Massaro JM, et al. Effect of a game-based intervention designed to enhance social incentives to increase physical activity among families: the BE FIT randomized clinical trial. JAMA Intern Med 2017 Nov 1;177(11):1586-1593 [FREE Full text] [doi: 10.1001/jamainternmed.2017.3458] [Medline: 28973115]
22. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
23. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006 Feb 1;163(3):197-203. [doi: 10.1093/aje/kwj036] [Medline: 16339049]
24. Manolio TA, Weis BK, Cowie CC, Hoover RN, Hudson K, Kramer BS, et al. New models for large prospective studies: is there a better way? Am J Epidemiol 2012 May 1;175(9):859-866 [FREE Full text] [doi: 10.1093/aje/kwr453] [Medline: 22411865]
25. Asch DA, Muller RW, Volpp KG. Automated hovering in health care--watching over the 5000 hours. N Engl J Med 2012 Jul 5;367(1):1-3. [doi: 10.1056/NEJMp1203869] [Medline: 22716935]
26. Coyne KS, Sexton CC, Kopp ZS, Luks S, Gross A, Irwin D, EpiLUTS Team. Rationale for the study methods and design of the epidemiology of lower urinary tract symptoms (EpiLUTS) study. BJU Int 2009 Aug;104(3):348-351 [FREE Full text] [doi: 10.1111/j.1464-410X.2009.08425.x] [Medline: 19291251]
27. Huybrechts KF, Mikkelsen EM, Christensen T, Riis AH, Hatch EE, Wise LA, et al. A successful implementation of e-epidemiology: the Danish pregnancy planning study 'Snart-Gravid'. Eur J Epidemiol 2010 May;25(5):297-304 [FREE Full text] [doi: 10.1007/s10654-010-9431-y] [Medline: 20148289]
28. Mishra GD, Hockey R, Powers J, Loxton D, Tooth L, Rowlands I, et al. Recruitment via the Internet and social networking sites: the 1989-1995 cohort of the Australian longitudinal study on women's health. J Med Internet Res 2014;16(12):e279 [FREE Full text] [doi: 10.2196/jmir.3788] [Medline: 25514159]
29. Kershaw KN, Liu K, Goff DC, Lloyd-Jones DM, Rasmussen-Torvik LJ, Reis JP, et al. Description and initial evaluation of incorporating electronic follow-up of study participants in a longstanding multisite cohort study. BMC Med Res Methodol 2016 Sep 23;16(1):125 [FREE Full text] [doi: 10.1186/s12874-016-0226-z] [Medline: 27664124]
30. Wyatt JC. When to use web-based surveys. J Am Med Inform Assoc 2000;7(4):426-429 [FREE Full text] [doi: 10.1136/jamia.2000.0070426] [Medline: 10887170]
31. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
32. Pew Research Center. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 6zOJhu5O3]
33. Spartano NL, Lunetta K, Trinquart L, Valentino M, Manders E, Pletcher M, et al. Abstract MP24: comparison of on-site versus remote support for a mobile-device pilot study: a collaboration between the Framingham Heart Study and Health eHeart Study (FHS-HeH). Circulation 2018;137(Suppl 1):AMP24 [FREE Full text]
34. O'Reilly-Shah VN. Factors influencing healthcare provider respondent fatigue answering a globally administered in-app survey. PeerJ 2017;5:e3785 [FREE Full text] [doi: 10.7717/peerj.3785] [Medline: 28924502]
35. Wijsman LW, Richard E, Cachucho R, de Craen AJ, Jongstra S, Mooijaart SP. Evaluation of the use of home blood pressure measurement using mobile phone-assisted technology: the iVitality proof-of-principle study. JMIR Mhealth Uhealth 2016 Jun 13;4(2):e67 [FREE Full text] [doi: 10.2196/mhealth.5485] [Medline: 27296628]
36. Silva de Lima AL, Hahn T, Evers LJW, de Vries NM, Cohen E, Afek M, et al. Feasibility of large-scale deployment of multiple wearable sensors in Parkinson's disease. PLoS One 2017;12(12):e0189161 [FREE Full text] [doi: 10.1371/journal.pone.0189161] [Medline: 29261709]

Abbreviations

BP: blood pressure
CES-D: Center for Epidemiologic Studies Depression Scale
e-cohort: electronic cohort

XSL•FO
RenderX

eFHS: electronic Framingham Heart Study
FHS: Framingham Heart Study
HR: heart rate
mHealth: mobile health
NOS: New Offspring Spouse Cohort
OS: operating system

Edited by G Eysenbach; submitted 13.09.18; peer-reviewed by R Shan; comments to author 25.10.18; revised version received 10.01.19; accepted 21.01.19; published 01.03.19.

Please cite as:
McManus DD, Trinquart L, Benjamin EJ, Manders ES, Fusco K, Jung LS, Spartano NL, Kheterpal V, Nowak C, Sardana M, Murabito JM
Design and Preliminary Findings From a New Electronic Cohort Embedded in the Framingham Heart Study
J Med Internet Res 2019;21(3):e12143
URL: http://www.jmir.org/2019/3/e12143/
doi:10.2196/12143
PMID:30821691

©David D McManus, Ludovic Trinquart, Emelia J Benjamin, Emily S Manders, Kelsey Fusco, Lindsey S Jung, Nicole L Spartano, Vik Kheterpal, Christopher Nowak, Mayank Sardana, Joanne M Murabito. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Automated Analysis of Domestic Violence Police Reports to Explore Abuse Types and Victim Injuries: Text Mining Study

George Karystianis[1], BSc, MSc, PhD; Armita Adily[1], BMedSc, MPH, PhD; Peter W Schofield[2], BSc (Hons), BS, MSc, MD; David Greenberg[3], MB, ChB, FFPsych (SA), FRCPC, FRANZCP, MMed; Louisa Jorm[4], BVSc, MSc, PhD; Goran Nenadic[5], BSc, MSc, PhD; Tony Butler[1], BSc, MSc, PhD

[1]The Kirby Institute, Faculty of Medicine, The University of New South Wales, Sydney, Australia

[2]Neuropsychiatry Service, Hunter New England Health, Newcastle, Australia

[3]School of Psychiatry, The University of New South Wales, Sydney, Australia

[4]Centre for Big Data Research in Health, The University of New South Wales, Sydney, Australia

[5]School of Computer Science, The University of Manchester, Manchester, United Kingdom

Corresponding Author:
George Karystianis, BSc, MSc, PhD
The Kirby Institute
Faculty of Medicine
The University of New South Wales
Level 6, Wallace Wurth Building
High Street, Kensington NSW
Sydney, 2052
Australia
Phone: 61 (2) 9385 0900
Email: gkarystianis@kirby.unsw.edu.au

Abstract

Background: The police attend numerous domestic violence events each year, recording details of these events as both structured (coded) data and unstructured free-text narratives. Abuse types (including physical, psychological, emotional, and financial) conducted by persons of interest (POIs) along with any injuries sustained by victims are typically recorded in long descriptive narratives.

Objective: We aimed to determine if an automated text mining method could identify abuse types and any injuries sustained by domestic violence victims in narratives contained in a large police dataset from the New South Wales Police Force.

Methods: We used a training set of 200 recorded domestic violence events to design a knowledge-driven approach based on syntactical patterns in the text and then applied this approach to a large set of police reports.

Results: Testing our approach on an evaluation set of 100 domestic violence events provided precision values of 90.2% and 85.0% for abuse type and victim injuries, respectively. In a set of 492,393 domestic violence reports, we found 71.32% (351,178) of events with mentions of the abuse type(s) and more than one-third (177,117 events; 35.97%) contained victim injuries. “Emotional/verbal abuse” (33.46%; 117,488) was the most common abuse type, followed by “punching” (86,322 events; 24.58%) and “property damage” (22.27%; 78,203 events). “Bruising” was the most common form of injury sustained (51,455 events; 29.03%), with “cut/abrasion” (28.93%; 51,284 events) and “red marks/signs” (23.71%; 42,038 events) ranking second and third, respectively.

Conclusions: The results suggest that text mining can automatically extract information from police-recorded domestic violence events that can support further public health research into domestic violence, such as examining the relationship of abuse types with victim injuries and of gender and abuse types with risk escalation for victims of domestic violence. Potential also exists for this extracted information to be linked to information on the mental health status.

(J Med Internet Res 2019;21(3):e13067) doi:10.2196/13067

KEYWORDS

domestic violence; injuries; abuse types; text mining; rule-based approach; police narratives

XSL•FO
RenderX

Introduction

Background

Domestic violence is a global social and public health phenomenon with important health consequences that affect thousands of lives each year [1-3]. It can be defined as "any incident of threatening behavior, violence (or psychological, physical, sexual, financial, emotional) abuse between adults who are or have been an intimate partner or family member, regardless of gender or sexuality" [4-6]. However, domestic violence can also occur in other relationship structures such as between a caregiver and a dependent person, including a child, or those living together in a household but not in an intimate relationship [4,5]. A multicountry violence study conducted by the World Health Organization estimates a prevalence of 15%-71% in physical and sexual partner violence toward women [1,3]. In Australia, in 2018, one of six women and one of 16 men experienced physical or sexual violence by a current or previous partner [7]. Domestic violence has various forms—from physical to emotional and verbal abuse. The type of abuse received and perpetrated may vary by gender, with each type bearing short- and long-term (physical and mental) health consequences for the victims [8-11]. Domestic violence bears a significant economic cost: Within Australia alone, the cost of violence against women was around Aus $22.2 billion in 2015-2016 [2,3,12].

The New South Wales Police Force (NSWPF) recorded 123,330 domestic violence–related events in 2017 in WebCOPS (Web Computerised Operational Policing System), a Web-based interface for the COPS, which enables the police to capture and analyze crime information on an organization-wide basis [13]. WebCOPS contains detailed information about domestic violence events as both structured fields (date of birth, Aboriginal status, whether weapons were used, etc) and free unstructured text called "event narratives." An event can contain more than one text narrative describing, in detail, alleged incident(s) that occurred between the person of interest (POI) and the victim, information regarding the circumstances of the event, and any action(s) taken by the police. Narratives are frequently written without a specific structure, featuring various misspellings, typographical and grammatical errors, and (sometimes informal) acronyms and abbreviations that can have different meanings depending on the context [13].

Domestic violence event narratives contain a wealth of important information regarding injuries and abuse types, which is not found in the medical records unless medical attention is sought, although even attainment of medical attention may not be flagged as related to domestic violence. However, the volume of the recorded data along with the associated long unstructured narratives makes it difficult to identify potentially meaningful information through traditional ethnographic/qualitative research methods involving eyeballing the records. One research paper recently commented that "...there is no systematic way to extract information from these [police] narratives other than by manual review" [14].

Prior Work

There is a need for methods that can automatically extract information of interest from large volumes of data in a short time. Text mining has been used for more than 30 years to harvest information from unstructured text in many fields, particularly in biomedicine [15-20]. Recent efforts have sought to text mine crime-related information from online media publications [21-23], with limited attempts to process police reports [13,24-28]. Previous work extracted data on the names, narcotic drugs, and weapons with varying degrees of success (F1-score ranging from 46% to 81%) through named entity extraction [24,25] and police report classification of events as domestic violence or nondomestic violence related, using an unsupervised clustering technique that correctly classified 44% of the reports set aside for manual inspection [26]. Other efforts included recognition of crime-related information (such as drugs, weapons, and facial features) from witness narratives through dictionaries and rules, with F1-scores ranging from 82% to 93% [27,28]. Recently, Karystianis et al applied a rule-based approach combined with manually crafted dictionaries to extract mentions of mental illnesses for POIs and victims from police text narratives of recorded domestic violence events with an average F1-score of 84% [13].

Aim

In this paper, we investigate whether the application of a text mining method can automatically extract abuse types (conducted by POIs) and sustained victim injuries from a large-scale corpus of 492,393 domestic violence events.

Methods

Data

We used a corpus of 492,393 domestic violence events provided to the researchers by the NSWPF, occurring from January 2005 to December 2016 [13]. The domestic violence events were flagged in WebCOPS as "domestic violence related," the description of violence was coded as "domestic," and the relationship between the victim and the POI included any of the following: "spouse/partner" (including ex-spouse/ex-partner), "boyfriend/girlfriend" (including ex-boyfriend/ex-girlfriend), "parent/guardian" (including step/foster), "child" (including step/foster), "sibling," "other member of family" (including kin), or "carer." These events covered the following categories: various types of assaults; breaches of Apprehended Violence Orders; homicides; malicious damage to property; and offense against another person such as intimidation, kidnapping, abduction, and harassment. These data included only events with recorded physical assaults and any cases with stalking, sexual assault, and young POIs were not included.

Permission to access the narratives was granted by the NSWPF following ethics approval from the University of New South Wales Human Research Ethics Committee (Ref: HC16558). Due to the inclusion of sensitive and personal information (eg, name, surname, and address) in the narratives, all processing was undertaken at the NSWPF headquarters. Only de-identified,

XSL•FO
RenderX

extracted outputs were allowed to be taken offsite for further analysis.

We used a total of 300 narratives for training, development (used to enhance the performance of the rules), and evaluation purposes (100 each). These sets are described in more detail in our previous work [13]. A hypothetical de-identified narrative is shown in Figure 1.

Categorizing Abuse Types

We categorized specific abuse types (ie, details of the abuse behavior) using several sources into nine categories [12,29,30] with 44 abuse types (Table 1). Although the provided data did not include domestic violence events involving sexual assault and stalking, there were still cases wherein these types of abuse were described in an event. Several nonspecific forms of violence (eg, "bashing," "smack," "assaulted," and "clipping") were categorized as "assault (unspecified)." A more detailed explanation of the abuse types is provided in Multimedia Appendix 1. A total of 17 common injury types were examined, including scratching, grazing, red mark/sign, tear off (nail), bruising, cut/abrasion, swelling, lump, other, fracture, black eye, broken tooth, burn mark, stab wound, bite mark, soreness, and bleeding.

Figure 1. A hypothetical example of a domestic violence event narrative as recorded by the New South Wales Police Force. Blue-highlighted terms indicate the annotated victim injuries, and yellow-highlighted terms indicate the abuse types.

-- Event XXXXXXXXX : -- Created : XXXXXXXXXXXXXXXXXXXXX -- Narrative 1 of 3 –

The victim named XXX and the defendant, XXX have been living together in XXX, XXX with the victims two children, age XXX and XXX. They have been involved romantically for three years. Due to his frequent drug abuse, the defendant has psychotic episodes in which he physically abuses the victim.

In XXX, around morning, the defendant asked the victim to make breakfast. After the victim went to the kitchen to start preparing breakfast, the defendant snapped and flipped over the dining table screaming " I told you to make breakfast!". The victims daughter started crying and the defendant yelled "make that brat quiet, or I will!". The victim asked the defendant if he has taken his medication to which he replied that he did. Just when the victim was going upstairs to put laundry, the defendant stood in front of her and slapped her with his left hand across her face. He then began to kick the victim before placing his hands around the victims throat and starting choking her.

The victim screamed " what the hell is wrong with you?". The defendant then started crying and saying that he was sorry and he did not mean anything. At this point, the victim called her mother and left the house with her two children going immediately at the police station in XXX. The police observed a large bruise on the right side of the victims face and red marks around her throat.

INJURIES: Victim has a bruise on her face and red marks in the throat area.

EVIDENCE AND EFFECTS OF ALCOHOL & DRUGS: Nil

MENTAL HEALTH & OTHER ISSUES: The defendant suffers from paranoid schizophrenia and is an alcoholic.

FIREARMS / DANGEROUS WEAPONS: kitchen knife

FEARS HELD BY THE VICTIM: Victim fears that the accused will continue to physically and emotionally abuse her and is afraid that he will hurt her children.

ACTIONS TAKEN BY THE POLICE: Victim statement recorded. The defendant was arrested and charged.

XSL•FO
RenderX

Table 1. Categories of abuse along with abuse types.

Abuse category	Abuse type
Physical assault	Assault (unspecified), biting, blocking, choking, ordered dog attack, dragging, elbowing, attempting to set fire to premises, gagging, grabbing, hair pulling, headbutting, head locking, kicking, kneeing, physical restraining, pulling, punching, pushing, scratching, shaking, slapping, spitting, stabbing, victim being thrown around, limb twisting, attempt to harm a victim with an object or weapon, and hitting the victim with an object or weapon
Threat	Intimidation (via body language) or stating explicit threat(s) to physically harm, sexually assault, and self-harm if the victim does not comply
Sexual assault	Sexual assault (eg, rape)
Emotional/verbal abuse	Self-harming when the victim does not comply, yelling profanities, and other emotional/verbal abuse
Stalking	Stalking, harassment, and forced entry
Financial abuse	Financial control (eg, no access to credit card)
Social abuse	Social restriction and prevent/limit child access
Unclassified	Apprehended Domestic Violence Order breach, chasing, lunging, other, and possession of personal effects (eg, phone and car keys)
Property damage	Property damage (ranging from breaking an item to causing damage to a house or vehicle)

Rule-Based System Development

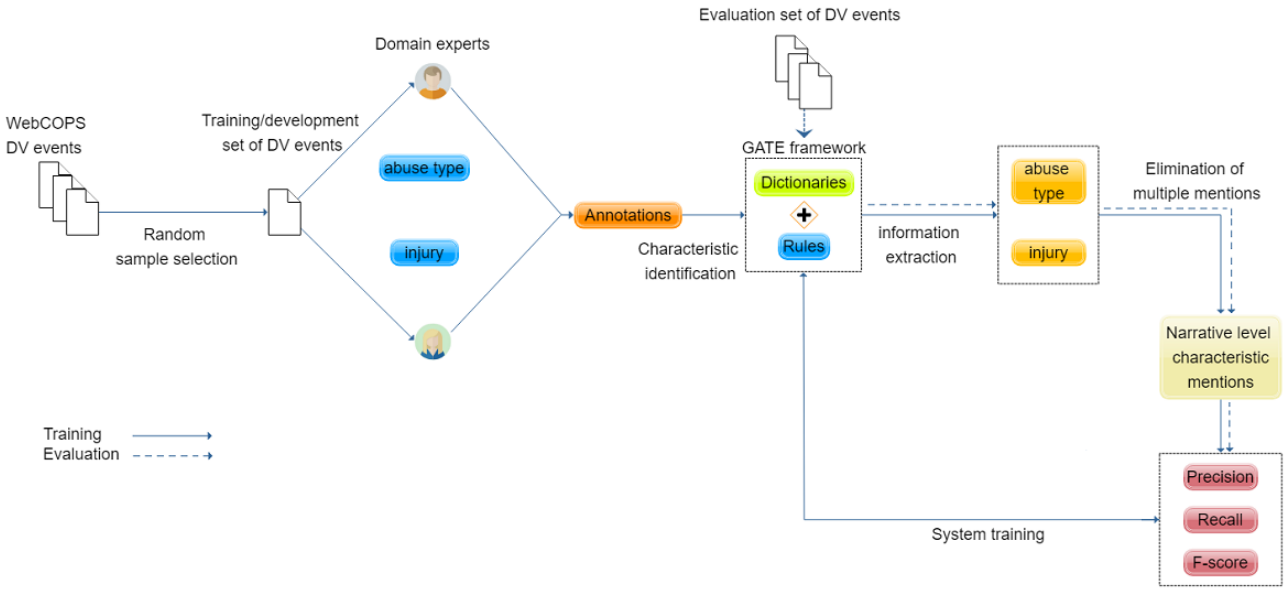
Overview

Our method involved the design and implementation of rule-based language expression patterns combined with dictionary terms for the recognition of abuse types and victim injuries at the narrative level. It consisted of the following steps (Figure 2): (1) creation of relevant dictionaries to recognize mentions of abuse types and victim injuries, (2) design and implementation of rules to capture abuse types and victim injuries mentions in context, and (3) aggregation of multiple mentions in each narrative to reach domestic violence event–level annotation.

Dictionaries

We recognized mentions of task-specific semantic groups through the development of 22 custom-made dictionaries (Table 2). The dictionaries were manually crafted by inspecting the training and the development sets for terms and expressions that describe abuse types (conducted by POIs) and victim injuries, by the first author (GK) and checked by two other authors (AA and PS) to ensure consistency. We used systematic variation (such as plural, past, and present tenses) and also included common misspellings (eg, "stuck" instead of "struck," "harassment," and "assalting") frequently present in the narratives. Although the majority of the terms are noun phrases, for the "threat" dictionary, we included verbal threats made by POIs and manually expanded variations by changing a noun (eg, "your kids are going to have no *father*" to "your kids are going to have no *mother*") and the surface expressions ("*your* dead" to "*you're* dead" or "*you are* dead").

Figure 2. An overview of the text-mining methodology used for the identification of abuse types and victim injuries from domestic violence police event narratives. DV: domestic violence; GATE: General Architecture for Engineering; WebCOPS: Web Computerised Operational Policing System.



XSL•FO
RenderX

Table 2. The manually crafted dictionaries and their respective size (number of terms included) used to identify abuse types and victim injuries.

Dictionary name	Size	Description	Examples
Anatomy	108	Anatomical parts of the human body in which a victim has been injured by the POI[a]	Chest, leg, head, neck
Assault	18	Verbs that indicate a nonspecific physical attack	Attacked, clipped, smacking, bashing
Attempt	6	Verbs that suggested a physical effort by the POI to harm the victim	Attempted, aimed, trying, tried
Be	4	Conjugations of the verb "be" in the present and past tense	Is, was, were, are
Confiscate	8	Verbs describing a confiscating act by an offender towards a victim	Confiscated, grabbed, snatched, grabbing
Damage	22	Verbs indicating an act of property damage by the POI	Cracked, burned, shuttering, ripping
Degree	14	Adjectives describing the victim's wound	Superficial, extensive, minor, major
Description	59	Terms (mostly adjectives) describing various attributes of an object such as color or type of made material	Yellow, wooden, serving, frying
Family	31	Various nouns indicating the relationship between individuals	Boyfriend, mother, father, cousin
First person threats	123	Threats made by the POI towards a victim	"I will kill you," "I am going to bury you," "I will hunt you down and kill you," "someone is going to kill you"
Force	8	Verbs describing an offender physically restrain a victim	Forcing, pinned, pinning, kept
Location	15	House locations that a DV[b] event occurred at	Toilet, loungeroom, wall, hallway
Number	10	Numbers in words suggesting the number of criminal counts charged at an offender	One, two, four, six
Object	174	Various objects that were broken or used in a DV event	Table leg, cup, rear door, window
POI	18	Terms that describe an offender in a DV event	Defendant, person of interest (offender), offender accused
Premises	6	Terms describing a residence	Unit, terrace, flat, premises
Preposition	44	Various prepositions suggesting the presence of a victim's injury in an anatomical part	Under left, lower, upper, front
Start	7	Verbs suggesting the initiation or continuation of an action by the offender	Begun, commenced, continuing, started
Trauma	14	Terms indicating a wound caused by a weapon/object used by the offender towards a victim	Wound, cut, trauma, fracture
Victim	19	Terms describing a victim in a DV event	Victim, vic, pinop (short for person in need for protection), pn (short for pinop),
Weapon	155	Objects used to cause harm or threaten to cause harm to a victim by an offender	Army knife, torch, book, shotgun

[a]POI: person of interest.

[b]DV: domestic violence.

Rules

We based our rules on syntactical patterns identified in the training and development sets, indicating the presence of an abuse type or victim injury. This work follows the same methodology that we previously developed [13]. The syntactical patterns included frozen syntactical expressions as anchors for certain elements built through specific verbs, noun phrases, and prepositions (eg, "commenced to choke") and semantic placeholders identifiable through the application of the manually crafted dictionaries (all possible synonyms describing a victim, such as "victim," "vic," and "pinop"). We specifically utilized concept enumeration, since it frequently appeared in the training and development sets (eg, "Injuries: *Swollen hand*, *soreness* and *scratch under left eye* [mentions of victim's injuries]").

General Architecture for Text Engineering (GATE) [31], a text mining framework for annotating and categorizing text, enabling

XSL•FO
RenderX

information recognition, was used to create and apply our rules. The observed syntactical patterns were converted into rules via Java Annotations Pattern Engine, GATE's pattern-matching language. A total of 64 rules were created (Multimedia Appendix 2).

Elimination of Multiple Mentions

More than one syntactical pattern may be matched in an event narrative and may refer to one or more mentions of abuse types of victim injuries (that can be duplicates). This led to the extraction of highly variable mentions of abuse types and victim injuries (eg, "punch," "punched," and "punching" are variations of the same abuse type ["punching"]; "bruised," "bruises," and "purple marks" are variations of the same injury ["bruising"]). Each mention is therefore mapped to its "canonical" representative, and only one mention for each abuse type or injury is kept and used to "tag" the domestic violence narrative. For example, if, in a domestic violence event report, we have extracted three mentions of the abuse type "punching" and two mentions of the abuse type "kicking," we only annotate two abuse types—"punching" and "kicking"—at the domestic violence event level.

Results

Evaluation

The text mining system was evaluated on a set of 100 previously unseen, randomly chosen domestic violence event reports. The set was manually inspected and annotated by the first and second authors (GK and AA) who identified the type(s) of abuse and victim injuries. The inter-annotator agreement calculated as the absolute agreement rate [32] was 91%, suggesting reliable annotations. Performance of the methodology was evaluated at the narrative level (after eliminating any multiple characteristic mentions). We calculated the precision (the number of true positives against the number of true positives and false positives), recall (the number of true positives against the number of true positives and false negatives), and F1-score (the harmonic mean between precision and recall) at the domestic violence event level using standard definitions [33]. We defined true positive as the detection of a correct mention in an event; false positive as the extraction of any unrelated mention that has not been annotated manually; false negative as the correct mention that was not detected by our method; and true negative as the case where our method did not identify any mentions when none were annotated.

The results are shown in Table 3. Injuries and abuse types returned F1-scores above 85%, suggesting reliable and consistent results with small but expected drops from the training (5.5% and 9.6%, respectively) and development sets (3.9% and 6.7%, respectively). In particular, the precision was 90.2% for abuse types and 85.0% for the victim injuries, with a small decrease from the development set (2.6% and 5.2%, respectively). In a similar fashion, recall was 89.6% and 86.3% for the abuse types and victim injuries, respectively, with a drop of 5.2% and 8.0%, respectively, when compared to the values of the development set. However, the evaluation set had a significantly smaller number of victim injury mentions (n=66) from the development set (n=88) and the training set (n=83); therefore, its recall value should be considered with caution.

Large-Scale Corpus Analysis

Given the relatively accurate results of the method in identifying abuse types and victims' injuries, we applied the method to the corpus of 492,393 domestic violence events. Over 71.32% of events (351,178) had an identified abuse type as mentioned in the report, whereas more than one-third of those events (177,607; 36.07%) contained a victim injury (Table 4).

Of the 44 abuse types, "emotional/verbal abuse" (117,488; 33.46%) was the most common, followed by "punching" (86,322; 24.58%) and "property damage" (78,203; 22.27%). A total of 35.45% (124,498 events) of domestic violence events contained only one identified abuse type, whereas 33.83% (118,819 events) of domestic violence events included three to five different abuse types (Table 5).

The most frequent injury type was "bruising" (51,455; 29.03%), followed by "cut/abrasion" (51,284; 28.93%) and "red marks/signs" (42,038; 23.71%) (Table 6). A total of 105,493 domestic violence events (59.56%) had only one form of injury, and 24.48% (43,373) of domestic violence events had two forms of injury (Table 7).

Table 3. Performance of the system on the training, development, and evaluation sets for the identification of abuse types and victim injuries with true positive, false positive, and false negative results.

Set and characteristic	Precision (%)	Recall (%)	F1-score (%)	True positive (%)	False positive (%)	False negative (%)
Evaluation						
Abuse type	90.2	89.6	89.8	259	28	30
Injury	85.0	86.3	85.6	57	10	9
Development						
Abuse type	92.8	94.8	93.7	310	24	17
Injury	90.2	94.3	92.3	83	9	5
Training						
Abuse type	93.9	96.3	95.3	293	19	11
Injury	93.1	97.5	95.2	81	6	2

XSL•FO
RenderX

Table 4. Number of domestic violence events containing various abuse types (n=351,178).

Abuse type	Events, n (%)
Assault (unspecified)	171,323 (48.79)
Emotional/verbal abuse	117,488 (33.46)
Punching	86,322 (24.58)
Property damage	78,203 (22.27)
Intimidation	75,662 (21.55)
Grabbing	66,728 (19.00)
Pushing	62,794 (17.88)
Scratching	20,493 (5.84)
Physical restraining	20,014 (5.70)
Kicking	19,435 (5.53)
Slapping	17,474 (4.98)
ADVO[a] breach	16,903 (4.81)
Attempting to hit with an object or weapon	13,592 (3.87)
Hair pulling/dragging by hair	13,048 (3.72)
Choking	11,325 (3.22)
Spitting	9341 (2.66)
Hitting with an object or weapon	8387 (2.39)
Other	7135 (2.03)
Pulling	6373 (1.81)
Victim being thrown around	5255 (1.50)
Lunging	4685 (1.33)
Possession of personal effects	3265 (0.93)
Blocking	3163 (0.90)
Harassment	3100 (0.88)
Stalking	2940 (0.84)
Self-harming	2597 (0.74)
Biting	2285 (0.65)
Dragging	2216 (0.63)
Shaking	2098 (0.60)
Stabbing	1903 (0.54)
Forced entry	1779 (0.51)
Headlocking	1482 (0.42)
Chasing	1324 (0.38)
Kneeing	1321 (0.38)
Gagging	1161 (0.33)
Elbowing	225 (0.06)
Limb twisting	173 (0.05)
Headbutting	148 (0.04)
Sexual assault	125 (0.04)
Prevent child access	91 (0.03)
Social restriction	40 (0.01)
Financial control	29 (0.01)

XSL•FO
RenderX

Abuse type	Events, n (%)
Attempting to set fire to premises	28 (0.01)
Ordered dog attack	1 (0.00)

[a]ADVO: Apprehended Domestic Violence Order.

Table 5. Domestic violence events according to the number of abuse types (n=351,178).

Number of abuse type(s)	Events, n (%)
1	124,498 (35.45)
2	89,342 (25.44)
3-5	118,819 (33.83)
6-9	17,951 (5.11)
>10	568 (0.16)
Total	351,178 (100.0)

Table 6. Number of events containing various injury types (n=177,607).

Injury type	Events, n (%)
Bruising	51,455 (29.03)
Cut/abrasion	51,284 (28.93)
Red mark(s)	42,038 (23.71)
Swelling	32,581 (18.38)
Soreness	26,729 (15.08)
Other	19,778 (11.16)
Bleeding	19,154 (10.81)
Fracture(s)	17,531 (9.89)
Lump	9482 (5.35)
Grazing	7305 (4.12)
Black eye(s)	2994 (1.69)
Scratching	2399 (1.35)
Bite mark(s)	2350 (1.33)
Stab wound(s)	2346 (1.32)
Burn mark(s)	1382 (0.78)
Broken tooth	620 (0.35)
Tear off nail(s)	7 (0.00)

Table 7. Domestic violence events according to the number of victim injury types (n=177,607).

Number of injury types	Events, n (%)
1	105,493 (59.56)
2	43,373 (24.49)
3-4	25,678 (14.49)
5-6	2484 (1.40)
≥7	89 (0.05)
Total	177,117 (100.0)

XSL•FO
RenderX

Discussion

Principal Results

To the best of our knowledge, this analysis represents the first attempt to capture domestic violence–related abuse and victim injuries using a large, population-level corpus of domestic violence events recorded by the police. The identification of abuse types conducted by POIs and various injuries sustained by victims in domestic violence disputes are not recorded in the structured information of the WebCOPS database fields. We therefore focused on the narrative part, where the application of our knowledge-driven approach has identified rich information and has the potential to be used for better understanding domestic violence and the development of related prevention interventions, surveillance, and reporting.

Our findings derived from text mining present a more detailed picture of the types of injuries and abuse occurring in domestic violence events. The most common abuse type in our dataset was nonphysical and involved "emotional/verbal abuse," which is consistent with the recent findings showing that nonphysical abuse types are more prevalent than physical ones [34] and that victims of domestic violence abuse are more likely to sustain certain types of injuries such as cuts and fractures than others [34,35]. Domestic violence can also take myriad physical forms, ranging from victim intimidation to cases where serious and grievous bodily harm is caused by a specific type of abuse (eg, "punching," "stabbing," and "choking"), which have both short- and long-term physical and mental health consequences [9-11].

Through the recognition of various abuse types and related victim injuries, potential exists to develop prevention and intervention guidelines by linking this information to diagnostic data held by health services, so that surveillance and monitoring of the victims can be performed. There is also a possibility to track any potential timelines in which the victim was abused. Moreover, the text mining method can be updated on an ongoing basis to monitor trends and inform risk stratification algorithms, which can drive domestic violence–prevention strategies targeting specific groups.

With the inclusion of domestic violence in the WHO's Sustainable Development Goals, the need for accurate reporting in this area will be necessary [36]. Text mining the police's domestic violence event narratives is possibly a source of obtaining very nuanced information on this topic including the cause of the event, the potential role of mental illness and substance (ab)use in the event, the types of abuse perpetrated, injuries sustained, weapons used, and information on relationship status. This invaluable information can then be used to target prevention strategies for use by those providing prevention services to particular groups and to identify warning signs for health care providers. A recent report indicated that in Australia, from 2012-2013 to 2013-2014, one woman was killed each week and one man was killed each month as a result of violence from a current or previous partner [7]. Subsequent analyses of this rich information will aim to examine these issues and identify early warning signs of abuse and domestic violence events, which may improve assistance in preventing homicides in domestic violence settings.

Error Analysis

Although the level of accuracy was acceptable for large-scale analysis to identify trends in domestic violence events, there were still some errors in both abuse types and victim injuries at the level of individual narrative reports. By inspecting the evaluation set, we observed that the system erroneously extracted few instances (five cases) of several POI injuries as victim injuries, since the rules were triggered for the POIs (eg, *"minor grazing to the right shoulder* [false positive for injury] of the POI"). In other instances (4 cases), victim injuries were incorrectly identified when they actually referred to property damage through ambiguous word and syntactical pattern combinations that indicated an injury (eg, "INJURIES/MEDICAL TREATMENT/DAMAGE TO PROPERTY: *Broken table leg* [false positive for victim injury]"). In 12 domestic violence events, when a victim fought back against a POI, any actions by the victim in self-defence were erroneously extracted as an abuse type (eg, "witness stepped in and *grabbed* [false positive for abuse type] the POI and *pinned him to the ground* [false positive for abuse type] until he calmed down" and "...has admitted she physically *pushed him* [false positive for abuse type] back after he *pushed* [true positive for abuse type] into her"). There were few occasions where an abuse type was recognized but had no domestic violence context (eg, "The Accused was closed inside the caged area, where he began *kicking* [false positive for abuse type] at the door and yelling at the police officers..."), while others had not occurred but were likely to happen in the future (eg, "The victim believes if she stayed at the residence she would definitely have been *bashed* [false positive for abuse type] by the accused and possibly *stabbed* [false positive for abuse type]").

Although we engineered the rules based on generic syntactical patterns that stated victim injuries and abuse types, these rules ignored a limited number of injury mentions, since they were not explicitly stated to have been sustained by the victim (eg, *"redness* [false negative for injury] and *grazes* [false negative for injury] sighted on back, *dried blood* [false negative for injury] on lips"). Some examples (eight cases) were more implicit and required additional inference using some related terms (eg, "the POI *placed his hand in the middle of the victim's sternum and applied force* [false negative for injury] causing her pain and shortness of breath"). Cases like these were the majority of false negatives for abuse types, suggesting that abuse types such as "grabbing" and "punching" can have quite a few lexical variations in the narratives, which indicate richness of the contexts.

Additionally, injury or abuse type mentions (six cases) that were accompanied by the victim's surname were excluded from our rule design, since there was no way to determine from the narrative who was the victim or POI without using the structured part of the record (eg, "xxx had a *bleeding nose* [false negative for injury]" and "xxx yelled *verbal abuse* [false negative for abuse type] at her").

Limitations

Our text mining system could have missed cases due to more specialized or explicit mentions of abuse types occurring in

XSL•FO
RenderX

domestic violence events, since we based our extraction rules on the information contained in only 200 narratives. Despite incorporation of all types of abuse, there are still likely to be cases in which we probably did not identify explicit types. The relatively smaller number of injury mentions in the evaluation set (when compared to one of the abuse types) could explain the lower performance for the injuries. Nonetheless, we designed our rules based on common syntactical patterns that would attribute abuse types/injury mentions toward POIs and victims, respectively, in order to avoid the generation of false negatives; hence, our recall was higher than the precision in all three datasets. Nevertheless, this approach was able to identify the victim's actions as types of POI's abuse as well as POI's injuries as those of the victim in some instances. This suggests that more specific engineered rules could address this issue. Similarly, although we included the basic and most common forms of injuries, there would be instances containing other causes of injuries or particular abuse types leading to specific injuries that probably have been excluded from our approach. Additionally, the implementation of spell-checking algorithms could assist in the identification of any misspelled abuse types or injuries and potentially elevate performance.

Our analysis of the results from the large corpus of domestic violence events is limited to the abuse types and victim's injuries. We plan to use this information in combination with administrative data collections on mental illness to further examine the nexus between mental illness and domestic violence and explore the relationship of abuse types with gender and victim injuries. It is pertinent to inquire whether domestic violence victims with mental illness are more vulnerable than those without mental illness in this large-scale dataset spanning 10 years, to identify new intel. Further analysis of the results combined with demographic variables can show interesting aspects of the data in the area of the domestic violence, from prevalence and incidence rates in specific cohorts to risk factors for the occurrence (or recurrence) of domestic violence events. The combination of victim injuries from clinical data resulting from health service contacts could assist in the early identification of victim abuse and the implementation of intervention strategies. Modelling will be used to investigate whether POI characteristics can predict severity of abuse and similarly, whether certain victim phenotypes are prone to particular types of abuse.

Conclusions

We demonstrated that a knowledge-driven approach can be used for the automated extraction of abuse types and victim injuries involved in domestic violence events. The performance was encouraging, with 90.2% and 85.0% precision for abuse types and injuries, respectively, further implicating that text mining can be used to extract meaningful information from these unstructured data on a large scale. The identified information has enabled us to confirm the magnitude of abuse that victims endure during domestic violence. The results can be used to support further public health research that aims to assess the profiling of POIs involved in domestic violence events and to alter existing intervention policies for victims of abuse.

Acknowledgments

The authors would like to thank the NSWPF for their assistance with this project, particularly Dr Chris Devery, Dr Christie Wallace, John Blanchette, Erin Sharland, and Nicole Grant. This research was supported by an Australian Institute of Criminology/Criminology Research Grant (34/15-16).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Brief description of the extracted abuse types.

[PDF File (Adobe PDF File), 215KB - jmir_v21i3e13067_app1.pdf]

Multimedia Appendix 2

Rule examples for recognition of abuse types and victim injuries.

[PDF File (Adobe PDF File), 192KB - jmir_v21i3e13067_app2.pdf]

References

1. Howard LM, Trevillion K, Khalifeh H, Woodall A, Agnew-Davies R, Feder G. Domestic violence and severe psychiatric disorders: prevalence and interventions. Psychol Med 2010 Jun;40(6):881-893. [doi: 10.1017/S0033291709991589] [Medline: 19891808]
2. Robinson L, Spilsbury K. Systematic review of the perceptions and experiences of accessing health services by adult victims of domestic violence. Health Soc Care Community 2008 Jan;16(1):16-30. [doi: 10.1111/j.1365-2524.2007.00721.x] [Medline: 18181812]
3. Trevillion K, Oram S, Feder G, Howard LM. Experiences of domestic violence and mental disorders: a systematic review and meta-analysis. PLoS One 2012;7(12):e51740 [FREE Full text] [doi: 10.1371/journal.pone.0051740] [Medline: 23300562]

XSL•FO
RenderX

4. Home Office Statistical Bulletin. 2008. Crime in England and Wales 2007/08: Findings from the British Crime Survey and police recorded crime URL: http://news.bbc.co.uk/2/shared/bsp/hi/pdfs/17_07_08_crime_statistics_200708.pdf [accessed 2019-02-22] [WebCite Cache ID 76MoKgYtI]
5. Briodi A. Sydney City Council and NSWPF. 2010. Domestic Violence is a crime electronic resource URL: https://catalogue.nla.gov.au/Record/4984779 [accessed 2018-07-11] [WebCite Cache ID 70p4JfaZT]
6. Australian Government - Department of Social Services. The National Plan to Reduce Violence against Women and their Children 2010-2022 URL: https://www.dss.gov.au/women/programs-services/reducing-violence/the-national-plan-to-reduce-violence-against-women-and-their-children-2010-2022 [accessed 2019-02-28] [WebCite Cache ID 76X39GmnR]
7. Australian Institute of Health and Welfare. 2018. Family, domestic and sexual violence in Australia URL: https://www.aihw.gov.au/reports/domestic-violence/family-domestic-sexual-violence-in-australia-2018/contents/summary [accessed 2018-12-07] [WebCite Cache ID 74TboPEi5]
8. Foshee V. Gender differences in adolescent dating abuse prevalence, types and injuries. Health Educ Res 1996;11(3):275-286. [doi: 10.1093/her/11.3.275-a]
9. Kelly J, Johnson M. Differentiation among types of intimate partner violence: Research update and implications for interventions. Family Court Review 2008 Jul;46(3):476-499. [doi: 10.1111/j.1744-1617.2008.00215.x]
10. Capaldi D, Shortt J, Kim H, Wilson J, Crosby L, Tucci S. Official incidents of domestic violence: Types, injury, and associations with nonofficial couple aggression. Violence and Victims 2009;24(4):502. [Medline: 19694354]
11. Cleak H, Schofield M, Axelsen L, Bickerdike A. Screening for Partner Violence Among Family Mediation Clients: Differentiating Types of Abuse. J Interpers Violence 2018 Apr;33(7):1118-1146. [doi: 10.1177/0886260515614559] [Medline: 26681786]
12. KPMG. 2016. The cost of violence against women and their children in Australia URL: https://www.dss.gov.au/sites/default/files/documents/08_2016/the_cost_of_violence_against_women_and_their_children_in_australia_-_summary_report_may_2016.pdf [accessed 2019-03-01] [WebCite Cache ID 74TbjRYfR]
13. Karystianis G, Adily A, Schofield P, Knight L, Galdon C, Greenberg D, et al. Automatic Extraction of Mental Health Disorders From Domestic Violence Police Narratives: Text Mining Study. J Med Internet Res 2018 Sep 13;20(9):e11548. [doi: 10.2196/11548] [Medline: 30213778]
14. Macdonald W, Fitzgerald J. NSW Government: Justice – Bureau of Crime Statistics and Research. 2014. Understanding fraud: The nature of fraud offences recorded by NSW Police URL: https://www.bocsar.nsw.gov.au/Pages/bocsar_pages/Fraud.aspx [accessed 2019-02-22] [WebCite Cache ID 76Moaevm8]
15. Abbe A, Grouin C, Zweigenbaum P, Falissard B. Text mining applications in psychiatry: a systematic literature review. Int J Methods Psychiatr Res 2016 Dec;25(2):86-100. [doi: 10.1002/mpr.1481] [Medline: 26184780]
16. Friedman C, Shagina L, Lussier Y, Hripcsak G. Automated encoding of clinical documents based on natural language processing. J Am Med Inform Assoc 2004;11(5):392-402 [FREE Full text] [doi: 10.1197/jamia.M1552] [Medline: 15187068]
17. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. J Am Med Inform Assoc 2010;17(5):507-513 [FREE Full text] [doi: 10.1136/jamia.2009.001560] [Medline: 20819853]
18. Spasić I, Livsey J, Keane JA, Nenadić G. Text mining of cancer-related information: review of current status and future directions. Int J Med Inform 2014 Sep;83(9):605-623 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.06.009] [Medline: 25008281]
19. Wang Y, Wang L, Rastegar-Mojarad M, Moon S, Shen F, Afzal N, et al. Clinical information extraction applications: A literature review. J Biomed Inform 2018 Jan;77:34-49 [FREE Full text] [doi: 10.1016/j.jbi.2017.11.011] [Medline: 29162496]
20. Karystianis G, Dehghan A, Kovacevic A, Keane JA, Nenadic G. Using local lexicalized rules to identify heart disease risk factors in clinical notes. J Biomed Inform 2015 Dec;58 Suppl:S183-S188 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.013] [Medline: 26133479]
21. Matto G, Mwangoka J. Detecting crime patterns from Swahili newspapers using text mining. IJKEDM 2017;4(2):145-156. [doi: 10.1504/IJKEDM.2017.086244]
22. Nokhbeh Zaeem R, Manoharan M, Yang Y, Barber K. Modeling and analysis of identity threat behaviors through text mining of identity theft stories. Computers & Security 2017 Mar;65:50-63. [doi: 10.1016/j.cose.2016.11.002]
23. Arulanandam R, Savarimuthu B, Purvis M. Extracting crime information from online newspaper articles. 2014 Jan 20 Presented at: Proceedings of the Second Australasian Web Conference-Volume 155; 2014; Auckland, New Zealand.
24. Chau M, Xu J, Chen H. Extracting meaningful entities from police narrative reports. 2002 May 19 Presented at: Proceedings of the annual national conference on Digital government research; 2002; Los Angeles, USA.
25. Ananyan S. AMCIS 2004 Proceedings. 2004. Crime pattern analysis through text mining URL: https://aisel.aisnet.org/amcis2004/236/ [accessed 2019-03-01] [WebCite Cache ID 76YY4WcQD]
26. Poelmans J, Elzinga P, Viaene S, Dedene G. Formally analysing the concepts of domestic violence. Expert Systems with Applications 2011 Apr;38(4):3116-3130. [doi: 10.1016/j.eswa.2010.08.103]

XSL•FO
RenderX

27. Ku C, Iriberri A, Leroy G. Crime information extraction from police and witness narrative reports. 2008 Presented at: Technologies for Homeland Security, IEEE Conference on; 2008; Boston, USA.
28. Iriberri A, Leroy G. Natural language processing and e-government: Extracting reusable crime report information. 2007 Presented at: Information Reuse and Integration, IEEE International Conference; 2007; Las Vegas, USA.
29. White Ribbon Australia. 2018. Physical Abuse URL: https://www.whiteribbon.org.au/understand-domestic-violence/types-of-abuse/physical-abuse/ [WebCite Cache ID 74TbzEdGO]
30. Mouzos J, Makkai T. Australian Institute of Criminology. 2004. Women's experiences of male violence: findings from the Australian component of the International Violence Against Women Survey (IVAWS) URL: https://aic.gov.au/publications/rpp/rpp56 [accessed 2019-03-01] [WebCite Cache ID 76YYLb70T]
31. Cunningham H, Tablan V, Roberts A, Bontcheva K. Getting more out of biomedical documents with GATE's full lifecycle open source text analytics. PLoS Comput Biol 2013;9(2):e1002854 [FREE Full text] [doi: 10.1371/journal.pcbi.1002854] [Medline: 23408875]
32. Ananiadou S, McNaught J, editors. Text Mining for Biology and Biomedicine. Boston, MA: Artec House Publishers; 2006.
33. Ananiadou S, Kell DB, Tsujii J. Text mining and its potential applications in systems biology. Trends Biotechnol 2006 Dec;24(12):571-579. [doi: 10.1016/j.tibtech.2006.10.002] [Medline: 17045684]
34. Outlaw M. No One Type of Intimate Partner Abuse: Exploring Physical and Non-Physical Abuse Among Intimate Partners. J Fam Viol 2009 Feb 27;24(4):263-272. [doi: 10.1007/s10896-009-9228-5]
35. Muelleman R, Lenaghan P, Pakieser R. Battered women: injury locations and types. Annals of emergency medicine 1996;28(5):486-492. [Medline: 8909268]
36. World Health Organization. The Sustainable Development Goals (SDG) and violence prevention: how do they connect? 2018; URL: https://www.who.int/violence_injury_prevention/violence/7th_milestones_meeting/Butchart_SDGs_and_violence_prevention.pdf?ua=1 [accessed 2018-12-07] [WebCite Cache ID 74Tc66boY]

Abbreviations

ADVO: Apprehended Domestic Violence Order
DV: domestic violence
GATE: General Architecture for Engineering
NSWPF: New South Wales Police Force
POI: person of interest
WebCOPS: Web Computerised Operational Policing System

Edited by G Eysenbach; submitted 11.12.18; peer-reviewed by I Spasic, A Davoudi; comments to author 05.01.19; revised version received 31.01.19; accepted 10.02.19; published 12.03.19.

Please cite as:
Karystianis G, Adily A, Schofield PW, Greenberg D, Jorm L, Nenadic G, Butler T
Automated Analysis of Domestic Violence Police Reports to Explore Abuse Types and Victim Injuries: Text Mining Study
J Med Internet Res 2019;21(3):e13067
URL: http://www.jmir.org/2019/3/e13067/
doi:10.2196/13067
PMID:30860490

©George Karystianis, Armita Adily, Peter W Schofield, David Greenberg, Louisa Jorm, Goran Nenadic, Tony Butler. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient Judgments About Hypertension Control: The Role of Variability, Trends, and Outliers in Visualized Blood Pressure Data

Victoria Anne Shaffer[1], PhD (Quantitative Psychology); Pete Wegier[2,3], PhD; KD Valentine[1], MS; Jeffery L Belden[4], MD; Shannon M Canfield[4], MPH; Sonal J Patil[4], MD; Mihail Popescu[5], PhD; Linsey M Steege[6], PhD; Akshay Jain[7], MS; Richelle J Koopman[4], MD, MS

[1]University of Missouri, Department of Psychological Sciences, Columbia, MO, United States
[2]Sinai Health System, Temmy Latner Center for Palliative Care, Toronto, ON, Canada
[3]Sinai Health System, Lunenfeld-Tanenbaum Research Institute, Toronto, ON, Canada
[4]University of Missouri, Department of Family & Community Medicine, Columbia, MO, United States
[5]University of Missouri, Department of Health Management & Informatics, Columbia, MO, United States
[6]University of Wisconsin, School of Nursing, Madison, WI, United States
[7]University of Missouri, Department of Electrical and Computer Engineering, Columbia, MO, United States

Corresponding Author:
Victoria Anne Shaffer, PhD (Quantitative Psychology)
University of Missouri
Department of Psychological Sciences
320 S. 6th St.
Columbia, MO, 65211
United States
Phone: 1 573 882 9094
Email: shafferv@missouri.edu

Abstract

Background: Uncontrolled hypertension is a significant health problem in the United States, even though multiple drugs exist to effectively treat this chronic disease.

Objective: As part of a larger project developing data visualizations to support shared decision making about hypertension treatment, we conducted a series of studies to understand how perceptions of hypertension control were impacted by data variations inherent in the visualization of blood pressure (BP) data.

Methods: In 3 Web studies, participants (internet sample of patients with hypertension) reviewed a series of vignettes depicting patients with hypertension; each vignette included a graph of a patient's BP. We examined how data visualizations that varied by BP mean and SD (Study 1), the pattern of change over time (Study 2), and the presence of extreme values (Study 3) affected patients' judgments about hypertension control and the need for a medication change.

Results: Participants' judgments about hypertension control were significantly influenced by BP mean and SD (Study 1), data trends (whether BP was increasing or decreasing over time—Study 2), and extreme values (ie, outliers—Study 3).

Conclusions: Patients' judgment about hypertension control is influenced both by factors that are important predictors of hypertension related-health outcomes (eg, BP mean) and factors that are not (eg, variability and outliers). This study highlights the importance of developing data visualizations that direct attention toward clinically meaningful information.

(J Med Internet Res 2019;21(3):e11366) doi:10.2196/11366

KEYWORDS

data visualization; hypertension; hypertension control; patients' judgment; primary care

Introduction

Uncontrolled hypertension is a significant health problem; there are 75 million adults in the United States alone with diagnosed hypertension [1-3]. Hypertension control is an important goal in primary care because uncontrolled hypertension is a major risk factor for morbidity and mortality and contributes to heart disease, stroke, and chronic kidney disease [2]. Multiple drugs

XSL•FO
RenderX

exist that effectively treat hypertension, yet hypertension remains uncontrolled in 46% of patients. Several national and regional health initiatives (eg, Healthy People 2020, Million Hearts Initiative, and Community Preventive Task Force) have focused on improving hypertension monitoring and management. However, despite these efforts, data from the National Health and Nutrition Examination Survey in 2011-2014 showed no marked change in the percentage of adults with controlled hypertension [4].

In their hypertension clinical action model, Kerr et al [5] identified 4 factors that predict uncontrolled hypertension—clinical uncertainty, competing demands/prioritization, medication-related factors (eg, side effects), and organizational factors (eg, lack of support to follow and reassess patients). However, the primary reason for failing to intensify medication in a clinic visit was uncertainty about the "true" blood pressure (BP) value. When multiple BP readings were recorded, there was often a discrepancy between the values with some readings falling inside the goal range and others falling outside, leading physicians and patients to question whether the action was warranted. BP measurements that occur both within and above the goal range can be especially perplexing to patients, leading to difficulty in making decisions about BP control [6-8].

To improve hypertension control in the primary care setting, our research team has developed a data visualization tool designed to support shared decision making about hypertension treatment. Our tool is a visual display of the patients' BP data over the last 2 years that will be embedded in the electronic health record (EHR) and will be jointly viewed by patients and clinicians during a primary care office visit. This tool aims to (1) reduce clinical uncertainty about BP data and hypertension control; and (2) increase patients' willingness to intensify their medication or comply with standing treatment plans, with the downstream benefit of improving hypertension control.

The data visualization tool was developed through a rapid prototyping process in which candidate visualizations were iteratively refined on the basis of regular feedback from patient and physician focus groups [9]. During the prototyping process, we concurrently conducted a series of vignette-based Web studies to inform the development of the data visualization tool. Previous research on the presentation of Web-based risk communications, including interactive graphics describing the risk of side effects for thyroid cancer treatments and EHR patient portal displays of laboratory blood test values, demonstrates the value of testing the effect of data visualization interventions on gist knowledge, perceptions of risk, and judgments about health status [10-13]. To date, no research has examined how data visualizations of BP designed for use in the EHR impact patients' perceptions of risk and hypertension control. This paper aims to report the results of 3 studies that examined how data visualizations that varied in BP mean and SD (Study 1), pattern of change over time (Study 2), and the presence of extreme values (Study 3) impact patients' judgments about hypertension control and the need for medication change.

Methods

Study Designs

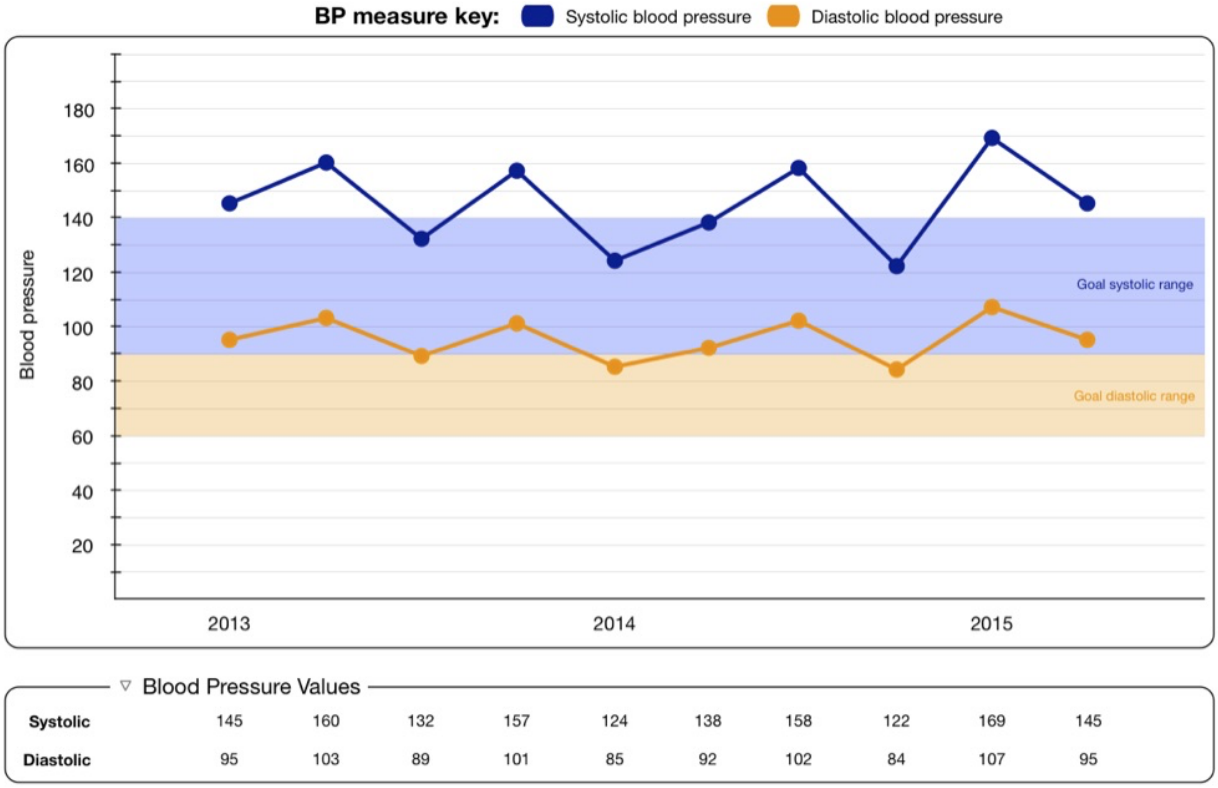
Three demographically diverse internet samples of patients with hypertension reviewed several brief vignettes describing fictitious patients with hypertension through Web survey; each vignette included a graph of the patients' BP data. All 3 studies used a within-subjects design, where all participants reviewed all vignettes, presented in random order, and provided judgments about the degree of hypertension control for every patient/vignette. The Web surveys are described below, and the results are reported in accordance with the Checklist for Reporting Results of Internet E-Surveys (ie, CHERRIES checklist) [14]. The Institutional Review Board at the University of Missouri approved all studies. The 3 studies had very similar methodologies; therefore, we will describe their methods and results together. All samples were recruited by Qualtrics, a survey company that maintains an opt-in demographically diverse internet panel that participates; details about participant recruitment are provided below.

Each vignette described a patient who was being treated for hypertension and included a graph of the patient's BP data over the past 2 years. Figure 1 presents an example of the data visualization tool. In Study 1, 9 vignettes systematically varied in the mean systolic BP (SBP; 130, 145, and 160 mm HG) and BP SD (5, 15, and 25) depicted in the graph; the slope was held constant (Multimedia Appendix 1). The mean SBP was chosen to represent clinical cases that included examples of controlled, uncontrolled, and borderline hypertension according to the 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults from the panel members appointed to the Eighth Joint National Committee [15]. The SDs were chosen to represent small, moderate, and large mean variability according to published SBP values [16].

In Study 2, 6 vignettes differed in the SBP mean (130, 145, and 160 mm HG) and pattern of change over time (ie, trend or slope of the data; increasing or decreasing), while holding the SD constant (15 mm HG; Multimedia Appendix 2). In Study 3, we used 10 vignettes that differed in their mean SBP (130 and 145 mm HG), presence of outliers (no outliers, 1 outlier, or 2 outliers), and positioning of those outliers (above or below the mean; Multimedia Appendix 3).

XSL•FO
RenderX

Figure 1. Sample visualization tool.



Outcomes

Primary outcomes were (1) perceived BP control; (2) need for medication change; (3) subjective risk of heart attack; and (4) subjective risk of stroke for each vignette. Perceived BP control and need for medication change were assessed using agree-disagree Likert scales, while the subjective likelihood of heart attack and stroke were measured using unlikely-likely Likert scales. In addition, we asked participants to estimate the proportion of SBP values out of range for each vignette: "What percentage of the patient's systolic blood pressure values (ie, top number) would you estimate to be out of the goal range?"

After evaluating all vignettes, participants completed the Subjective Numeracy Scale (SNS) [17] and a Single-Item Literacy Screener for health literacy (SILS) [18]. In addition, participants provided demographic information (age, gender, race/ethnicity, education, and income) and responded to 2 additional items about how often they monitor and graph their own BP; no identifying information was collected or stored.

Web Survey

All data were collected by open Web surveys developed using Qualtrics survey development software [19]. Informed consent, task instructions, practice task, vignettes, and study outcomes were all presented through a Web survey. The research team generated vignettes, BP displays, and outcomes for the studies. Of note, no adaptive items were used in the Web surveys. Responses to primary outcomes described above were recorded through slider bar. Responses to the SNS, SILS, and demographic items were all recorded using radio buttons on the Web survey, except age, which was recorded as free text. Responses to all items were required. Furthermore, back buttons on the browsers were disabled during the survey so participants could not go back and review previous webpages or change their responses on previous pages.

In all 3 Web surveys, participants completed one practice vignette that included directions for the practice task and 5 practice items; this section of the Web surveys utilized 3 webpages. Each vignette and its corresponding 5 questionnaire items (hypertension control, need for a medication change, 10-year heart attack risk, 10-year stroke risk, and percentage of SBP values out of goal range) added 2 additional pages to each Web survey (per vignette). The SILS and SNS scales added 9 items, and the demographic variables added 7 items to each Web study. Furthermore, we provided a free textbox at the end of the study for participants to share any additional feedback on the data visualizations presented in the Web studies.

In Study 1, the Web survey included 9 vignettes; therefore, participants completed 68 total items across 27 total webpages. In Study 2, the Web survey included 6 vignettes; participants completed 53 total items across 21 total webpages. In Study 3, the Web survey included 10 vignettes; participants completed 73 total items across 29 total webpages. The presentation order of the vignettes and their associated outcome measures was randomized within subjects. In addition, the team pilot-tested each Web survey themselves for functionality and usability before the surveys were deployed.

Participant Recruitment

Qualtrics identified participants with hypertension through a single self-reported measure: "Has your doctor ever diagnosed you with hypertension, also known as high blood pressure?"; similar self-report items have been used to identify patients with

XSL•FO
RenderX

hypertension in other epidemiological studies [20,21]. Informed consent was received from all participants after they were apprised about the purpose of the survey and the maximum completion time for the survey (30 minutes); participants were also provided with the name and contact information for the Project Director. Data for all 3 studies were collected between May 2016 and July 2016.

We used Internet Protocol addresses and Qualtrics identification numbers to determine unique site visitors. The survey completion rate was 86% (51/59) for Study 1, 79% (50/63) for Study 2, and 66% (55/83) for Study 3. All completed questionnaires were used in the analyses.

Power and Statistical Analyses

We planned to recruit 50 patients with hypertension for each of the 3 Web studies. The sample size was determined *a priori* using G-Power [22,23] with the following data characteristics: 80% power to detect a significant effect (f=0.25) at an alpha of .05, with a minimum correlation of .50 between repeated measures. All outcomes were treated as continuous variables. We examined the effects of data variations on primary and secondary outcomes by conducting a series of multivariate analysis of variance tests for repeated measures. Predictors included the level of the mean SBP (all studies), SBP SD (Study 1), presence and direction of slope (Study 2), and the number and direction of outliers (Study 3). All tests were conducted in SPSS version 24 and R version 3.5 and were considered statistically significant when P<.05.

Results

Participants

Across the 3 experiments, a total of 156 patients with hypertension participated in this research. Participants were majority female (102/156, 65.4%) and white (116/156, 74.4%), with a mean age of 47.30 (range 19-79) years; Table 1 summarizes additional participants' characteristics. In accordance with the sex and gender reporting guidelines [24], we present the data for all 3 studies disaggregated by gender in Multimedia Appendix 4.

In contrast, when the patient had BP that was objectively well controlled (mean 130 mm HG), the variability of the BP data had a significant impact on judgments about hypertension control and related health risks. For example, when the variability was low (SD 5 mm HG), BP was considered reasonably well controlled with little need for a medication change, and there was a much lower perceived risk of heart attack and stroke. In contrast, when variability in BP data was high (SD 25 mm HG), the perception of BP control was similar to those patients who had much higher mean (160 mm HG) BP values. This is noteworthy because evidence suggests that BP mean, much more than variability, is predictive of outcomes that matter to patients [25] (ie, heart attack and stroke). The effects of home BP variability on cardiovascular events and mortality are based on post-hoc analyses of 2 studies that looked at multiple indices of home BP variability [26]. The SD of BP is highly dependent on the mean BP, and it is unclear if indices of BP variability independent of the mean BP incrementally predict cardiovascular mortality or total mortality beyond mean SBP [27-29]. In addition, BP variability has unclear prognostic significance as varying methods or indices have been used to quantify the BP variability in all studies with no current standard or optimal indices available to quantify the BP variability.

Patients in Study 1 were also asked to recall the percentage of SBP values that were "out of range" (ie, exceeded 140 mm HG) in the graph. The recall was largely inaccurate but varied significantly by the SBP mean and SD ($F_{4,200}$=8.73; P<.001; generalized η^2=0.04; Figure 2). With low variability (SD 5), participants overestimated the percent of SBP values out of range when hypertension was controlled (mean 130) and underestimated the percent of SBP values out of range when hypertension was borderline (mean 145) or uncontrolled (mean 160). The estimates of SBP values out of range were more accurate when there was moderate to high variability in the BP data. Figure 2 shows a jellyfish plot of error in SBP recall. Dots above the dotted line represent participant overestimation, and dots below represent underestimation. For each vignette, we show the following: (1) the mean (large black dot) and 95% CI (black line) in the center; (2) a dot plot of all data points on the left; and (3) a kernel-density plot of the distribution of answers on the right.

Study 1—Variations in SBP Mean and Variability

In Study 1, we observed a significant interaction between the SBP mean and SBP SD on perceived BP control ($F_{4,200}$=16.94; P<.001; generalized η^2=0.08), need for medication change ($F_{4,200}$=16.19; P<.001; generalized η^2=0.08), heart attack risk ($F_{4,200}$=8.88; P<.001; generalized η^2=0.04), and stroke risk ($F_{4,200}$=11.70; P<.001; generalized η^2=0.05; see Tables 2 and 3). When the mean SBP was high (160 mm HG), variability in the BP data did not impact the perception of BP control. Across all 3 SDs, participants reported that hypertension was *not* well controlled, patients *should* change their medication, and patients had an *elevated* risk of heart attack and stroke.

XSL•FO
RenderX

Table 1.

Characteristics	Study 1, (n=51)	Study 2, (n=50)	Study 3, (n=53)
Gender, male, n (%)	19 (36)	17 (34)	18 (33)
Age (years)			
Mean (SD)	47.0 (13.1)	44.7 (13.6)	50.2 (14.5)
Range	19-72	20-79	19-77
Race or ethnicity, n (%)			
White or Caucasian	44 (86)	34 (68)	38 (69)
Black or African American	3 (6)	10 (20)	9 (16)
Asian or Pacific Islander	2 (4)	1 (2)	3 (5)
American Indian or Alaskan Native	0	1 (2)	1 (2)
Hispanic or Latino/a	2 (4)	3 (6)	2 (4)
Other	0	1 (2)	2 (4)
Single-Item Literacy Screener for health literacy, mean (SD)	2.0 (1.0)	1.8 (0.8)	2.0 (0.9)
Subjective Numeracy Scale, mean (SD)	4.3 (0.9)	4.3 (1.1)	4.3 (1.0)
Education, n (%)			
Some high school	1 (2)	1 (2)	2 (4)
High school graduate	10 (20)	14 (28)	13 (24)
Some college	16 (32)	12 (24)	16 (29)
Vocational training	4 (8)	7 (14)	4 (7)
Associate’s degree	4 (8)	6 (12)	7 (13)
Bachelor’s degree	9 (18)	5 (10)	8 (15)
Master’s degree	0	3 (6)	4 (7)
Professional degree	5 (10)	1 (2)	0
Doctoral degree	1 (2)	1 (2)	1 (2)
Income (US $), n (%)			
<10k	5 (10)	1 (2)	1 (2)
10-19k	4 (8)	7 (14)	6 (11)
20-29k	7 (13)	7 (14)	13 (24)
30-39k	5 (10)	7 (14)	10 (18)
40-49k	9 (17)	4 (8)	5 (9)
50-59k	4 (8)	9 (18)	8 (15)
60-69k	4 (8)	5 (10)	2 (4)
70-79k	3 (6)	0	3 (5)
80-89k	2 (4)	2 (4)	4 (7)
90-99k	1 (2)	2 (4)	0
100-149k	6 (12)	5 (10)	1 (2)
>149k	1 (2)	1 (2)	2 (4)
How often do you monitor your BP at home?, n (%)			
Never	11 (22)	7 (14)	8 (15)
Annually	4 (8)	4 (8)	4 (7)
Monthly	13 (25)	19 (38)	14 (25)
Weekly	9 (18)	12 (24)	21 (38)
Daily	14 (27)	8 (16)	8 (15)

XSL•FO
RenderX

Characteristics	Study 1, (n=51)	Study 2, (n=50)	Study 3, (n=53)
How often do you graph your home BP measurements?, n (%)			
Never	29 (57)	32 (64)	33 (60)
Annually	4 (8)	3 (6)	3 (5)
Monthly	6 (12)	6 (12)	9 (16)
Weekly	6 (12)	7 (14)	7 (13)
Daily	6 (12)	2 (4)	3 (5)

Table 2. Results of Study 1—blood pressure control.

Level of agreement with the following statements (0 “Strongly Disagree”-100 “Strongly Agree”)	SD 5, mean (95% CI)	SD 15, mean (95% CI)	SD 25, mean (95% CI)
This patient’s blood pressure is well controlled			
Mean BP[a,b]=130	79.37 (73.38-85.36)	56.82 (48.87-64.78)	29.18 (21.47-36.88)
Mean BP=145	39.57 (30.80-48.33)	23.45 (15.97-30.93)	21.8 (13.85-29.76)
Mean BP=160	22.75 (14.51-30.98)	16.61 (9.44-23.77)	17.57 (10.28-24.85)
This patient needs to change their medication			
Mean BP=130	29.75 (21.41-38.12)	47.39 (39.46-55.32)	71.27 (63.31-79.24)
Mean BP=145	58.16 (49.39-66.92)	76.53 (69.19-83.87)	85.04 (79.42-90.66)
Mean BP=160	86.47 (80.52-92.42)	85.25 (78.51-92.00)	83.43 (76.12-90.74)

[a]BP: blood pressure.

[b]All BPs provided in mm HG.

Study 2—Data Trends

In Study 2, we observed that patient ratings of hypertension control, need for medication changes, and 10-year risk of heart attack and stroke were significantly affected by trends visible in the BP data (see Tables 4 and 5) . Participants judged a positive data trend (depicting SBP values increasing over time) to be significantly less well controlled ($F_{1,49}$=107.53; P<.001; generalized η^2=0.45), in greater need of medication change ($F_{1,49}$=129.07; P<.001; generalized η^2=0.51), and at greater 10-year risk for a heart attack ($F_{1,49}$=112.79; P<.001; generalized η^2=0.47) and stroke ($F_{1,49}$=111.52; P<.001; generalized η^2=0.47) than a negative data trend (depicting SBP values decreasing over time) with the same SBP mean. Consistent with Study 1, participants’ recall of the percentage of SBP measurements out of range was largely inaccurate (Figure 3). Although the graphs depicting an increasing and decreasing trend had the same number of out-of-range measurements, participants recalled seeing a greater percentage of out-of-range values when the SBP was increasing over time than when the slope was decreasing ($F_{1,49}$=33.04; P<.001; generalized η^2=0.16).

Study 3—Outliers

Another way of examining the effect of variability in judgments about BP control is to consider the presence of extreme values (ie, outliers) independent of the overall measurement variance. In Study 3, we compared judgments about uniform data patterns (ie, no discernable outliers) with those with either 1 or 2 outliers. In addition, we systematically varied whether those outliers were above or below the mean BP of the depicted time period. The presence and number of outliers significantly affected judgments of hypertension control ($F_{4,\ 216}$=17.98; P<.001; generalized η^2=0.14), need for medication change ($F_{4,216}$=13.38; P<.001; generalized η^2=0.11), and perceived risk of heart attack ($F_{4,216}$=12.85; P<.001; generalized η^2=0.10), and stroke ($F_{4,216}$=13.81; P<.001; generalized η^2=0.11; Tables 6 and 7). When any extreme values were present, participants judged the patient to have hypertension that was significantly less well controlled, to be in greater need of medication change, and to be at greater 10-year risk for a heart attack or stroke than when the data had a more uniform distribution. Furthermore, 2 outliers (whether above or below the mean) were considered significantly more concerning than a single outlier.

Finally, recall for the percent of SBP values out of range was also significantly impacted by the presence and magnitude of

XSL•FO
RenderX

extreme values ($F_{4,216}$=5.54; P<.001; generalized η^2=0.04; Figure 4). When hypertension was controlled (mean 130), participants overestimated the percent of SBP values out of range with a uniform distribution and outlier(s) above the mean but more accurately recalled the percent of SBP values out of range when the outlier(s) were below the mean. In contrast, with borderline hypertension control (mean 145), recall for percent of SBP values out of range was fairly accurate with a uniform distribution and when there was only one outlier above the mean. However, participants consistently underestimated the percent of SBP values out of range when there was a single outlier below the mean or 2 outliers in either direction.

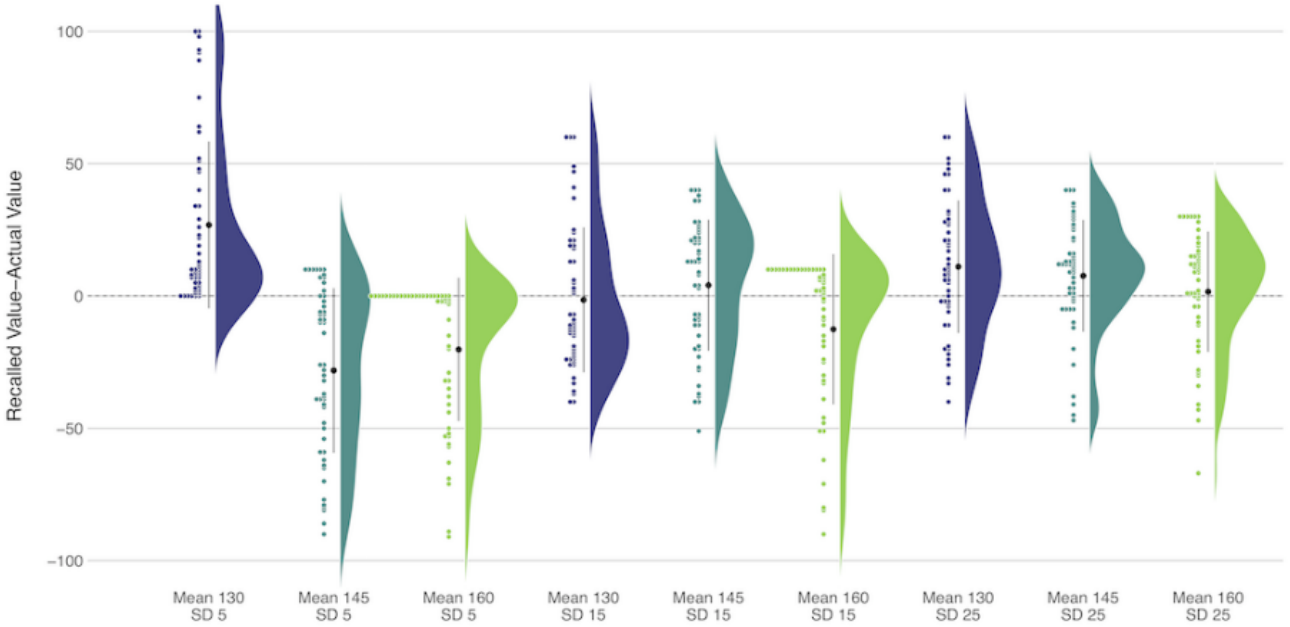
Table 3. Results of Study 1—risk perception and recall.

Perceived likelihood of the following events (1 "Extremely Unlikely"-10 "Extremely Likely")	SD 5, mean (95% CI)	SD 15, mean (95% CI)	SD 25, mean (95% CI)
Heart attack in the next 10 years			
Mean BP[a,b]=130	3.82 (3.26-4.39)	4.29 (3.67-4.92)	6.33 (5.73-6.94)
Mean BP=145	5.59 (4.90-6.27)	7.22 (6.57-7.86)	7.61 (7.09-8.13)
Mean BP=160	7.67 (7.05-8.28)	7.98 (7.32-8.64)	8.12 (7.65-8.59)
Stroke in the next 10 years			
Mean BP=130	3.76 (3.17-4.36)	4.45 (3.77-5.13)	6.63 (5.95-7.30)
Mean BP=145	5.76 (5.08-6.45)	7.43 (6.84-8.03)	8.00 (7.45-8.62)
Mean BP=160	8.04 (7.45-8.62)	8.14 (7.49-8.78)	8.35 (7.91-8.80)
% of systolic blood pressure points out of range (0%-100%)			
Mean BP=130 (Actual)	0	40	40
Mean BP=130 (Recalled)	26.84 (17.98-35.71)	38.53 (30.83-46.23)	51.06 (44.01-58.10)
Mean BP=145 (Actual)	90	60	60
Mean BP=145 (Recalled)	61.84 (53.10-70.59)	64.08 (57.12-71.04)	67.63 (61.69-73.57)
Mean BP=160 (Actual)	100	90	70
Mean BP=160 (Recalled)	79.78 (72.17-87.39)	77.42 (69.45-85.42)	71.63 (65.24-78.02)

[a]BP: blood pressure.

[b]All BPs provided in mm HG.

Figure 2. Error in the systolic blood pressure recall—Study 1, mean (large black dot) 95% CI.



XSL•FO
RenderX

Table 4. Results of Study 2—blood pressure control.

Level of agreement with the following statements (0 “Strongly Disagree”-100 “Strongly Agree”)	Increasing, mean (95% CI)	Decreasing, mean (95% CI)
This patient’s blood pressure is well controlled		
Mean BP[a,b]=130	30.88 (21.93-39.83)	85.72 (79.2-92.24)
Mean BP=145	20.40 (13.06-27.74)	73.80 (66.18-81.42)
Mean BP=160	11.26 (4.90-1762)	49.86 (40.45-59.27)
This patient needs to change their medication		
Mean BP=130	82.26 (75.85-88.67)	21.70 (13.47-29.93)
Mean BP=145	85.78 (79.37-92.19)	27.32 (19.41-35.23)
Mean BP=160	90.80 (84.15-97.45)	45.50 (35.72-55.28)

[a]BP: blood pressure.

[b]All BPs provided in mm HG.

Table 5. Results of Study 2—risk perception and recall.

Perceived likelihood of the following events (1 “Extremely Unlikely”-10 “Extremely Likely)	Increasing, mean (95% CI)	Decreasing, mean (95% CI)
Heart attack in the next 10 years		
Mean BP[a,b]=130	7.18 (6.58-7.78)	3.02 (2.42-3.62)
Mean BP=145	7.96 (7.39-8.53)	3.66 (3.01-4.31)
Mean BP=160	9.00 (8.56-9.44)	5.60 (4.89-6.31)
Stroke in the next 10 years		
Mean BP=130	7.18 (6.55-7.81)	2.96 (2.34-3.58)
Mean BP=145	7.98 (7.38-8.58)	3.60 (2.98-4.22)
Mean BP=160	9.08 (8.60-9.56)	5.56 (4.80-6.32)
% of systolic blood pressure points out of range (0%-100%)		
Mean BP=130 (Actual)>	40	40
Mean BP=130 (Estimate)	56.48 (49.30-63.66)	29.4 (21.45-37.35)
Mean BP=145 (Actual)	60	60
Mean BP=145 (Estimate)	63.86 (56.89-70.83)	44.82 (37.15-52.49)
Mean BP=160 (Actual)	90	90
Mean BP=160 (Estimate)	84.66 (78.94-90.38)	61.78 (52.23-71.33)

[a]BP: blood pressure.

[b]All BPs provided in mm HG.

XSL•FO
RenderX

Figure 3. Error in the systolic blood pressure recall—Study 2, mean (large black dot) 95% CI.

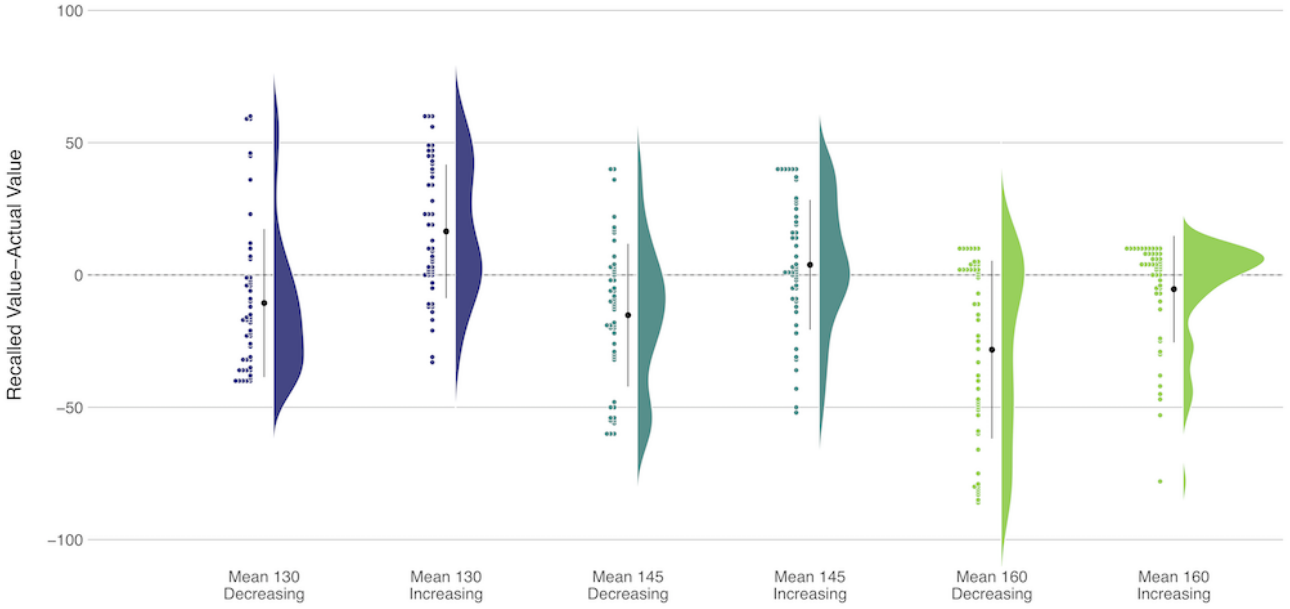


Table 6. Results of Study 3— blood pressure control.

Level of agreement with the following statements (0 “Strongly Disagree”-100 “Strongly Agree”)	Uniform, mean (95% CI)	1 Up, mean (95% CI)	1 Down, mean (95% CI)	2 Up, mean (95% CI)	2 Down, mean (95% CI)
This patient’s blood pressure is well controlled					
Mean BP[a,b] =130	89.22 (84.12-94.32)	68.91 (61.41-76.41)	67.71 (59.49-75.93)	53.07 (44.47-61.68)	47.02 (38.60-55.44)
Mean BP=145	37.04 (28.15-45.92)	40.65 (32.40-48.91)	28.49 (20.33-36.65)	39.27 (30.43-48.12)	19.09 (11.84-26.34)
This patient needs to change their medication					
Mean BP=130	20.87 (12.67-29.07)	33.18 (24.89-41.48)	32.47 (24.61-40.33)	49.87 (40.89-58.85)	53.98 (46.44-61.52)
Mean BP=145	70.93 (62.51-79.35)	65.04 (56.26-73.81)	76.93 (69.52-84.34)	73.25 (65.42-81.09)	84.51 (77.91-91.11)

[a]BP: blood pressure.

[b]All BPs provided in mm HG.

XSL•FO
RenderX

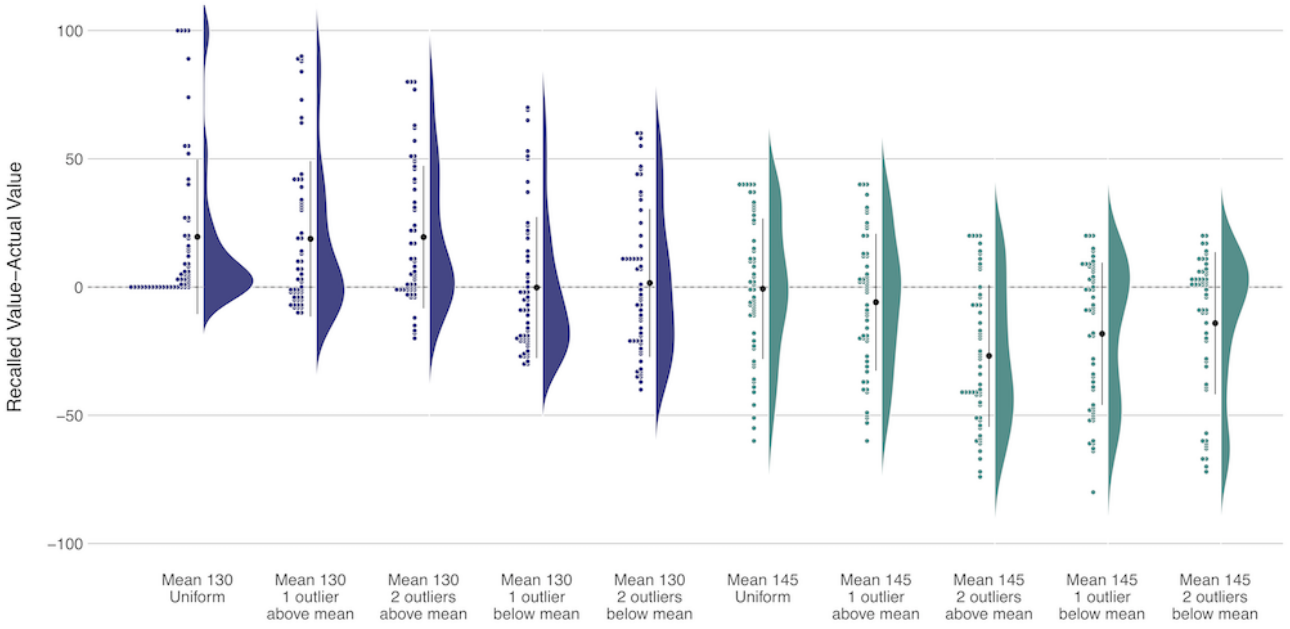
Table 7. Results of Study 3—risk perception and recall.

Perceived likelihood of the following events (1 “Extremely Unlikely”-10 “Extremely Likely”)	Uniform, mean (95% CI)	1 Up, mean (95% CI)	1 Down, mean (95% CI)	2 Up, mean (95% CI)	2 Down, mean (95% CI)
Heart attack in the next 10 years					
Mean BP[a,b]=130	2.65 (2.12-3.19)	4.02 (3.32-4.72)	3.87 (3.28-4.46)	5.29 (4.57-6.01)	4.85 (4.22-5.49)
Mean BP=145	6.42 (5.75-7.08)	6.24 (5.63-6.84)	7.07 (6.48-7.66)	6.82 (6.10-7.54)	7.69 (7.11-8.27)
Stroke in the next 10 years					
Mean BP=130	2.58 (2.08-3.08)	4.14 (3.40-4.88)	3.90 (3.24-4.56)	5.46 (4.69-6.23)	4.76 (4.06-5.46)
Mean BP=145	6.55 (5.87-7.22)	6.42 (5.83-7.01)	7.35 (6.73-7.97)	6.95 (6.22-7.67)	7.98 (7.39-8.57)
% of systolic blood pressure points out of range (0%-100%)					
Mean BP=130 (Actual)	0	10	30	20	40
Mean BP=130 (Estimate)	19.58 (11.44-27.72)	28.80 (20.61-36.99)	29.27 (21.89-36.65)	39.47 (31.96-46.98)	41.58 (33.79-49.38)
Mean BP=145 (Actual)	60	60	80	80	80
Mean BP=145 (Estimate)	59.33 (51.92-66.74)	54.09 (46.88-61.31)	61.78 (54.28-69.29)	53.18 (45.69-60.67)	65.91 (58.41-73.41)

[a]BP: blood pressure.

[b]All BPs provided in mm HG.

Figure 4. Error in the systolic blood pressure recall—Study 3, mean (large black dot) 95% CI.



Discussion

Principal Findings

To inform the development of a physician-patient shared data visualization tool for BP values, we conducted 3 vignette-based Web studies to understand better how patients interpret the variability in visualizations of BP data. In all studies, we observed that patients with hypertension consistently judged variations in BP data as meaningful indicators of hypertension control. In Study 1, increased variability in SBP data was associated with increasingly negative judgments about hypertension control, even when the mean SBP remained constant and within goal range. In Study 2, we demonstrated that the pattern of change in SBP values over time also significantly influenced judgments about hypertension control. While graphs depicting an increase in the SBP over time were appropriately judged to be of greater concern than graphs showing a decrease in the SBP over time, recall for the number of out-of-range BP values was inaccurate. When the SBP values increased over time, patients remembered more out-of-range values than when the slope decreased, even though the number of out-of-range values did not differ in our study. Finally, the presence of outliers (Study 3) also significantly impacted patients’ judgments about hypertension control, the need for a medication change, and the risks associated with uncontrolled hypertension. Observing even 1 or 2 outliers in BP data leads

XSL•FO
RenderX

patients to inappropriately determine that hypertension is not well controlled.

The use of data visualization in electronic medical records has the potential to transform clinical encounters. While the technology to develop these tools is available, little is known about how these data displays will influence patients' judgments about hypertension control and subsequent decisions regarding treatment. These 3 studies have demonstrated that judgments about hypertension control are strongly and inappropriately influenced by the presence of outliers and variability in the data. Outliers and variability mask mean BP values and the presence of data trends—important predictors of BP-related health outcomes (eg, heart attack or stroke) [25]. These findings are important for the development of interventions to promote shared decision making in primary care, which must direct attention to clinically meaningful information, that is, mean BP and trends rather than variability in BP or outliers [25]. It should be stressed that these conclusions about the potential benefits of visualization techniques that minimize the impact of outliers and variability would apply only to measurements obtained from well-calibrated BP measurement devices. These recommendations should not be applied to devices where variability is derived from inaccurate measurement.

Limitations

There are several limitations to these studies that potentially constrain their generalizability. One limitation is the use of internet patient samples. When patients are making decisions about treatment for hypertension, they are typically made in conjunction with their physician during a clinic visit. In addition, while our sample of patients is more demographically diverse than typical internet samples, it is not representative of the population of patients with hypertension. Furthermore, we focused only on patients in these studies; therefore, future work should examine the effect of data visualization on physicians' judgments about hypertension control, as well as the effect on shared patient-physician decisions. It is possible that physicians will perform similarly to patients because we are examining judgments that stem from common perceptual and cognitive processes [30,31]. On the other hand, physicians may have greater knowledge about the relative importance of the BP mean and variability than patients, which could alter their judgments.

Conclusions

Health information technologies provide an opportunity for patients to become more engaged in decision making about hypertension control. We are endeavoring to design a data visualization tool for BP that can be jointly used by physicians and patients in this decision-making process. This tool aims to make the limited time shared in the exam room more efficient and effective. Defining how data elements, such as trends, variability, and outliers, support or detract from an understanding of the data will aid in the design of data visualizations that highlight meaningful characteristics of the data; this may, in turn, result in shared decisions that are better informed. Areas for future study include understanding how these parameters influence physician judgments about hypertension control and how information acquisition from data visualizations is affected by numeracy, health literacy, and graph literacy of patients and physicians.

Acknowledgments

This project was supported by the Agency for Healthcare Research and Quality (grant number R01HS023328). 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 data from this manuscript were previously presented at the 2016 and 2017 Annual Meetings of the Society for Medical Decision Making.

The authors would like to thank Joey Burant (@jbburant) and John Burn-Murdock (@jburnmurdock) for inspiring the "jellyfish" plots used in this manuscript. Joey posted R code via Twitter that was helpful in the construction of these plots, and John coined the plot name.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study 1 Materials.

[PDF File (Adobe PDF File), 805KB - jmir_v21i3e11366_app1.pdf]

Multimedia Appendix 2

Study 2 Materials.

[PDF File (Adobe PDF File), 516KB - jmir_v21i3e11366_app2.pdf]

XSL•FO
RenderX

Multimedia Appendix 3

Study 3 Materials.

[PDF File (Adobe PDF File), 888KB - jmir_v21i3e11366_app3.pdf]

Multimedia Appendix 4

Data for Studies 1-3 disaggregated by gender.

[PDF File (Adobe PDF File), 136KB - jmir_v21i3e11366_app4.pdf]

References

1. Go A, Mozaffarian D, Roger V, Benjamin EJ, Berry JD, Borden WB, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2013 update: a report from the American Heart Association. Circulation 2013 Jan 01;127(1):e6-e245 [FREE Full text] [doi: 10.1161/CIR.0b013e31828124ad] [Medline: 23239837]
2. Roger V, Go A, Lloyd-Jones D. Heart disease and stroke statistics--. Dallas, TX: American Heart Association; 2012:a.
3. Merai R, Siegel C, Rakotz M, Basch P, Wright J, Wong B, DHSc, et al. CDC Grand Rounds: A Public Health Approach to Detect and Control Hypertension. MMWR Morb Mortal Wkly Rep 2016 Nov 18;65(45):1261-1264 [FREE Full text] [doi: 10.15585/mmwr.mm6545a3] [Medline: 27855138]
4. Yoon SSS, Carroll MD, Fryar CD. Hypertension Prevalence and Control Among Adults: United States, 2011-2014. NCHS Data Brief 2015 Nov(220):1-8 [FREE Full text] [Medline: 26633197]
5. Kerr E, Zikmund-Fisher B, Klamerus M, Subramanian U, Hogan M, Hofer TP. The role of clinical uncertainty in treatment decisions for diabetic patients with uncontrolled blood pressure. Ann Intern Med 2008 May 20;148(10):717-727. [Medline: 18490685]
6. Lebeau J, Cadwallader J, Aubin-Auger I, Mercier A, Pasquet T, Rusch E, et al. The concept and definition of therapeutic inertia in hypertension in primary care: a qualitative systematic review. BMC Fam Pract 2014 Jul 02;15:130 [FREE Full text] [doi: 10.1186/1471-2296-15-130] [Medline: 24989986]
7. Crowley MJ, Smith VA, Olsen MK, Danus S, Oddone EZ, Bosworth HB, et al. Treatment intensification in a hypertension telemanagement trial: clinical inertia or good clinical judgment? Hypertension 2011 Oct;58(4):552-558 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.174367] [Medline: 21844490]
8. Kolata G. The New York Times. 2015. Blood Pressure, a Reading With a Habit of Straying URL: https://www.nytimes.com/2015/12/01/health/blood-pressure-a-reading-with-a-habit-of-straying.html
9. Wegier P, Belden J, Canfield S. Designing data visualizations to support shared decision making about blood pressure. 2016 Oct 23 Presented at: The Meeting of the Society for Medical Decision Making; 2016; Vancouver, BC p. 38.
10. Zikmund-Fisher BJ, Exe NL, Witteman HO. Numeracy and literacy independently predict patients' ability to identify out-of-range test results. J Med Internet Res 2014 Aug 08;16(8):e187 [FREE Full text] [doi: 10.2196/jmir.3241] [Medline: 25135688]
11. Zikmund-Fisher B, Witteman H, Fuhrel-Forbis A, Exe N, Kahn V, Dickson M. Animated graphics for comparing two risks: a cautionary tale. J Med Internet Res 2012 Jul 25;14(4):e106 [FREE Full text] [doi: 10.2196/jmir.2030] [Medline: 22832208]
12. Zikmund-Fisher BJ, Dickson M, Witteman HO. Cool but counterproductive: interactive, Web-based risk communications can backfire. J Med Internet Res 2011 Aug 25;13(3):e60 [FREE Full text] [doi: 10.2196/jmir.1665] [Medline: 21868349]
13. Zikmund-Fisher B, Scherer A, Witteman H, Solomon J, Exe N, Fagerlin A. Effect of Harm Anchors in Visual Displays of Test Results on Patient Perceptions of Urgency About Near-Normal Values: Experimental Study. J Med Internet Res 2018 Mar 26;20(3):e98 [FREE Full text] [doi: 10.2196/jmir.8889] [Medline: 29581088]
14. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
15. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). JAMA 2014 Feb 05;311(5):507-520. [doi: 10.1001/jama.2013.284427] [Medline: 24352797]
16. Juhanoja EP, Niiranen TJ, Johansson JK, Puukka PJ, Jula AM. Agreement between ambulatory, home, and office blood pressure variability. J Hypertens 2016 Jan;34(1):61-67. [doi: 10.1097/HJH.0000000000000772] [Medline: 26630214]
17. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making 2007 Sep;27(5):672-680. [doi: 10.1177/0272989X07304449] [Medline: 17641137]
18. Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. BMC Fam Pract 2006 Mar 24;7:21 [FREE Full text] [doi: 10.1186/1471-2296-7-21] [Medline: 16563164]
19. Qualtrics. Utah, USA: Provo; 2018. URL: https://www.qualtrics.com/ [accessed 2019-02-21] [WebCite Cache ID 76LlJQc6G]

XSL•FO
RenderX

20. Centers for Disease ControlPrevention (CDC). Vital signs: awareness and treatment of uncontrolled hypertension among adults--United States, 2003-2010. MMWR Morb Mortal Wkly Rep 2012 Sep 07;61:703-709 [FREE Full text] [Medline: 22951452]
21. Diaz VA, Mainous AG, Koopman RJ, Geesey ME. Undiagnosed obesity: implications for undiagnosed hypertension, diabetes, and hypercholesterolemia. Fam Med 2004 Oct;36(9):639-644 [FREE Full text] [Medline: 15467941]
22. 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]
23. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [Medline: 17695343]
24. Heidari S, Babor TF, De Castro P, Tort S, Curno M. [Sex and gender equity in research: rationale for the SAGER guidelines and recommended use]. Gac Sanit 2018 May 03:--- [FREE Full text] [doi: 10.1016/j.gaceta.2018.04.003] [Medline: 29731296]
25. Hansen TW, Thijs L, Li Y, Boggia J, Kikuya M, Björklund-Bodegård K, International Database on Ambulatory Blood Pressure in Relation to Cardiovascular Outcomes Investigators. Prognostic value of reading-to-reading blood pressure variability over 24 hours in 8938 subjects from 11 populations. Hypertension 2010 Apr;55(4):1049-1057 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.109.140798] [Medline: 20212270]
26. Stergiou GS, Ntineri A, Kollias A, Ohkubo T, Imai Y, Parati G. Blood pressure variability assessed by home measurements: a systematic review. Hypertens Res 2014 Jun;37(6):565-572. [doi: 10.1038/hr.2014.2] [Medline: 24553366]
27. Asayama K, Kikuya M, Schutte R, Thijs L, Hosaka M, Satoh M, et al. Home blood pressure variability as cardiovascular risk factor in the population of Ohasama. Hypertension 2013 Jan;61(1):61-69 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.00138] [Medline: 23172933]
28. Rothwell PM, Howard SC, Dolan E, O'Brien E, Dobson JE, Dahlöf B, ASCOT-BPLAMRC Trial Investigators. Effects of beta blockers and calcium-channel blockers on within-individual variability in blood pressure and risk of stroke. Lancet Neurol 2010 May;9(5):469-480. [doi: 10.1016/S1474-4422(10)70066-1] [Medline: 20227347]
29. Diaz KM, Tanner RM, Falzon L, Levitan EB, Reynolds K, Shimbo D, et al. Visit-to-visit variability of blood pressure and cardiovascular disease and all-cause mortality: a systematic review and meta-analysis. Hypertension 2014 Nov;64(5):965-982 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.114.03903] [Medline: 25069669]
30. Belden J, Lowrance N, Patel J. 2014. Inspired EHRs: Designing for Clinicians URL: http://inspiredehrs.org/designing-for-clinicians/e-prescribing.php
31. Ware C. Visual Queries: The Foundation of Visual Thinking. Knowledge and Information Visualization. Berlin Heidelberg: Springer; 2005.

Abbreviations

BP: blood pressure
EHR: electronic health record
SILS: Single-Item Literacy Screener for health literacy
SBP: systolic blood pressure
SNS: Subjective Numeracy Scale

Edited by G Eysenbach; submitted 09.07.18; peer-reviewed by H Witteman, J Goris; comments to author 22.08.18; revised version received 27.09.18; accepted 29.09.18; published 26.03.19.

Please cite as:
Shaffer VA, Wegier P, Valentine KD, Belden JL, Canfield SM, Patil SJ, Popescu M, Steege LM, Jain A, Koopman RJ
Patient Judgments About Hypertension Control: The Role of Variability, Trends, and Outliers in Visualized Blood Pressure Data
J Med Internet Res 2019;21(3):e11366
URL: https://www.jmir.org/2019/3/e11366/
doi:10.2196/11366
PMID:30912759

©Victoria Anne Shaffer, Pete Wegier, KD Valentine, Jeffery L Belden, Shannon M Canfield, Sonal J Patil, Mihail Popescu, Linsey M Steege, Akshay Jain, Richelle J Koopman. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete

XSL•FO
RenderX

bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Role of Frontline Leaders in Building Health Professional Support for a New Patient Portal: Survey Study

Sari Kujala[1], PhD; Iiris Hörhammer[2], DSc; Tarja Heponiemi[3], PhD; Kim Josefsson[3], PhD

[1]Department of Computer Science, Aalto University, Espoo, Finland

[2]Department of Industrial Engineering and Management, Aalto University, Espoo, Finland

[3]National Institute for Health and Welfare, Helsinki, Finland

Corresponding Author:
Sari Kujala, PhD
Department of Computer Science
Aalto University
PO Box 15400
Espoo,
Finland
Phone: 358 503862768
Email: sari.kujala@aalto.fi

Abstract

Background: Effective leadership and change management are thought to contribute to the successful implementation of health information technology innovations. However, limited attention has been paid to the role of frontline leaders in building health professional support for new technical innovations.

Objective: First, we examined whether frontline leaders' positive expectations of a patient portal and perceptions of its implementation were associated with their support for the portal. Second, we explored whether leaders' positive perceptions influenced the same unit's health professional support for the portal.

Methods: Data were collected through an online survey of 2067 health professionals and 401 frontline leaders working in 44 units from 14 health organizations in Finland. The participating organizations run a joint self-care and digital value services project developing a new patient portal for self-management. The survey was conducted before the piloting and implementation of the patient portal.

Results: The frontline leaders' perception of vision clarity had the strongest association with their own support for the portal (ß=.40, *P*<.001). Results also showed an association between leaders' view of organizational readiness and their support (ß=.15, *P*=.04). The leaders' positive perceptions of the quality of informing about the patient portal was associated with both leaders' own (ß=.16, *P*=.02) and subordinate health professionals' support for the portal (ß=.08, *P*<.001). Furthermore, professional participation in the planning of the portal was positively associated with their support (ß=.57, *P*<.001).

Conclusions: Findings suggest that assuring good informing, communicating a clear vision to frontline leaders, and acknowledging organizational readiness for change can increase health professional support for electronic health (eHealth) services in the pre-implementation phase. Results highlight the role of frontline leaders in engaging professionals in the planning and implementation of eHealth services and in building health professionals' positive attitudes toward the implementation of eHealth services.

(J Med Internet Res 2019;21(3):e11413) doi:10.2196/11413

KEYWORDS

patient portal; implementation; expectations; organizational readiness; leadership

Introduction

Effective leadership and change management are seen as particularly important to the successful implementation of health information technology innovations [1-5]. The leaders' role in implementation projects can involve mitigating possible risks [6], but leaders can also move proactively to ensure the success of the implementation. For example, Ingebrigtsen et al [7] identified seven categories of leadership behaviors associated with successful outcomes: (1) communicating clear visions and

XSL•FO
RenderX

goals, (2) leadership support, (3) establishing a governance structure and coordinating teamwork, (4) arranging training, (5) identifying and appointing champions, (6) addressing work process change, and (7) follow-up. Also, ensuring the ongoing involvement of key stakeholders as well as reserving extra time and reducing workload during the initial implementation are often identified as tasks for leaders [2].

Successful implementation does not only depend on the strategic level of management, but the operational and frontline levels also have their own roles and tasks [7,8]. Implementations often happen in complex organizations, and changes in these environments require a clear vision at all levels [7]. The high-level strategies and objectives need to be integrated with the underlying processes that operationalize the objectives [9]. After all, health care professionals and the attitudes they present are key to preventing resistance and ensuring the active use of new innovations [10,11].

In theory, effective leadership should lead to organizational readiness for change during the pre-implementation stage [12]: those involved are individually and collectively primed, motivated, and technically capable of executing the change [13]. In particular, leaders and change agents need to communicate the benefits of the change. For example, Paré et al [12] found that nurses' perceptions of vision clarity (why a change is needed) and change appropriateness (that the proposed change is the correct one) had a significant positive influence on organizational readiness in two cross-sectional surveys.

The pre-implementation phase is critical, as the successes and risks tend to accumulate over time [14,15]. Based on their systematic literature review, Ingebrigtsen et al [7] suggest that good communication should spread to all organizational leadership levels before implementation. However, they identified that there is a lack of evidence regarding the role of clinical leaders within different organizational levels and implementation phases.

In this paper, we focus on clinical leaders who are responsible for leadership within an organization that delivers care as defined by Ingebrigtsen et al [7]. In particular, we are interested in frontline leaders such as physicians and nurses working and supervising health professionals in frontline units (eg, wards or care units). These leaders seem to play a critical role in conveying information and motivating health professionals in the early phases of implementation. Although leadership support has been identified as important, it is not known how leaders influence the attitudes of health professionals.

This study aimed to examine whether the health care leaders' positive expectations of a patient portal and perceptions of its implementation are associated with their own and the subordinate professional support for the portal in the pre-implementation phase. Specifically, leader support for the portal and their perceptions of the portal benefits, the readiness and implementation practices of their units, and the quality of informing about the portal were studied.

Methods

A survey study was conducted to capture health professional and leader expectations about a national patient portal for self-management that was developed by a self-care and digital value services (*omahoito ja digitaaliset arvopalvelut*, or ODA) project. The data were collected before piloting and implementing the patient portal.

Study Setting

At the time of the study in spring 2017, the role of information and communication technology (ICT) was widening in Finnish public health care. The objective of the national eHealth (electronic health) and eSocial strategy 2020 is to support the active role of citizens in promoting their own well-being by improving information management and implementing self-management and online services [16]. The aim of the strategy is to support the prevention of health problems, self-assessment of the need for services, and independent coping.

To build online services for citizens, the government has funded an ODA project. The ODA project is run as a joint one comprising 14 municipalities and hospital districts in Finland—including the largest cities of Helsinki, Espoo, Tampere, Turku, and Oulu—totaling 1,766,334 inhabitants in 2014 (32.3% of the Finnish population). The ultimate aim of the ODA project is to provide a national patient portal for self-management and self-service in 2018. In addition to the technical development, particular emphasis has been placed on changing the operational processes by using a participatory approach.

The three main services of the patient portal are patient self-assessment with online well-being coaching programs, a symptom checker for patients to support care navigation and provide information on medical conditions, and a personalized shared care plan tethered to the electronic health records. Using the portal would provide health professionals with additional structured information from patient self-assessment and well-being coaching, symptom checker, and shared care plan. At the time of this study, it was unclear how health professional workflow would be affected. However, the aim was to handle routine tasks automatically, provide patients self-managements tools, and provide more time to personal care.

The ODA project has trained groups of 10 health professionals representing each city and health care region in 2016 and 2017 in a lean process development approach. These groups have planned the operational processes to be supported by the technical solutions in individual services. At the time of the study, none of the organizations had yet started to pilot the services. The first pilot projects started in June 2017, and the whole patient portal entity was planned to start in autumn 2018.

Questionnaire

There were two versions of the questionnaire, one for professionals and one for their supervisors and leaders. Both surveys collected data on respondent support for the patient portal, background information, respondent's health organization (city or hospital district), work unit (hospital, health center, etc) and other variables such as efficiency improvements, benefits

XSL•FO
RenderX

for patients, the organization's implementation practices, and the quality of informing. From the professional questionnaire, only the data on their background information and support for a patient portal were included in this study.

The focus of this study was on the leader questionnaire (see Multimedia Appendix 1). The questionnaire was targeted to official supervisors and leaders. Informal leaders such as clinical champions were not included. Five previously validated Information Systems Expectations and Experiences survey items were applied for measuring support for a patient portal [17-19]. An essential part of leader support for a patient portal is that they also support change in work processes. The 5-item organizational readiness and 3-item vision clarity scales [12] were also applied.

In addition, leaders were asked to rate whether they agreed with the positive influences that the new patient portal was planned to have. There were 4 statements about the expected efficiency improvements and 6 statements about the expected benefits for patients. In addition, we asked leaders to evaluate personnel readiness (ie, how willing and able the personnel are to adopt the new portal) using 5 items. They were also asked about the organization's implementation practices by rating whether they agreed that 9 good implementation practices would take place in the future implementation of the new patient portal. The 9 good implementation practices were identified from the literature [2,6,14,20-23] (see Multimedia Appendix 1). Thus, the content validity of these survey items was established through a literature review [24].

The quality of informing including communicating the goals of the portal was ascertained with 3 questions, and another 3 questions focused on participation in the design of the services. All items were measured on a 5-point Likert-type scale from strongly disagree (1) to strongly agree (5) and included option 6 (I don't know), which in the statistical analyses was handled as a missing value. User participation was a categorical variable formulated based on 3 questions, showing whether a respondent had participated in the planning of the services of the patient portal.

Before the data were gathered, the content of the questionnaire was reviewed by a group of 3 researchers in the field and the ODA management team. In addition, as recommended [25-27], we tested the reliability of the questionnaire with 4 leaders who completed the questionnaire and talked aloud at the same time about how they understood the questions. Based on the iterative pilot testing, the questionnaire was revised by clarifying wording and slightly modifying some items.

Data Collection

The data were gathered in the spring of 2017 by using a Web-based questionnaire tool. The project managers of each of the 14 organizations participating in the ODA project sent survey invitations and reminders via email to all health professionals. Based on estimations of the project managers, respondents represented 6.1% of health professionals working in the organizations. As the project managers did not know the number of leaders or how many of the professionals received the survey information, we were not able to calculate the exact response rate. The first page of the questionnaire included a screenshot of one ODA page to illustrate the professional view of the patient portal. To encourage participation, 10 pairs of movie tickets were raffled off among the respondents. The study protocol was reviewed and approved by the Ethical Review Board of Aalto University.

Analysis

Because the data did not allow linking health professionals with their respective leaders, leader variables were averaged over their organizational unit and assigned to the unit's professionals. Two small health organizations with units including fewer than two leaders were excluded from the analysis. Units were defined by health organization (city or hospital district) and work unit (hospital, health center, etc). To restrict the analysis to leaders who supervise professionals working with patients, administrative units were excluded from the analyses. The minimum number of professionals in the included units was four.

Descriptive statistics and reliability analyses were performed and mean sum scores were computed for all study variables (see Multimedia Appendix 2). Cronbach alpha scores were all well above .84 (for the leader questioner) and .80 (for the professional questionnaire), indicating good internal consistency [28]. A multicollinearity analysis of the study variables was performed. The variance inflation factors for independent variables in the leader support regression were all below 2.8 and for independent variables in the professional support regression below 5.6, indicating that multicollinearity was not a concern in this study [29].

Associations of leader support with other leader variables were tested with multiple linear regression that included leader support for the patient portal as a dependent variable (DV) and leader age, gender, vision clarity, expected efficiency improvements, expected benefits for patients, personnel readiness, organizational readiness, quality of informing, implementation practices, and participation in the planning of the new patient portal as independent variables (IV).

In the analysis of associations between leader variables and a professional's support for services, multilevel linear regression was used. This enabled us to control for the natural clustering of employees in organizational units (ie, the dependency between observations of professionals from the same unit). Individual-level professional support was used as the DV; unit-level leader variables and individual-level participation in the planning of the patient portal were used as IVs; and the age of professionals was used as a control variable. Robust estimators for standard errors were used. We first tested the associations of the leader predictors separately using univariate regression. Second, to test the relative contribution of each of the IVs to the total variance of DV explained, we used multiple regression analysis. Predictors that were significant in univariate regression were included in the multiple regression analysis. Backward elimination was used to assess the independence of the predictors. All statistical analyses were performed using Stata version 15.0 (StataCorp LLC).

XSL•FO
RenderX

Results

Respondents

Responses from 44 organizational units (eg, primary care in Helsinki) were included in the analysis. All together, 401 leader and 2067 health professional respondents working in these organizational units responded to the questionnaire. The respondents represented 12 health organizations and 6 work units (primary care health center, hospital, psychiatric outpatient clinic, elementary school health care, emergency care, dental care, and other).

Tables 1 and 2 show the background information of the respondents. The majority of leaders (324/401, 80.8%) and professionals (776/2067, 85.92%) were women, and the mean ages of the leaders and professionals were 51.6 and 45.1 years, respectively. The leader respondents were nurses, physicians, counselors, and others whose titles did not reveal profession. A greater proportion of leaders (63/401, 15.7%) than professionals (78/2067, 3.77%) had participated in the planning of the patient portal services.

Table 1. Leader background information (n=401).

Characteristic	Value
Gender, n (%)	
Male	64 (16.0)
Female	324 (80.8)
Not reported	13 (3.2)
Age in years, mean (SD)	51.6 (7.9)
Profession, n (%)	
Nurse leader	177 (44.1)
Physician leader	94 (23.4)
Counseling leader	74 (18.5)
Other	56 (14.0)
Participated in the planning, n (%)	63 (15.7)
Years of work experience in similar tasks, mean (SD)	10.2 (8.5)

Table 2. Professional background information (n=2067).

Characteristics	Value
Gender, n (%)	
Male	228 (11.0)
Female	1776 (85.9)
Not reported	63 (3.1)
Age in years, mean (SD)	45.1 (11.1)
Profession, n (%)	
Hospital nurse	651 (31.5)
Practical nurse	405 (19.6)
Doctor/dentist	189 (9.1)
Public health nurse	167 (8.1)
Physio and other therapists	124 (6.0)
Dental nurse	47 (2.3)
Social worker	42 (2.0)
Midwife	28 (1.4)
Administrator	27 (1.3)
Other	387 (18.7)
Participated in the planning, n (%)	78 (3.7)
Years of work experience in the field, mean (SD)	16.2 (10.8)

XSL•FO
RenderX

Factors Associated With Leader Support for a New Patient Portal

Multimedia Appendix 2 presents the mean and standard deviation of the study variables. Table 3 presents the results of the multiple linear regression analysis, examining the associations of independent leader variables with the leader support for the patient portal. All IVs except for gender and age were significantly associated with leader support in the univariate analyses (P<.001; Multimedia Appendix 3), and therefore they were all included as independent predictors in the multiple regression. Age was included as a control variable as previous studies have suggested associations of age [30] and gender [31] with the use of eHealth technologies.

Results show that the leaders' clear vision of the future patient portal was moderately associated with their support for the portal (ß=.40, P<.001). Moreover, good quality of information (ß=.16, P=.02) and perceived organizational readiness (ß=.15, P=.04) were modestly positively associated with leader support for the patient portal. In the multiple regression model, expected efficiency improvements, benefits for patients, personnel readiness, and implementation practices did not have a significant independent association with leader support. Overall, the model explained 43% of the variance in leader support for the portal.

Associations of Unit-Level Leader Factors With Professional Support for the Patient Portal

To test how the perceptions of the leaders were associated with professional support for the patient portal in the pre-implementation phase, we used multilevel modeling. First, in the null model, we tested the effect of clustering on professional support. As indicated in Table 4, the intraclass correlation was 0, indicating that variation in professional support occurred at the individual level rather than at the unit level. Model A includes all potential predictors of professional support that were significant (P<.001) in the univariate analyses (Multimedia Appendix 4).

In Model A, only the individual-level variables professional participation in the planning of the services (ß=.57, P<.001) and professional age (ß=.06, P=.04) showed statistically significant association with professional support. To further test which independent variables could best explain professional support, we used backward elimination. Model B shows that the individual-level variable professional participation in the planning (ß=.57, P<.001), unit-level leader view of informing (ß=.08, P<.001), and individual-level professional age (ß=–.06, P=.05) alone explain 17% of the variation in professional support with rather modest magnitudes of association. The analysis therefore suggests that there is an association between leader perception of the quality of informing provided and disseminated about the new patient portal and professional support for the portal. The quality of informing variable included aspects of how well the leaders received information and how well the leaders informed their own subordinates. Thus, the results suggest that if the leaders had received information about the portal, they also informed their subordinates. In addition, professionals' own participation in the planning of the services is associated with their support.

Table 3. Multiple linear regression analysis for leader support showing the associations of independent leader variables; R^2=.43.

Variables	ß	Standard error	P value
Vision clarity	.40	.06	<.001
Efficiency improvements	.04	.07	.58
Benefits for patients	.03	.07	.65
Personnel readiness	.01	.07	.90
Organizational readiness	.15	.07	.04
Quality of informing	.16	.07	.02
Implementation practices	–.08	.07	.25
Participation in the planning (category reference: no participation)	.07	.13	.63
Age	–.05	.04	.30

XSL•FO
RenderX

Table 4. Multilevel model of the association of unit-level leader views with professional support (N=1532; 44 units). Continuous variables were used as continuous standardized variables.

Variable	Null model	Model A	Model B
Intercept, regression coefficient (robust standard error)	–0.00 (.03)	–0.03 (.02)	–0.02 (.03)
Unit-level leader variables, regression coefficient (robust standard error)			
Support for services	—	0.01 (.04)	—
Vision clarity	—	–0.04 (.04)	—
Expected efficiency improvements	—	0.00 (.04)	—
Expected benefits for patients	—	0.05 (.06)	—
Personnel readiness	—	0.01 (.05)	—
Organizational readiness	—	0.04 (.03)	—
Quality of informing	—	0.11 (.06)	0.08 (.02)[b]
Implementation practices	—	–0.08 (.05)	—
Individual-level variable, regression coefficient (robust standard error)			
Professional's age	—	–0.06 (.03)[a]	–0.06 (.03)
Professional's participation (category reference: no participation)	—	0.57 (.10)[b]	0.57 (.10)[b]
R^2	0.00	0.24	0.17
Rho (intraclass correlation)	0.00	0.00	0.00

[a]$P<.05$.

[b]$P<.001$.

Discussion

Principal Findings and Comparison With Prior Work

Effective leadership has been identified as important in the successful implementation of health information technology innovations [1-5]. This survey study aimed at clarifying the role of frontline leaders especially in motivating health professionals in the pre-implementation phase of a patient portal. Frontline leaders' own support for the new patient portal had strongest association with their perception of vision clarity, information shared in the organization about the portal, and organizational readiness. These factors explained close to half of the variation in leader support, which suggests that special attention should be paid to communicating the vision clearly among the leaders, supporting the communication concerning the new eHealth services in all employee groups, and choosing the timing of implementation wisely to support the organizational readiness. Frontline leaders' positive perception of quality of information about the patient portal in the organization was also associated with professional support for the portal, which further accentuates the significance of good communication. The multilevel analysis further showed that professionals' own participation in the planning of the patient portal had positive association with their support for the patient portal. This finding implies that engaging professionals early on in the development can have a notable effect on implementation success.

Overall, the quality of informing was evaluated to be rather low. In the early pre-implementation phase, the health organizations had not yet efficiently informed the personnel that a new patient portal will be implemented. However, informing is important especially in the initiation phase because potential users need to be aware of new services in order to adopt them [20]. Moreover, reaching hundreds or thousands of health professionals takes time. Realistic positive expectations are also known to support later adoption of new services [32].

The findings suggest that good communication and being well informed are particularly important in building frontline leader support for a new eHealth service in the pre-implementation phase. If top management communicates a clear vision and sufficiently conveys information about the new service, frontline leaders are able to inform their subordinate health professionals. During this early implementation phase, the provision of sufficient information about the future patient portal and professional participation in its planning were also positively associated with the health professional support. Thus, frontline leaders have a critical role in engaging health professionals in planning and supporting their positive attitudes, which will in turn ensure active use of the new services [10].

Other study variables were positively associated with leader support in the univariate regression analysis. Most of the study variables were related to implementation practices associated with successful health IT adoption by earlier research [7]. As some of the implementation practices were not yet topical in the pre-implementation phase, the leaders seemed to form expectations based on their previous experiences in their organization. Thus, results suggest that the way technical innovations had previously been implemented influences support for a new service. In addition to implementation practices, change appropriateness regarding how the new portal influences

XSL•FO
RenderX

work processes and customers was positively associated with leader support as expected based on earlier research [12].

Ingebrigtsen et al [7] suggest that leaders at all levels of the organization support successful IT adoption. For example, all leadership levels should communicate a clear vision and the goals of the new system. Our research results suggest that frontline leaders are more prepared to support the new patient portal and communicate a clear vision to their subordinates if they have received information about the vision from the upper levels. As the strategic level is not in direct contact with health professionals, it is the frontline level that has a critical role in communicating the vision and providing leadership support. On the other hand, the frontline needs information and resources from the upper levels to be effective.

In summary, the results support earlier research findings. The contribution of our study is that it clarifies the role of the frontline leaders in the early phases of implementing a new patient portal. Our results highlight the importance of vision clarity and conveying information to the frontline leaders so they can create support among health professionals for a new patient portal. In large and complex organizations, frontline leaders have a critical role in communicating and spreading the vision for the health professionals, engaging them in planning and supporting positive attitudes among the professionals.

Limitations

In this study, we focused on the pre-implementation phase and a specific application, the patient portal for self-management. Due to the cross-sectional study design, it is not possible to infer causality relationships between variables. Future longitudinal studies should clarify the role of frontline leaders' pre-implementation perceptions in the later phases of implementation. The importance of top management support has been highlighted in earlier research [5] but was not directly studied here.

Furthermore, the variable quality of informing was not one-dimensional as it embedded aspects on how well the leaders received information and how well the leaders informed their own subordinates. The Cronbach alpha of the measure was very high (.92) supporting the assumption that the frontline leaders cannot inform their subordinates if they have not received the information from the organization.

We were not able to link leaders specifically with their subordinate health professionals. Instead, we averaged leader variables over the unit they worked in. However, the statistically significant association found between leader and professional variables in our study can be considered strong, since only averaged values of leader variables over a work unit could be used in the analysis. It is possible that associations would have been stronger and weaker associations could have been found if individual leaders could have been linked to their respective subordinates.

Acknowledgments

The authors are grateful to Hanna Nordlund and all the ODA project managers for their collaboration and help in conducting this study. This work was supported by the Strategic Research Council at the Academy of Finland, decision number 303606.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire.

[PDF File (Adobe PDF File), 123KB - jmir_v21i3e11413_app1.pdf]

Multimedia Appendix 2

Mean and standard deviation of the key leader variables.

[PDF File (Adobe PDF File), 97KB - jmir_v21i3e11413_app2.pdf]

Multimedia Appendix 3

Univariate regression analyses of the leader variables associated with leader support.

[PDF File (Adobe PDF File), 102KB - jmir_v21i3e11413_app3.pdf]

Multimedia Appendix 4

Univariate regression analyses of the leader variables associated with professional support.

[PDF File (Adobe PDF File), 98KB - jmir_v21i3e11413_app4.pdf]

XSL•FO
RenderX

References

1. Sligo J, Gauld R, Roberts V, Villa L. A literature review for large-scale health information system project planning, implementation and evaluation. Int J Med Inform 2017 Dec;97:86-97. [doi: 10.1016/j.ijmedinf.2016.09.007] [Medline: 27919399]
2. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. Int J Med Inform 2013 May;82(5):e73-e86. [doi: 10.1016/j.ijmedinf.2012.10.007] [Medline: 23146626]
3. Damberg CL, Ridgely MS, Shaw R, Meili RC, Sorbero MES, Bradley LA, et al. Adopting information technology to drive improvements in patient safety: lessons from the Agency for Healthcare Research and Quality health information technology grantees. Health Serv Res 2009 Apr;44(2 Pt 2):684-700 [FREE Full text] [Medline: 21456111]
4. Hsiao J, Chang H, Chen R. A study of factors affecting acceptance of hospital information systems: a nursing perspective. J Nurs Res 2011 Jun;19(2):150-160. [doi: 10.1097/JNR.0b013e31821cbb25] [Medline: 21586992]
5. Chen R, Hsiao J. An investigation on physicians' acceptance of hospital information systems: a case study. Int J Med Inform 2012 Dec;81(12):810-820. [doi: 10.1016/j.ijmedinf.2012.05.003] [Medline: 22652011]
6. Ludwick DA, Doucette J. Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries. Int J Med Inform 2009 Jan;78(1):22-31. [doi: 10.1016/j.ijmedinf.2008.06.005] [Medline: 18644745]
7. Ingebrigtsen T, Georgiou A, Clay-Williams R, Magrabi F, Hordern A, Prgomet M, et al. The impact of clinical leadership on health information technology adoption: systematic review. Int J Med Inform 2014 Jun;83(6):393-405. [doi: 10.1016/j.ijmedinf.2014.02.005] [Medline: 24656180]
8. Avgar AC, Litwin AS, Pronovost PJ. Drivers and barriers in health IT adoption: a proposed framework. Appl Clin Inform 2012;3(4):488-500 [FREE Full text] [doi: 10.4338/ACI-2012-07-R-0029] [Medline: 23646093]
9. Kuziemsky CE. Review of social and organizational issues in health information technology. Healthc Inform Res 2015 Jul;21(3):152-160 [FREE Full text] [doi: 10.4258/hir.2015.21.3.152] [Medline: 26279951]
10. Lapointe L, Rivard S. Getting physicians to accept new information technology: insights from case studies. CMAJ 2006 May 23;174(11):1573-1578 [FREE Full text] [doi: 10.1503/cmaj.050281] [Medline: 16717265]
11. 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]
12. Paré G, Sicotte C, Poba-Nzaou P, Balouzakis G. Clinicians' perceptions of organizational readiness for change in the context of clinical information system projects: insights from two cross-sectional surveys. Implement Sci 2011 Feb 28;6:15 [FREE Full text] [doi: 10.1186/1748-5908-6-15] [Medline: 21356080]
13. Holt DT, Vardaman JM. Toward a comprehensive understanding of readiness for change: the case for an expanded conceptualization. J Change Manag 2013 Mar;13(1):9-18. [doi: 10.1080/14697017.2013.768426]
14. Sicotte C, Paré G. Success in health information exchange projects: solving the implementation puzzle. Soc Sci Med 2010 Apr;70(8):1159-1165. [doi: 10.1016/j.socscimed.2009.11.041] [Medline: 20137847]
15. Keshavjee K, Bosomworth J, Copen J, Lai J, Kucukyazici B, Lilani R, et al. Best practices in EMR implementation: a systematic review. AMIA Annu Symp Proc 2006:982 [FREE Full text] [Medline: 17238601]
16. Information to support well-being and service renewal: eHealth and eSocial strategy 2020. URL: http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74459/URN_ISBN_978-952-00-3575-4.pdf [accessed 2018-06-21] [WebCite Cache ID 70L10R44o]
17. Ward MM, Vartak S, Schwichtenberg T, Wakefield DS. Nurses' perceptions of how clinical information system implementation affects workflow and patient care. Comput Inform Nurs 2011 Sep;29(9):502-511. [doi: 10.1097/NCN.0b013e31822b8798] [Medline: 21825972]
18. Wakefield DS, Halbesleben JRB, Ward MM, Qiu Q, Brokel J, Crandall D. Development of a measure of clinical information systems expectations and experiences. Med Care 2007 Sep;45(9):884-890. [doi: 10.1097/MLR.0b013e3180653625] [Medline: 17712259]
19. Kirkendall ES, Goldenhar LM, Simon JL, Wheeler DS, Spooner S. Transitioning from a computerized provider order entry and paper documentation system to an electronic health record: expectations and experiences of hospital staff. Int J Med Inform 2013 Nov;82(11):1037-1045. [doi: 10.1016/j.ijmedinf.2013.08.005] [Medline: 24041453]
20. Esmaeilzadeh P, Sambasivan M. Health Information Exchange (HIE): a literature review, assimilation pattern and a proposed classification for a new policy approach. J Biomed Inform 2016 Dec;64:74-86 [FREE Full text] [doi: 10.1016/j.jbi.2016.09.011] [Medline: 27645322]
21. Halbesleben JRB, Wakefield DS, Ward MM, Brokel J, Crandall D. The relationship between super users' attitudes and employee experiences with clinical information systems. Med Care Res Rev 2008 Sep 16;66(1):82-96. [doi: 10.1177/1077558708325984] [Medline: 19047764]
22. Peute LW, Aarts J, Bakker PJM, Jaspers MWM. Anatomy of a failure: a sociotechnical evaluation of a laboratory physician order entry system implementation. Int J Med Inform 2010 Apr;79(4):e58-e70. [doi: 10.1016/j.ijmedinf.2009.06.008] [Medline: 19640778]

XSL•FO
RenderX

23. Kruse CS, Kothman K, Anerobi K, Abanaka L. Adoption factors of the electronic health record: a systematic review. JMIR Med Inform 2016 Jun 01;4(2):e19 [FREE Full text] [doi: 10.2196/medinform.5525] [Medline: 27251559]
24. Straub D, Gefen D. Validation guidelines for IS positivist research. CAIS 2004;13:63. [doi: 10.17705/1CAIS.01324]
25. Kitchenham B, Pfleeger SL. Principles of survey research part 4. SIGSOFT Softw Eng Notes 2002 May 01;27(3):20-23. [doi: 10.1145/638574.638580]
26. Bowden A, Fox-Rushby JA, Nyandieka L, Wanjau J. Methods for pre-testing and piloting survey questions: illustrations from the KENQOL survey of health-related quality of life. Health Policy Plan 2002 Sep;17(3):322-330. [Medline: 12135999]
27. Boynton PM. Administering, analysing, and reporting your questionnaire. BMJ 2004 Jun 05;328(7452):1372-1375 [FREE Full text] [doi: 10.1136/bmj.328.7452.1372] [Medline: 15178620]
28. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.2307/3151312]
29. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. Qual Quant 2007 Mar 13;41(5):673-690. [doi: 10.1007/s11135-006-9018-6]
30. Menachemi N, Brooks RG. EHR and other IT adoption among physicians: results of a large-scale statewide analysis. J Healthc Inf Manag 2006;20(3):79-87. [Medline: 16903665]
31. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of ehealth interventions in inpatient routine care. J Health Commun 2017 Mar;22(3):274-284. [doi: 10.1080/10810730.2017.1284286] [Medline: 28248626]
32. Kujala S, Mugge R, Miron-Shatz T. The role of expectations in service evaluation: a longitudinal study of a proximity mobile payment service. Int J Hum-Comput Stud 2017 Feb;98:51-61. [doi: 10.1016/j.ijhcs.2016.09.011]

Abbreviations

DV: dependent variable
eHealth: electronic health
ICT: information and communication technology
IV: independent variable
ODA: omahoito ja digitaaliset arvopalvelut

Edited by G Eysenbach; submitted 28.06.18; peer-reviewed by D Klein, M Keller, M Rigby; comments to author 07.10.18; revised version received 20.11.18; accepted 29.12.18; published 22.03.19.

Please cite as:
Kujala S, Hörhammer I, Heponiemi T, Josefsson K
The Role of Frontline Leaders in Building Health Professional Support for a New Patient Portal: Survey Study
J Med Internet Res 2019;21(3):e11413
URL: http://www.jmir.org/2019/3/e11413/
doi:10.2196/11413
PMID:30901003

©Sari Kujala, Iiris Hörhammer, Tarja Heponiemi, Kim Josefsson. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient Engagement as Measured by Inpatient Portal Use: Methodology for Log File Analysis

Timothy Huerta[1,2,3,4*], PhD, MS; Naleef Fareed[2,4*], PhD, MBA; Jennifer L Hefner[1,3,4], PhD, MPH; Cynthia J Sieck[1,4], PhD, MPH; Christine Swoboda[4], PhD, MS; Robert Taylor[4], MA; Ann Scheck McAlearney[1,2,3,4], ScD, MS

[1]Department of Family Medicine, College of Medicine, The Ohio State University, Columbus, OH, United States

[2]Department of Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, OH, United States

[3]Division of Health Services Management and Policy, College of Public Health, The Ohio State University, Columbus, OH, United States

[4]CATALYST: Center for the Advancement of Team Science, Analytics, and Systems Thinking in Health Services and Implementation Science Research, College of Medicine, The Ohio State University, Columbus, OH, United States

[*]these authors contributed equally

Corresponding Author:
Timothy Huerta, PhD, MS
Department of Family Medicine
College of Medicine
The Ohio State University
460 Medical Center Drive, Suite 532
Columbus, OH, 43210
United States
Phone: 1 614 293 8973
Email: tim.huerta@osumc.edu

Abstract

Background: Inpatient portals (IPPs) have the potential to increase patient engagement and satisfaction with their health care. An IPP provides a hospitalized patient with similar functions to those found in outpatient portals, including the ability to view vital signs, laboratory results, and medication information; schedule appointments; and communicate with their providers. However, IPPs may offer additional functions such as meal planning, real-time messaging with the inpatient care team, daily schedules, and access to educational materials relevant to their specific condition. In practice, IPPs have been developed as websites and tablet apps, with hospitals providing the required technology as a component of care during the patient's stay.

Objective: This study aimed to describe how inpatients are using IPPs at the first academic medical center to implement a system-wide IPP and document the challenges and choices associated with this analytic process.

Methods: We analyzed the audit log files of IPP users hospitalized between January 2014 and January 2016. Data regarding the date/time and duration of interactions with each of the MyChart Bedside modules (eg, view lab results or medications and patient schedule) and activities (eg, messaging the provider and viewing educational videos) were captured as part of the system audit logs. The development of a construct to describe the length of time associated with a single coherent use of the tool—which we call a session—provides a foundational unit of analysis. We defined frequency as the number of sessions a patient has during a given provision day. We defined comprehensiveness in terms of the percentage of functions that an individual uses during a given provision day.

Results: The analytic process presented data challenges such as length of stay and tablet-provisioning factors. This study presents data visualizations to illustrate a series of data-cleaning issues. In the presence of these robust approaches to data cleaning, we present the baseline usage patterns associated with our patient panel. In addition to frequency and comprehensiveness, we present considerations of median data to mitigate the effect of outliers.

Conclusions: Although other studies have published usage data associated with IPPs, most have not explicated the challenges and choices associated with the analytic approach deployed within each study. Our intent in this study was to be somewhat exhaustive in this area, in part, because replicability requires common metrics. Our hope is that future researchers in this area will avail themselves of these perspectives to engage in critical assessment moving forward.

(J Med Internet Res 2019;21(3):e10957) doi:10.2196/10957

XSL•FO
RenderX

KEYWORDS

patient portals; health records, personal; health information technology; inpatient portals

Introduction

Background

The use of ambulatory patient portals has substantially expanded since 2011, coincident with requirements for patient engagement placed on health care facilities as a component of meaningful use. Ambulatory patient portals allow patients to view their medical information, schedule appointments, manage their medications, and communicate with their doctors [1]. Research on ambulatory portals has found that patients want to adopt the use of these tools at a rate faster than that in which health care facilities have promulgated the technology [2]. Patient portals can foster increased patient engagement, and frequent users of outpatient health applications, portals, and personal health records show improvement in both risk factors for chronic diseases [3-7] and health outcomes [8,9]. As patient engagement technology shows positive results, some health care facilities have begun to explore other technologies that may improve the patient's experience with the hope of impacting both clinical outcomes and satisfaction with care—inpatient portals (IPPs) are an example of such a technology.

Interest in IPPs stems from the benefits found in the use of outpatient portals as opposed to a response to regulation. Given the primary driver of health care costs at the national level is hospital-based care, facilities that can use technology to improve quality at a lower cost to the populations they serve—2 components of The Triple Aim [10]—are able to maintain a sustainable competitive advantage over their competition and, presumably, capture a greater component of the markets that they serve. An IPP provides a hospitalized patient with similar functions to those found in outpatient portals, including the ability to view vital signs, laboratory results, and medication information; schedule appointments; and communicate with their providers. However, IPPs may offer additional functions such as meal planning, real-time messaging with the inpatient care team, daily schedules, and access to educational materials relevant to their specific conditions [11,12]. In practice, IPPs have been developed as websites and tablet applications, with hospitals providing the required technology as a component of care during the patient's stay.

Although studies that have explored the effect of IPPs in specific contexts go back to 2011 [13-15], adoption of large-scale commercial IPPs has generally been slow. One of the first such implementations occurred in 2013 at the Ohio State University (OSU) Wexner Medical Center with Epic Systems' MyChart Bedside (MCB) product. As the first academic medical center to implement a system-wide IPP, OSU sought to use its unique position to document its journey and provide guidance on IPP adoption [1,11,12]. However, only recently has enough usage data been available upon which to explore outcomes through a *big data* approach—specifically through the analysis of audit log files.

Audit log files are server-side records of actions taken by a user. Every button pushed on a website or mobile app creates a record on the host computer, which can be used as the basis of an analysis. Log file analysis can be used to assess both when technology was accessed and what features of a program were used [16,17]. These log files are routinely assessed for operational purposes to help developers of websites and apps understand the behaviors of users to further improve these programs [18]. Log file analysis in the context of outpatient portals has been used to track the number of people using health applications in the patient portals; however, many of these studies measure only log-ins to the portal and not use of individual functions [6,8,19,20]. There has been limited research measuring the rate of use of one or two specific features of a tool, such as secure provider messaging or medication refill requests [21-23]. Log file analysis in the more recent IPP literature also shares similar constraints, where studies typically use crude measures to quantify IPP usage. These range from survey responses about use [24] to counts of overall use [25] or focus on the use of a few individual functions [26]. These kinds of studies report the total patients who used IPP and the total frequency of use across individual IPP functions; they typically do not report the limitations of using such measures.

The concept of use of a portal is multifaceted. For example, a researcher may define use in terms of the frequency with which a patient logs into the portal, or they may define use in terms of comprehensiveness, a function of the number of features a patient uses. Furthermore, although some may desire to measure use in a particular way, doing so may be hindered because of the manner in which the data are collected. As a result, the use of log files is not a straightforward task, and different definitions can be constructed to describe different elements of engagement. As a nascent area for research, there is limited scholarship and guidance on how one might approach the analysis of these data and explicate the challenges of analyzing log files.

Objectives

This study sought to address 2 aims. Primarily, it provides a documentation of our approach to processing IPP log files using data from our IPP implementation in January 2014 through the subsequent 2 years ending and in January 2016. This time frame represents a period with relatively stable use and lower distortions caused by implementation issues. Although other studies have published usage data, these previous studies did not describe the challenges and choices associated with the analytic approach deployed within each study. Therefore, our primary goal was to address and explicate methodological issues one might find in a similar analysis of log files in an institutional local data setting. Furthermore, to foster a standardized approach to this research, we included our analytic files as appendices that can be used to analyze Epic patient portal log files. Our intent was also to provide guidance for future reviewers related to ensuring that studies adhere to the highest quality of data analysis.

The secondary aim of our study was to provide descriptive results from our institution's data on IPP usage to present the implications of our methodological approach and the

XSL•FO
RenderX

assumptions around the decisions we chose. Our results also provided a glimpse into the usage of the IPP and its functionality in the context of care and presented initial insights into how patients use this tool. The interpretation of our data could help to identify patterns that others may validate in their own patient populations, based on the implementation experience with our IPP. Altogether, this study sought to encourage research rigor and replicability with an eye toward informing practice based on the first large-scale implementation in place. We addressed the primary aim of our study in the Methods section, and the secondary aim has been presented in the Results section with the help of descriptive statistics.

Methods

Data Source and Data Model

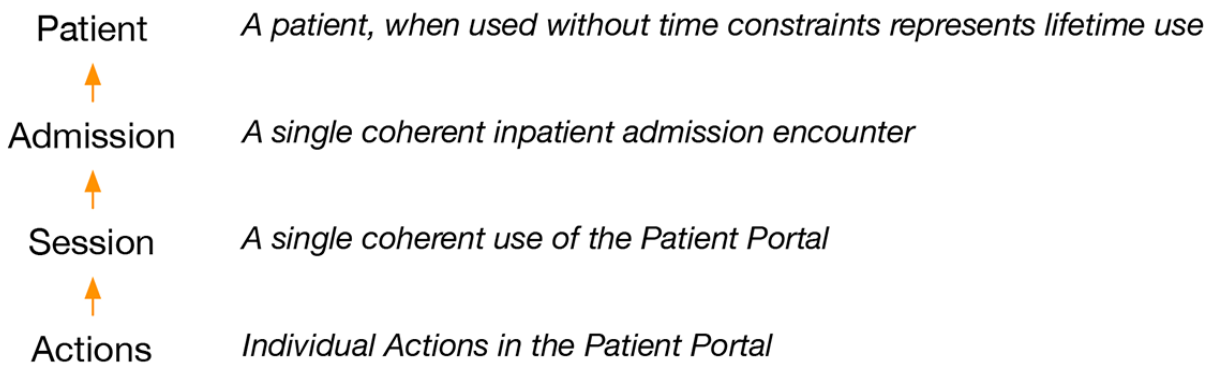
IPP use information is extracted from a limited dataset composed of retrospective data on MCB use in the form of audit log files from the OSU Information Warehouse, secured under an Institutional Review Board approval. The data for this study were gathered between January 2014 and January 2016. Data regarding the date/time and duration of interactions with each of the MCB modules (eg, view lab results or medications and patient schedule) and activities (eg, messaging the provider and viewing educational videos) are captured as part of the system audit logs. All events are recorded as a triad of data that includes the following:

1. A medical record number (MRN).
2. An activity code (WPR 530: UA-TYPE), augmented by the extended information (WPR 550: UA–EXTENDED INFO).
3. A timestamp including the date and time (MDY HMS) when the activity took place (WPR 520: UA - INSTANT).

To safeguard the privacy of the patient, the MRN was replaced by a study identifier by an honest broker in accordance with institutional policy for research-related requests forming a limited data set. The activity code represents the action taken by the patient (eg, a log-in, a log-out, accessing education materials, or sending a secure message). The parentheticals provide the variable name in the Epic data model for ease of identification. Other data sets that we merged with our audit log files are admission, discharge, and transfer (ADT) data; patient summary; and hospital charge data.

Our data model was primarily guided by 2 key objectives that determine the dimensionality of our final data set: (1) level of analysis of IPP use and (2) measure of IPP use. The level of analysis of IPP use indicates the rows of the data set, and the measure of IPP use represents the columns of our dataset (see Figure 1). From a conceptual standpoint, the level of analysis can be viewed as (1) session: a continuous period of portal use from the moment a patient logs into the IPP to the moment they logout; (2) admission: a period of continuous use during a clinical encounter, from admission to discharge from a unit; and (3) patient: all IPP use for a given patient across all inpatient encounters within the medical center. Similarly, the measure of IPP use can be viewed as (1) frequency: the count of use or the count of a particular action over a level of analysis; (2) comprehensiveness: a count of the number of unique actions or the number of unique activities over a level of analysis; and (3) duration: the amount of time spent using the portal over a level of analysis. Other important concepts that need to be considered for our data model are presented in Table 1.

Figure 1. Data aggregation model.



XSL•FO
RenderX

Table 1. Other important concepts related to the data model.

Concept	Description
User action or task	The recorded server actions that appear in the audit logs. Some represent an intentional action of the patient, whereas others represent background functioning of the portal (eg, an action will be recorded in the log file if the information on a given page is refreshed as the patient is viewing it).
Function	A group of user actions or tasks.
Provisioning	When the patient is provisioned with the hospital tablet, within a hospital encounter. This marks the point at which the portal is first available to the patient.
Length of stay	The number of calendar days a patient was in the care of the medical center.
Encounter	A patient's interaction with the medical center.
Contact serial number	A unique identification for each patient encounter.

Data Processing

The approach we deployed to process our data to structure it to fit our data model objectives described in the previous section involved 6 overarching modules. For each of these modules, we explicated their goals and the critical assumptions around the data processing below.

Module 1

Process raw ADT dataset (refer to Stata Module 1 in Multimedia Appendix 1): The goal of this module was to establish the start and end date of each specific inpatient clinical encounter during which a patient received continuous medical care. To achieve this, we first inspected the ADT patterns for a patient at a facility within OSU using the ADT data set. For the data to fit our definition of a clinical encounter or admission, we performed 2 transformations of the raw ADT data:

I. Combine overlapping encounters: Patients may have more than one recorded ADT event in the raw data within a time period. For example, a patient may be admitted for an inpatient stay, admitted to another hospital within the system for a procedure, and then return to the original room. The result is multiple admissions but a single discharge. We treated these as a single admission.
II. Combine adjacent encounters: Patients may have encounters that occur within a short time frame of each other. This is generally attributable to patients moving between different OSU facilities. Without a way to differentiate true discharges from transfers between facilities, we merged all encounters that occurred within 4 hours of each other (which covers approximately 80% of all transfers at our institution) to form a single admission event.

Module 2

Process raw hospital charges dataset (refer to Stata Module 2 in Multimedia Appendix 2): The goal of this module was to obtain information about patients' clinical diagnoses. As MCB is an IPP, initial processing of the hospital charges' data requires that data be restricted to only inpatient encounters (ie, excluding any instances of outpatient or emergency visits). As a single hospital charge may cover multiple encounters, we created a new variable: *hospital account id*, which links specific charges to an admission period. An artifact of creating this variable is the presence of duplicate *hospital account ids* with different admission and discharge times in the raw source data. As we used the raw ADT data to define admission times, we retained only one of the duplicate observations of *hospital account id* in the hospital charges files to obtain information about patients' clinical diagnoses and validated that this information was indeed the same across the duplicate observations.

Module 3

Process raw Audit Log dataset (refer to Stata Module 3 in Multimedia Appendix 3): The goal of this module was to apply existing category labels for the MCB function to users' actions. A review of raw audit log data found that the log files include activity codes, which are generated by the computational environment that represents administrative actions not initiated by the patient (eg, pushed data to ensure that the tablet does not display stale data). A review of system documentation, conversations with the implementation team, discussion with the vendor (Epic), and subsequent data visualizations were used to assist in the identification of data in the activity code associated with the idiosyncratic choices of the institution in its implementation, and those actions were removed from the analysis. We removed 2 user actions: (1) Get Menu Items and (2) Get Provider, as the former was a navigational action and the latter appeared as multiple separate actions, one for each Care Team member a patient has, whenever the Get Care Team action was performed. We also removed Get Wallpaper Data actions, which represented a page refresh action and occurred every 5 min (discussed in greater detail in Module 5).

Our approach resulted in 9 MCB functions (see Table 2). Analysis of the raw data also found that the logout and login variables are not reliable indicators of a contiguous MCB use period. Hence, we did not use these variables for our classification of users' actions. Table 3 provides the description and use case examples for each of the 9 functions identified by our study team.

XSL•FO
RenderX

Table 2. Active and inactive MyChart Bedside users' actions.

Functions	Users' action	Activity
Access educational materials	Get 1 patient-prescribed education title	Active
	Update education status	Active
	Get patient-prescribed education titles	Not active
Access personal notes	Create patient note	Active
	Delete patient media	Active
	Delete patient note	Active
	Update patient note	Active
	Get patient notes	Not active
Administrative	Identify user with lock	Active
	Make Bedside link	Active
	Send MyChart email	Active
	Accept terms and conditions	Not active
	Handshake	Not active
	Load terms and conditions	Not active
	Set lock for user	Not active
	Update photo for user	Not active
	Login/logout	Not active
Check secure messages	Get messages	Active
Happening soon	Create user-created event	Active
	Delete user-created event	Active
	Get appointment event detail	Active
	Get medication administration details	Active
	Get surgery event detail	Active
	Load schedule	Active
I would like	Delete patient request	Active
	Save patient request	Active
	Get patient requests	Not active
MyChart Ambulatory	Create MyChart account	Active
	Load MyChart info	Active
	Validate MyChart login	Not active
	Validate MyChart password	Not active
Review current care team	Get Care Team	Active
Review vitals and lab results	Get lab result comments	Active
	Get health metrics	Not active
	Switch bedside admission	Not active
Send a secure message	Save message	Active

XSL•FO
RenderX

Table 3. MyChart Bedside functions.

Function	Description	Use case example
Happening soon	Review scheduled upcoming tests or procedures	The patient wants to see what time they can expect the respiratory therapist to come for their breathing treatment.
I would like	Request one of a number of ancillary services	The patient wants to see a representative from Pastoral Care.
MyChart Ambulatory tasks	Access the ambulatory patient portal through the MCB[a] conduit	The patient wants to create an outpatient MyChart account or change the password on an existing MyChart account.
Review vitals and lab results	Review vital signs including blood pressure, heart rate, and temperature	The patient wants to see the results of their morning blood tests.
Send a secure message	Send a secure message to the care team	The patient wants to ask a question about their treatment.
Check secure messages	Check whether a secure message has been received	The patient wants to see whether a member of the care team has responded to their earlier question.
Access personal notes	Record and review personal notes (audio and written)	The patient has a question for their doctor that they want to remember to ask during rounds.
Review current care team	Review active members of the care team	The patient wants to find out the name of the nurse that just came in, but was too shy to ask.
Access educational materials	Access training materials through a link to an external health information content provider	The patient was assigned educational material by the care team. The care team will engage in teach-back once the patient is done or will ensure the content is discussed no later than 24 hours after assignment.

[a]MCB: MyChart Bedside.

Textbox 1. MyChart Bedside external actions.

External actions:

- Dining on demand
- Welcome video
- Getting started
- MyChart Bedside (MCB) patients' rights and responsibilities
- MCB patients' tutorial

In addition to the features available directly through MCB, the hospital also provides external tools, such as the ability to order food, through the portal. These external actions are identified in the log as *Media/Web content*, which the software uses for any remote internet call. The *extended info* variable (WPR 550) provides a unique code that identifies each of these external actions. Observations with these actions were replaced with specific MCB active tasks (see Textbox 1).

In some cases, multiple login events appeared sequentially, with no other actions recorded. It was determined that these did not provide any meaningful information, and all but one of these sequential logins was dropped. There were also instances where the same actions occurred multiple times with the same time stamp. We retained only one observation in such a case, although it was difficult to delineate 2 close observations in time as they may truly be unique or artifacts of system idiosyncrasies.

Module 4

Merge ADT, Audit Log, and Hospital Charges data sets (refer to Stata Module 4 in Multimedia Appendix 4): The goal of this module was to first link the data from the processed ADT and audit log data sets and then to link the new dataset with the Hospital Charges data set. As there was no unique identifier across the ADT and audit log data sets, we linked our data sets by matching all the records using the patient's MRN and keeping only the audit log observations that fell within a clinical encounter period. This process did yield some log observations that fell outside of a clinical encounter, and we dropped these observations in our linked data set. The dropped observations fell into 2 general categories, those that were (1) very close to encounter periods and (2) far from encounter periods. The former may be the result of potential inconsistencies in our session variables, and the latter phenomenon is harder to diagnose as use may appear days or even years after being removed from any patient encounter. These observations were removed from the data. Finally, we used the *hospital account id* from the ADT and Hospital Charges data sets to link the hospital charges file to our merged data set.

Module 5

Generate levels of analysis (refer to Stata Module 5 in Multimedia Appendix 5): The goal of this module was to create variables to flag different temporal dimensions within which

XSL•FO
RenderX

IPP use could be measured. Our first step was to create a variable that flags the beginning of each session. We defined a session to be a continuous period of portal use without 15 min of inactivity. It should be noted that our institution logs individuals out of the system after 10 min of inactivity; however, a review of the data showed the logout system was not always effective. As a result, we chose 15 min as a conservative estimate to represent sufficient inactive time to justify the end of a session. Using our definition of a patient admission, established in Module 1, we also created a variable to flag unique patient admissions. We were also able to identify unique patients who used the IPP by their MRN.

With respect to IPP sessions, we encountered additional idiosyncrasies in the data that required processing. As previously noted, the Get Wallpaper Data action resulted in page refreshes at intervals of 5 min. Figure 2 illustrates this phenomenon with session data, and Figure 3 illustrates the same data after we omit this action.

Module 6

Generate measures of use (refer to Stata Module 6 in Multimedia Appendix 5): The goal of this module was to create variables to measure IPP use from several different dimensions. As we attempted to develop our use measures, we first encountered challenges with patients using MCB who had extremely short or long lengths of stay (LOSs). Figure 4 resents the LOS associated with MCB users. The data approximate a log normal distribution; however, a small number of provisioned patients were long-term admissions to the hospital. Given the disproportionate effect of outliers on descriptive statistics, we sought to identify a point in the distribution where we no longer felt that we had sufficient data to make robust approximations. In the OSU case, we did not include individuals who were hospitalized longer than 30 days as the number of patients at that point fell below our *a priori* group size of 30 (indicating that at least 30 people had that total LOS). As we collected more data over time, we expected to push this boundary into longer LOSs.

Additionally, as illustrated in Figure 4, we identified another artifact in the audit log data. In cases where an individual logs in to review their current status on the home screen, the result is a *zero-time session*, one in which no additional functions are activated. If the goal is to create a metric for the duration of time the patient uses the tools, then that measure of duration would be systematically biased. As a result, the development of a metric of *duration*, defined as *the time spent in a particular session*, was determined to be unable to be constructed using audit log files from MCB. Hence, we focused on the creation of IPP use measures of frequency and comprehensiveness measures.

Furthermore, provisioning of the tablet with MCB occurs subsequent to admission and, as a result, the LOS and the number of provision days may not be equivalent. This was an important consideration because the use of LOS as a comparative construct creates systematic bias in the result. Consider a hypothetical patient who presented in the emergency room, was deemed nonresponsive, and was then admitted to the hospital. It may be that the patient becomes coherent on day 3 and provisioned with the IPP on day 5, based on established protocols. This difference between day first provisioned and LOS was visualized to assess the practice pattern associated with the provisioning process.

Each tile in Figure 5 represents the intersection of a patient's LOS on the *x*-axis with the day they were provisioned a tablet during that admission on the *y*-axis. In the context of this graph, LOS is constructed such that if a patient is admitted and discharged on the same day, their LOS would be equal to zero and is not based on an hourly calculation of days. The number in each tile represents the total number of patients at a given intersection of LOS and day of provisioning. For example, 719 patients had an LOS of 3 and were provisioned a tablet on the second day (day of provisioning=1 in Figure 5). The color gradient represents the percentage of patients provisioned on a given day within each LOS. It should be noted that if individuals were always provisioned on the day they were admitted, the graph would be the line of red tiles at the day of provisioning zero. What we see in fact is that there is significant variability in the time from admission to tablet provision. It may take several days for a patient to get their tablet and, as such, LOS as the basis of analysis may suffer from distortions associated with provisioning practice.

The research team also identified a number of exogenous factors that impacted the use of the IPP. For example, tablets are assigned to the patient while in a unit and, as a result, the patient could experience a discontinuity in access attributable to a change in unit, as was commonly the case when a patient was moved between units after a procedure (eg, prepartum to postpartum). Figure 5 focuses on the first 10 days because over 90% of all patients are provisioned for 10 days or less. As a result, the presentation of data in tables using the first 10 provision days represents a relatively robust approximation of usage for the majority of patients.

As a result of our processing, although duration, frequency, and comprehensiveness could offer a holistic view of IPP usage, only the latter 2 of those measures are robust. We have presented the preliminary results of this application of our methods in the next section using the frequency and comprehensiveness measures of IPP use.

XSL•FO
RenderX

Figure 2. Session length (n=139,181).

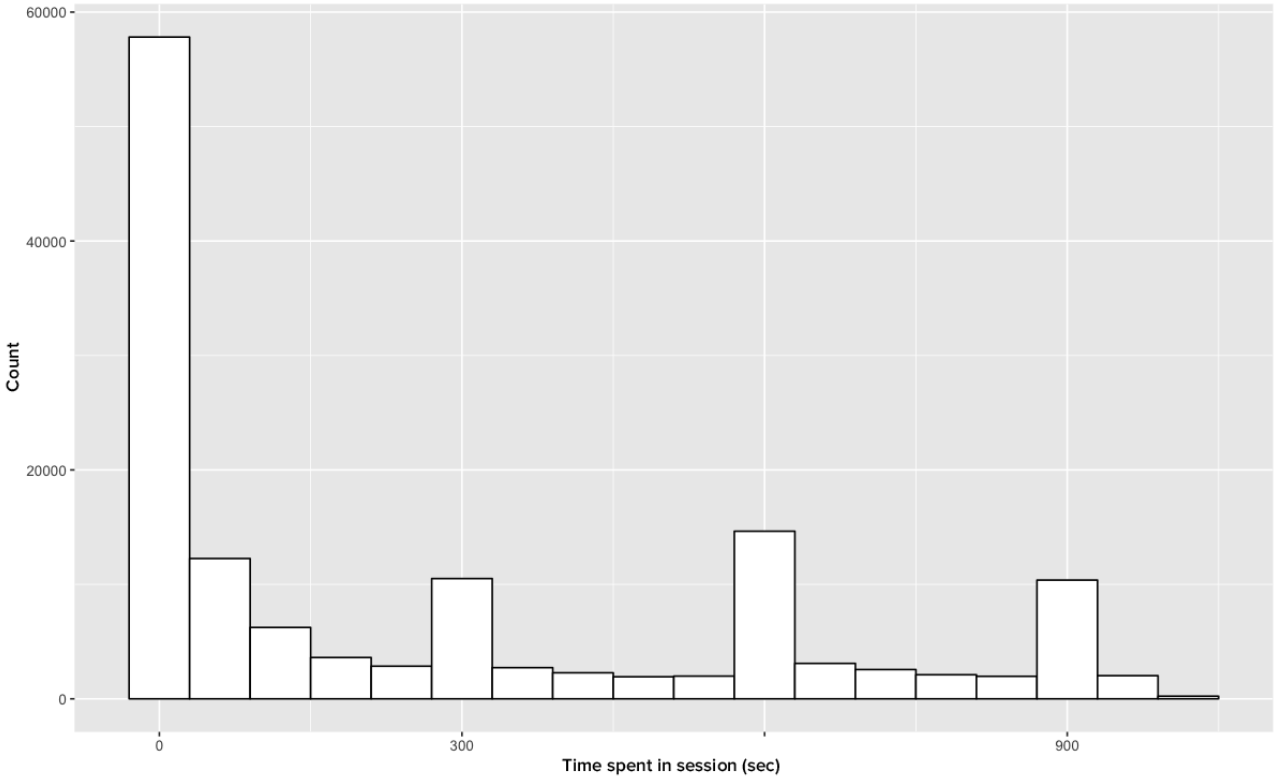
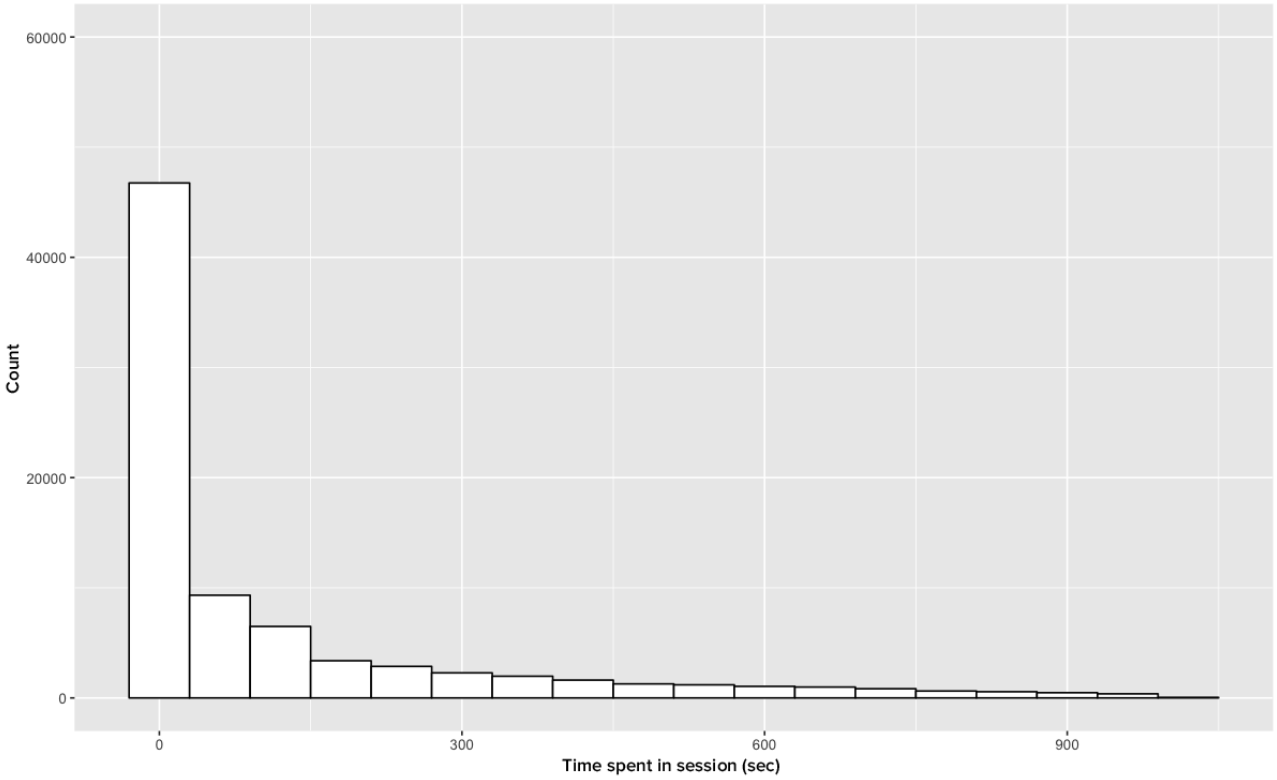


Figure 3. Session length (n=82,117).



XSL•FO
RenderX

Figure 4. Length of stay (n=6575).

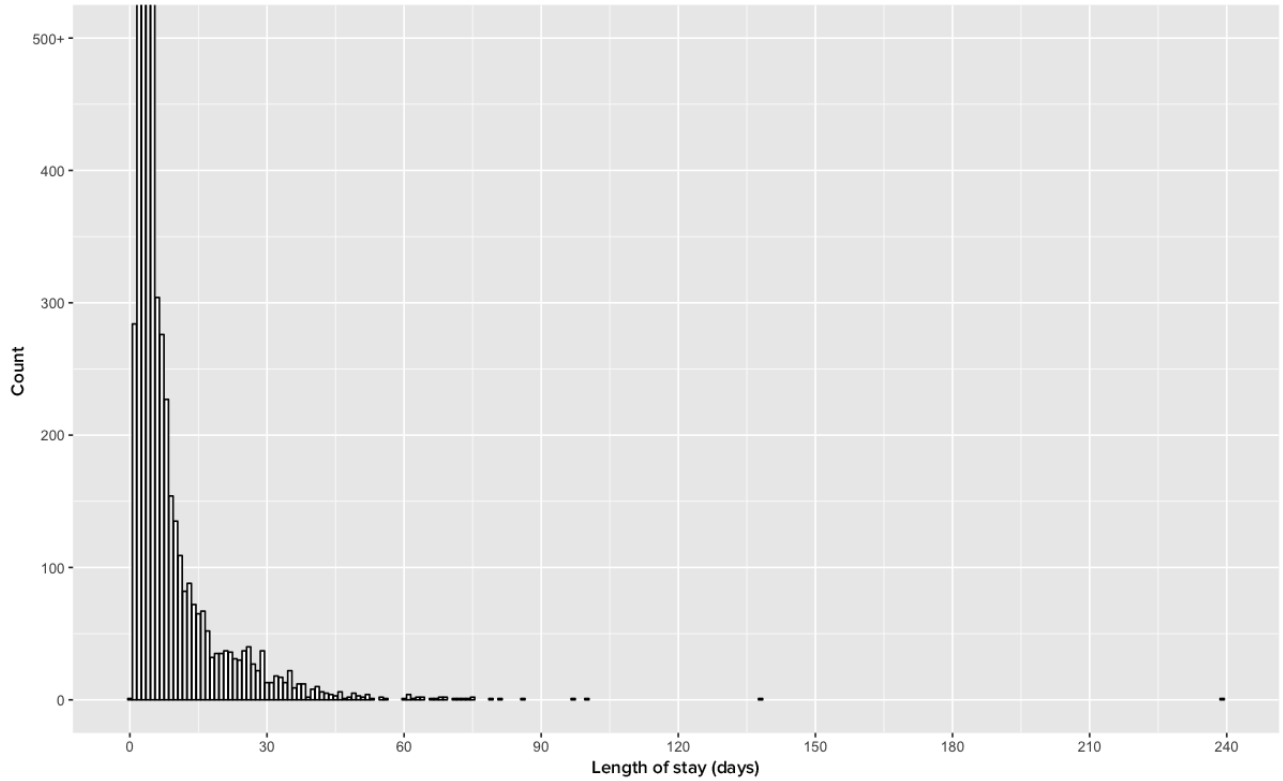
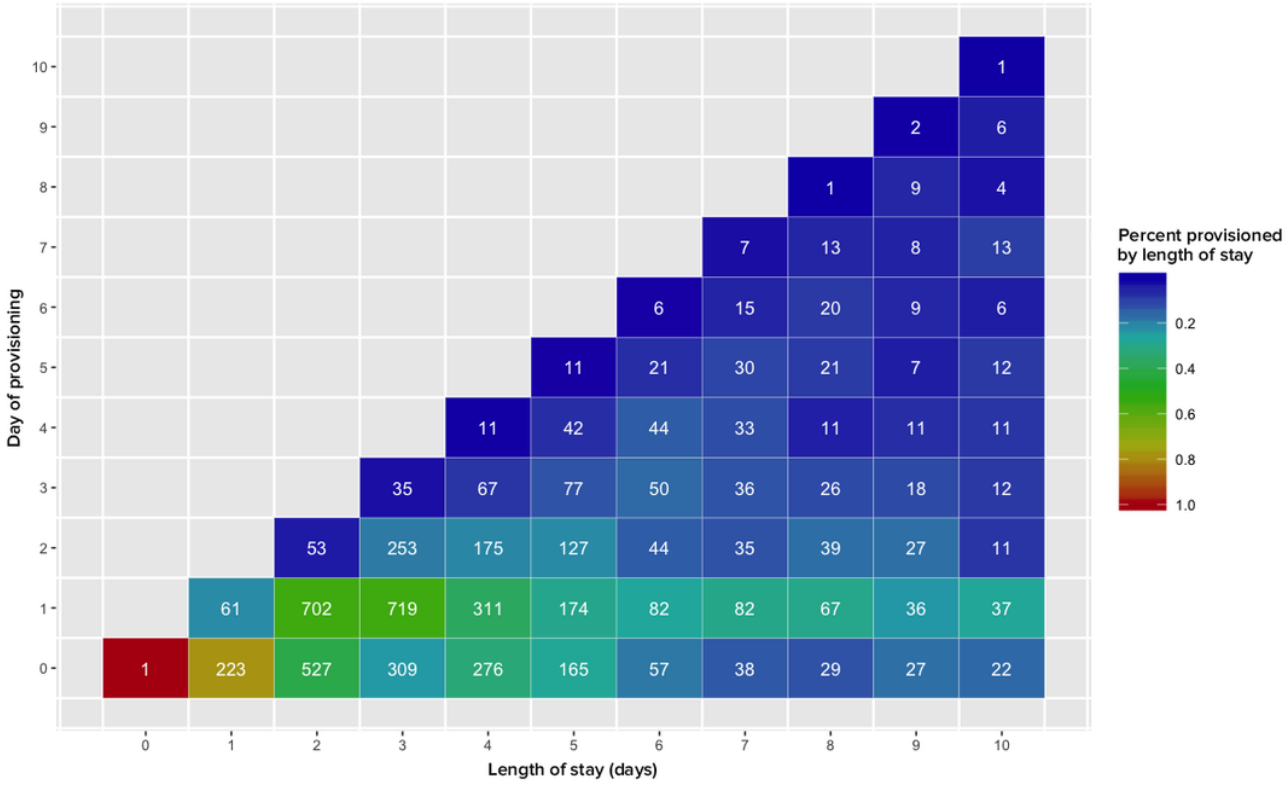


Figure 5. Day of provisioning by length of stay (n=5418).



XSL•FO
RenderX

Results

Overview

We applied our methodology for processing and merging the data sets described in the Methods section to MCB use data at OSU for the time periods between January 2014 and January 2016. Table 4 provides context for our study sample in relation to the patients seen by the OSU general patient population. Our sample of MCB users consisted of patients who were generally younger, more likely to be female, and had similar Charlson Comorbidity Index scores (range 0 to 19) to the OSU general population. They also had a slightly lower count of diagnoses than the general population.

The results below enumerate the implications of our decisions in the previous section, provide a glimpse into the usage of the IPP and its functionality in the context of care, and present initial insights into how patients use these tools. We offered the interpretation of our data so as to identify patterns that others may validate in their own patient populations, based on the implementation experience with our IPP. The results we presented below are based on the first 10 tablet provision days within an admission.

Frequency

Table 5 presents descriptive statistics on the frequency of IPP use across the various MCB functions based on our 3 levels of analysis: (1) session, (2) admission, and (3) patient. With respect to session, our results indicated that the median number of active tasks across sessions for all MCB functions was zero. This suggests that the majority of sessions contained only administrative/navigational tasks. We found that the top quartile for Happening Soon and View Care Team had 2 or more active tasks. Within a session, we found that the Send a Secure Message function was rarely used, albeit the Check Secure Messages function was used relatively more. The View Care Team function was the most popular function used. As previously noted in the Data Processing section, variables such as View Care Team may be inflated because of how the observations were recorded. For the admission and patient levels of analysis, the View Care Team and Happening Soon functions appear to be used more than the other MCB functions. These 2 functions also had a high level of spread based on their interquartile ranges. Like the session level, the Send a Secure Message was the least used function. Across the statistics, it is important to recognize that the admission and patient levels of analysis contain a high level of variability because of, for example, different numbers of admissions for a patient.

Table 4. Summary statistics of patient characteristics.

Characteristics	Study sample (n=5305)	Ohio State University general population (N=69,761)
Age (years), mean (SD)	45.06 (16.91)	53.64 (18.01)
Female (%)	75	55
Charlson Comorbidity Index, mean (SD)	3.06 (3.55)	2.77 (3.12)
Diagnoses, mean (SD)	13.05 (8.98)	15.49 (9.40)

Table 5. Frequency of MyChart Bedside functions.

Functions	Sessions (n=59,802)		Admissions (n=6105)		Patients (n=4979)	
	Median	IQR[a] (Min-Max)	Median	IQR (Min-Max)	Median	IQR (Min-Max)
Happening soon	0	2 (0-170)	6	19 (0-1011)	6	21 (0-3562)
I would like	0	0 (0-11)	0	1 (0-29)	0	1 (0-50)
MyChart Ambulatory tasks	0	0 (0-23)	0	1 (0-42)	0	1 (0-48)
View vitals and lab results	0	0 (0-60)	1	4 (0-226)	2	4 (0-312)
Check secure messages	0	0 (0-187)	1	3 (0-190)	1	3 (0-190)
Send a secure message	0	0 (0-4)	0	0 (0-12)	0	0 (0-12)
Access personal notes	0	0 (0-11)	0	1 (0-34)	0	1 (0-42)
View care team	0	2 (0-191)	6	13 (0-536)	6	14 (0-1963)
Access educational materials	0	0 (0-44)	1	2 (0-67)	1	3 (0-86)

[a]IQR: interquartile range.

XSL•FO
RenderX

Table 6. The number of distinct MyChart Bedside functions used.

Number of functions	Sessions (n=59,802), frequency (%)	Admissions (n=6105), frequency (%)	Patients (n=4979), frequency (%)
0	26,257 (43.91)	494 (8.09)	324 (6.51)
1	8,259 (13.81)	278 (4.55)	211 (4.24)
2	11,450 (19.15)	568 (9.30)	425 (8.54)
3	6041 (10.10)	651 (10.66)	519 (10.42)
4	3697 (6.18)	905 (14.82)	704 (14.14)
5	1766 (2.95)	813 (13.32)	633 (12.71)
6	945 (1.58)	588 (9.63)	485 (9.74)
7	852 (1.42)	663 (10.86)	576 (11.57)
8	533 (0.89)	1033 (16.92)	975 (19.58)
9	2 (0.00)	112 (1.83)	127 (2.55)

Comprehensiveness

Table 6 presents the descriptive statistics on the count of unique IPP actions performed by patients based on our 3 levels of analysis. As the session level, the results concur with our analysis using the frequency measure in the sense that the most common activities were administrative/navigational tasks. The admission and patient levels of analysis both reflect a similar pattern, where there is an even distribution between 2 and 8 functions. At 8 functions, we also found that the Check Secure Messages function was the least commonly used task. It should be noted that, across all levels, the proportion of data for using all 9 functions was very low.

We generated similar statistics for the frequency and comprehensive measures using the first 30 days within an admission (see supplementary tables in Multimedia Appendix 6). The patterns in these results did not vary significantly from the use measures presented using the first 10 tablet provision days.

Discussion

Summary

IPPs offer patients access to unique features not required in outpatient portals. Although significant research has been done to study outpatient portals, the nascent nature of IPPs lends them to exploration. As we noted in the Introduction section, the purpose of this study was twofold. The goal with our first aim was to address and explicate methodological issues that one might find in a similar analysis of audit log files from an institutional local data setting. Our intent was to provide guidance for future reviewers related to ensuring that studies adhere to the highest quality of data analysis. The second aim of our study was to offer a sense of the implications of our methodological approach, provide a glimpse into the usage of the IPP and its functions in the context of care, and present initial insights into how patients use these tools. Our interpretation of the data could help identify patterns that others may validate in their own patient populations.

With respect to the first aim, we provided 6 modules that described the goals and critical assumptions around our methodological approach. We supplemented these discussions with the Stata code (available as appendices in the Multimedia Appendix 1) used within each module so that researchers can apply them to their own data. In addition, below we have highlighted 4 key lessons that we learned related to this aim. These lessons should guide and inform future research using audit log files with the intention of helping researchers overcome many of the complex, tedious, and resource-intensive tasks involved with parsing through and combining this type of data [27]. Our hope is also that future researchers in this area will avail themselves of our perspectives to engage in critical assessment moving forward.

Lessons Learned

First, as the precision of recording of IPP user activity in the audit log files is imperfect, significant experience and familiarity are needed with the data to trace the flow of user activity [27-29]. To address this major challenge, we recognized, very early in the process, the need for a data model with prime directives that would enable us to reframe or restructure the data for the purposes of our study. This was especially important for our study given that the original intent of the data collection was not for research purposes but rather for our institution's operational needs. We found that explicating the levels of analysis and the measures of use were 2 useful objectives around which our data needed to be redefined. The need for strong assumptions around the data model (eg, constraining our analysis to 10 or fewer IPP provision days) was also necessary.

Second, we found that using different levels of analysis involved information that contained both unique and common aspects. Assessing IPP use from the session to the patient level provided a view of IPP utilization that ranged from a granular to a more aggregate perspective. Notably, our data indicated that IPP use at the admission level was strongly correlated with the relatively longer-term view of use at the patient level. This phenomenon needs to be further explored to confirm that this pattern exists beyond our institutions' data.

Third, and as other researchers have noted [30-33], we found that the choice of a specific IPP measure of use is complex and involves tradeoffs. The use of frequency provided us a count of the unique actions within a level of analysis, albeit the number

XSL•FO
RenderX

was sensitive to patients' LOSs (as noted by the extreme maximum points in our descriptive statistics for this measure). Although the comprehensiveness measure reflected whether patients used all the active IPP functions, this measure does not demonstrate users' intensity of use of specific functions.

Finally, we recognized the utility of higher-level categorizations of the over 30 IPP user actions. By making these data parsimonious, we were able to effectively analyze and generate statistics about the data that may have been more tedious and complicated by the many types of user actions. We, nonetheless, recommend that future research explore other approaches to studying users' actions to be able to generate robust findings about IPP use. These approaches may include techniques such as clustering, Gaussian mixture models, and multidimensional scaling.

With respect to the second aim, we offered an interpretation of our data based on the implementation experience with our IPP to identify patterns that others may validate in their own implementations and within their own patient populations. The analysis primarily demonstrates how the decisions we made around our data model's assumptions and the methodological steps we undertook influenced the presentation of the output for our results; this subsequently could determine the narrative used by researchers and practitioners to communicate the patterns of IPP user actions found in their data.

As organizations move forward with their implementation of these tools, the need for benchmarks is critical to understand the relative success or challenges of a specific implementation. We also noted that use of the technology is a long-tailed phenomenon impacted by the presence of different types of users. For instance, 1 user, who was provisioned for less than three days, viewed members of the care team 111 times. In practice, wide differences in use may stem from a number of reasons, including comfort with technology as well as a patient's disease state. Future research should attempt to identify the factors that contribute to distinguishing superusers (highly proficient users), hyperusers (high-frequency users), and intermittent users. In this instance, on provision day 1, 1 hyperuser had 41 separate sessions. To this end, we presented data in our results with both the interquartile ranges and the minimum and maximum to allow researchers to see the need for a classification schema moving forward.

Limitations

Key limitations should be noted with respect to our experience working with IPP audit log file data. Although we presented our methodology as a general approach to parsing and analyzing log file data, there are idiosyncrasies to such data that may exist specifically to an institution or the source of the data. It is impossible to consider all of these contingencies, but researchers should be mindful of the ones we have listed—along with the possibility of others—when using our approach.

Another limitation involves potential confounders that can influence the results presented. For example, the use of whiteboards as a means to communicate care team member information, changes to the care team members during a patient encounter, the times/days a patient uses the IPP, and the facility/unit within which the IPP is used may all influence the patterns identified in our data. However, this type of analysis is beyond the scope of our preliminary analysis and lends itself to future explorations that need to capture all of these and other possible contingencies in IPP use.

Finally, we submit that there may be other decisions that may be relevant or undiscovered with respect to how log file data need to be parsed or analyzed; these may have been potentially overlooked or they may be functions of the time and experience we have had with our data. Our hope is that we have motivated a conversation to advance future research to uncover these important aspects of the data that can help further improve the methodology used to manage these data.

Conclusions

This study represented the first volley into the IPP space by studying the experience of a health care system that has fully integrated the technology into its care processes. As we explore these data moving forward, we expect to offer some of the first glimpses into how such systems will be used in the context of care. Models for behavior change, such as the Health Belief Model [34], propose that systems such as IPPs may support patient engagement through education provided via the portal, increasing patients' knowledge about their condition as well as helping them assess the benefits of and barriers to taking action. However, our results would suggest that the longer the patient is in the hospital, the less likely they are to use the IPP to gain a greater understanding of their condition. Furthermore, although patients and providers can use the IPP to exchange information and educational messages can be tailored to patients' needs, it would seem that these phenomena are not yet happening in practice.

Patient portals offered in the inpatient setting present patients with targeted education and tools for managing health at a time when their perception of the threat from not managing their health is likely to be high, and therefore they would be more likely to engage with the technology [35]. Information and technology alone, however, are insufficient to fully engage patients in their care—patients also need motivation to engage [36-40]. A common element of health behavior change theories is the need for a trigger to action [34,41]. This is supported in studies of individual behavior change across a variety of health behaviors [42-46]. For many patients, hospitalization is often caused by an exacerbation of 1 or more conditions. We assert that hospitalization can serve as the necessary trigger that engages these patients in managing their care [34,41]. In other areas, times of acute crisis have been linked to a greater perception of risk and increased focus on health behaviors [42,43,45,47]. Therefore, hospitalization may create a window of higher motivation to engage, to initiate behavior change, and to foster interest in tools such as IPPs for managing health.

The use of IPPs may come to redefine how patients experience care in the hospital setting, and the use of this new technology may represent a paradigm-shifting change in the way care is delivered. Although the hospital experience is often about moving a patient out of crisis, it also represents an opportunity to influence a patient's assessment of the benefits and barriers to taking action (seriousness and risk) as well as their confidence

XSL•FO
RenderX

about being able to accomplish the necessary behavioral change required to achieve the desired consequence (self-efficacy). Patient portals in this setting may provide a means to increase patient self-efficacy during a particularly receptive time that can be continued after they transition to the outpatient environment.

Acknowledgments

This work was supported by the Agency for Healthcare Research and Quality (AHRQ) Grants R01 HS024091-01 and R21 HS024349-01 as well as the AHRQ Grant P30-HS024379 through the OSU Institute for the Design of Environments Aligned for Patient Safety.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Stata module 1.

[TXT File, 7KB - jmir_v21i3e10957_app1.txt]

Multimedia Appendix 2

Stata module 2.

[TXT File, 2KB - jmir_v21i3e10957_app2.txt]

Multimedia Appendix 3

Stata module 3.

[TXT File, 11KB - jmir_v21i3e10957_app3.txt]

Multimedia Appendix 4

Stata module 4.

[TXT File, 1KB - jmir_v21i3e10957_app4.txt]

Multimedia Appendix 5

Stata module 5.

[TXT File, 18KB - jmir_v21i3e10957_app5.txt]

Multimedia Appendix 6

Statistics for the frequency and comprehensive measures using the first 30 days within an admission.

[XLSX File (Microsoft Excel File), 12KB - jmir_v21i3e10957_app6.xlsx]

References

1. Huerta TR, McAlearney AS, Rizer MK. Introducing a patient portal and electronic tablets to inpatient care. Ann Intern Med 2017 Dec 05;167(11):816-817. [doi: 10.7326/M17-1766] [Medline: 29059690]
2. Ford EW, Hesse BW, Huerta TR. Personal health record use in the United States: forecasting future adoption levels. J Med Internet Res 2016 Mar 30;18(3):e73 [FREE Full text] [doi: 10.2196/jmir.4973] [Medline: 27030105]
3. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a Web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. J Med Internet Res 2011;13(3):e71 [FREE Full text] [doi: 10.2196/jmir.1603] [Medline: 21959968]
4. Kelders SM, van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
5. Sieverink F, Kelders SM, Braakman-Jansen LM, van Gemert-Pijnen JE. The added value of log file analyses of the use of a personal health record for patients with type 2 diabetes mellitus: preliminary results. J Diabetes Sci Technol 2014 Mar 4;8(2):247-255 [FREE Full text] [doi: 10.1177/1932296814525696] [Medline: 24876574]

XSL•FO
RenderX

6. Lau M, Campbell H, Tang T, Thompson DJ, Elliott T. Impact of patient use of an online patient portal on diabetes outcomes. Can J Diabetes 2014 Feb;38(1):17-21. [doi: 10.1016/j.jcjd.2013.10.005] [Medline: 24485208]
7. van Gemert-Pijnen JE, Kelders SM, Bohlmeijer ET. Understanding the usage of content in a mental health intervention for depression: an analysis of log data. J Med Internet Res 2014;16(1):e27 [FREE Full text] [doi: 10.2196/jmir.2991] [Medline: 24486914]
8. Phelps RG, Taylor J, Simpson K, Samuel J, Turner AN. Patients' continuing use of an online health record: a quantitative evaluation of 14,000 patient years of access data. J Med Internet Res 2014;16(10):e241 [FREE Full text] [doi: 10.2196/jmir.3371] [Medline: 25344621]
9. Tulu B, Trudel J, Strong DM, Johnson SA, Sundaresan D, Garber L. Patient portals: an underused resource for improving patient engagement. Chest 2016 Jan;149(1):272-277. [doi: 10.1378/chest.14-2559] [Medline: 26066707]
10. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769 [FREE Full text] [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
11. McAlearney AS, Sieck CJ, Hefner JL, Aldrich AM, Walker DM, Rizer MK, et al. High touch and high tech (HT2) proposal: transforming patient engagement throughout the continuum of care by engaging patients with portal technology at the bedside. JMIR Res Protoc 2016 Nov 29;5(4):e221 [FREE Full text] [doi: 10.2196/resprot.6355] [Medline: 27899338]
12. Hefner JL, Sieck CJ, Walker DM, Huerta TR, McAlearney AS. System-wide inpatient portal implementation: survey of health care team perceptions. JMIR Med Inform 2017 Sep 14;5(3):e31 [FREE Full text] [doi: 10.2196/medinform.7707] [Medline: 28912115]
13. Kelly MM, Hoonakker PLT, Dean SM. Using an inpatient portal to engage families in pediatric hospital care. J Am Med Inform Assoc 2017 Dec;24(1):153-161. [doi: 10.1093/jamia/ocw070] [Medline: 27301746]
14. Wilcox L, Woollen J, Prey J, Restaino S, Bakken S, Feiner S, et al. Interactive tools for inpatient medication tracking: a multi-phase study with cardiothoracic surgery patients. J Am Med Inform Assoc 2016 Jan;23(1):144-158 [FREE Full text] [doi: 10.1093/jamia/ocv160] [Medline: 26744489]
15. Vawdrey DK, Wilcox LG, Collins SA, Bakken S, Feiner S, Boyer A, et al. A tablet computer application for patients to participate in their hospital care. AMIA Annu Symp Proc 2011;2011:1428-1435 [FREE Full text] [Medline: 22195206]
16. Han JY. Transaction logfile analysis in health communication research: challenges and opportunities. Patient Educ Couns 2011 Mar;82(3):307-312. [doi: 10.1016/j.pec.2010.12.018] [Medline: 21277146]
17. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 7;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
18. Jansen BJ, Spink A, Taksai I. Handbook of Research on Web Log Analysis. Hershey, PA: IGI Global; 2009.
19. Riippa I, Linna M, Rönkkö I, Kröger V. Use of an electronic patient portal among the chronically ill: an observational study. J Med Internet Res 2014 Dec 8;16(12):e275 [FREE Full text] [doi: 10.2196/jmir.3722] [Medline: 25488754]
20. Toscos T, Daley C, Heral L, Doshi R, Chen Y, Eckert GJ, et al. Impact of electronic personal health record use on engagement and intermediate health outcomes among cardiac patients: a quasi-experimental study. J Am Med Inform Assoc 2016 Jan;23(1):119-128. [doi: 10.1093/jamia/ocv164] [Medline: 26912538]
21. North F, Hanna BK, Crane SJ, Smith SA, Tulledge-Scheitel SM, Stroebel RJ. Patient portal doldrums: does an exam room promotional video during an office visit increase patient portal registrations and portal use? J Am Med Inform Assoc 2011 Dec;18(Suppl 1):i24-i27 [FREE Full text] [doi: 10.1136/amiajnl-2011-000381] [Medline: 21976028]
22. Sarkar U, Lyles CR, Parker MM, Allen J, Nguyen R, Moffet HH, et al. Use of the refill function through an online patient portal is associated with improved adherence to statins in an integrated health system. Med Care 2014 Mar;52(3):194-201 [FREE Full text] [doi: 10.1097/MLR.0000000000000069] [Medline: 24374412]
23. Shimada SL, Allison JJ, Rosen AK, Feng H, Houston TK. Sustained use of patient portal features and improvements in diabetes physiological measures. J Med Internet Res 2016 Jul 1;18(7):e179 [FREE Full text] [doi: 10.2196/jmir.5663] [Medline: 27369696]
24. O'Leary KJ, Lohman ME, Culver E, Killarney A, Randy SG, Liebovitz DM. The effect of tablet computers with a mobile patient portal application on hospitalized patients' knowledge and activation. J Am Med Inform Assoc 2016 Jan;23(1):159-165. [doi: 10.1093/jamia/ocv058] [Medline: 26078412]
25. Kidwell KM, Peugh J, Westcott E, Nwankwo C, Britto MT, Quinn CT, et al. Acceptability and feasibility of a disease-specific patient portal in adolescents with sickle cell disease. J Pediatr Hematol Oncol 2018 Oct 18:-. [doi: 10.1097/MPH.0000000000001341] [Medline: 30339657]
26. Dalal AK, Dykes PC, Collins S, Lehmann LS, Ohashi K, Rozenblum R, et al. A web-based, patient-centered toolkit to engage patients and caregivers in the acute care setting: a preliminary evaluation. J Am Med Inform Assoc 2016 Jan;23(1):80-87. [doi: 10.1093/jamia/ocv093] [Medline: 26239859]
27. Oliner A, Ganapathi A, Xu W. Advances and challenges in log analysis. Commun ACM 2012 Feb 1;55(2):55-61. [doi: 10.1145/2076450.2076466]
28. Villella P, Petersen C. Google Patents. 2010. Log collection, structuring and processing URL: https://patents.google.com/patent/US20120246303A1/en [accessed 2019-02-22] [WebCite Cache ID 76NDjOFDc]

XSL•FO
RenderX

29. Gajula R, Prabhu V, Manyam B, Hunur N. Google Patents. 2015. Method and system for log file processing and generating a graphical user interface based thereon URL: https://patents.google.com/patent/US20090249250 [accessed 2019-02-22] [WebCite Cache ID 76NE4uXYF]
30. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
31. 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;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
32. Roehrs A, da Costa CA, Righi RD, de Oliveira KS. Personal health records: a systematic literature review. J Med Internet Res 2017 Jan 6;19(1):e13 [FREE Full text] [doi: 10.2196/jmir.5876] [Medline: 28062391]
33. Shimada SL, Petrakis BA, Rothendler JA, Zirkle M, Zhao S, Feng H, et al. An analysis of patient-provider secure messaging at two Veterans Health Administration medical centers: message content and resolution through secure messaging. J Am Med Inform Assoc 2017 Sep 1;24(5):942-949. [doi: 10.1093/jamia/ocx021] [Medline: 28371896]
34. Rosenstock I. The health belief model and preventive health behavior. Health Educ Behav 1974 Dec;2(4):354-386. [doi: 10.1177/109019817400200405]
35. Prey JE, Woollen J, Wilcox L, Sackeim AD, Hripcsak G, Bakken S, et al. Patient engagement in the inpatient setting: a systematic review. J Am Med Inform Assoc 2014;21(4):742-750 [FREE Full text] [doi: 10.1136/amiajnl-2013-002141] [Medline: 24272163]
36. Coulter A. Patient engagement--what works? J Ambul Care Manage 2012;35(2):80-89. [doi: 10.1097/JAC.0b013e318249e0fd] [Medline: 22415281]
37. Coulter A, Ellins J. Effectiveness of strategies for informing, educating, and involving patients. Br Med J 2007 Jul 7;335(7609):24-27 [FREE Full text] [doi: 10.1136/bmj.39246.581169.80] [Medline: 17615222]
38. Hibbard J, Greene J, Tusler M. Improving the outcomes of disease management by tailoring care to the patient's level of activation. Am J Manag Care 2009 Jun;15(6):353-360 [FREE Full text] [Medline: 19514801]
39. Laurance J, Henderson S, Howitt PJ, Matar M, al Kuwari H, Edgman-Levitan S, et al. Patient engagement: four case studies that highlight the potential for improved health outcomes and reduced costs. Health Aff (Millwood) 2014 Sep;33(9):1627-1634. [doi: 10.1377/hlthaff.2014.0375] [Medline: 25201668]
40. Simmons LA, Wolever RQ, Bechard EM, Snyderman R. Patient engagement as a risk factor in personalized health care: a systematic review of the literature on chronic disease. Genome Med 2014;6(2):16 [FREE Full text] [doi: 10.1186/gm533] [Medline: 24571651]
41. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Aug 26;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
42. Holtrop JS, Corser W, Jones G, Brooks G, Holmes-Rovner M, Stommel M. Health behavior goals of cardiac patients after hospitalization. Am J Health Behav 2006;30(4):387-399. [doi: 10.5555/ajhb.2006.30.4.387] [Medline: 16787129]
43. Lau K, Freyer-Adam J, Gaertner B, Rumpf H, John U, Hapke U. Motivation to change risky drinking and motivation to seek help for alcohol risk drinking among general hospital inpatients with problem drinking and alcohol-related diseases. Gen Hosp Psychiatry 2010;32(1):86-93. [doi: 10.1016/j.genhosppsych.2009.10.002] [Medline: 20114133]
44. Merrill RM, Aldana SG, Garrett J, Ross C. Effectiveness of a workplace wellness program for maintaining health and promoting healthy behaviors. J Occup Environ Med 2011 Jul;53(7):782-787. [doi: 10.1097/JOM.0b013e318220c2f4] [Medline: 21670705]
45. Sieck CJ, Heirich M. Focusing Attention on Substance Abuse in the Workplace: A Comparison of Three Workplace Interventions. J Workplace Behav Health 2010 Feb;25(1):72-87. [doi: 10.1080/15555240903358744]
46. Sieck CJ, Wickizer TM, Geist L. Population Health Management in Integrated Physical and Mental Health Care. In: Hefner JL, Huerta TR, McAlearney AS, editors. Population Health Management in Health Care Organizations. London: Emerald; 2014:139-150.
47. Sieck CJ, Heirich M, Major C. Alcohol counseling as part of general wellness counseling. Public Health Nurs 2004;21(2):137-143. [doi: 10.1111/j.0737-1209.2004.021207.x] [Medline: 14987213]

Abbreviations

ADT: admission, discharge, and transfer
IPP: inpatient portal
LOS: length of stay
MCB: MyChart Bedside
MRN: medical record number
OSU: Ohio State University

XSL•FO
RenderX

Edited by G Eysenbach; submitted 24.05.18; peer-reviewed by M Kelly, A Dalal, M Whetstone; comments to author 01.09.18; revised version received 28.12.18; accepted 29.12.18; published 25.03.19.

Please cite as:
Huerta T, Fareed N, Hefner JL, Sieck CJ, Swoboda C, Taylor R, McAlearney AS
Patient Engagement as Measured by Inpatient Portal Use: Methodology for Log File Analysis
J Med Internet Res 2019;21(3):e10957
URL: http://www.jmir.org/2019/3/e10957/
doi:10.2196/10957
PMID:30907733

©Timothy Huerta, Naleef Fareed, Jennifer L. Hefner, Cynthia J Sieck, Christine Swoboda, Robert Taylor, Ann Scheck McAlearney. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Disparities in Patient-Reported Interest in Web-Based Patient Portals: Survey at an Urban Academic Safety-Net Hospital

Shobha Sadasivaiah[1], MD, MPH; Courtney R Lyles[2], PhD; Stephen Kiyoi[1], MLIS, MS HAIL, AHIP; Piera Wong[3], RN, MS, AOCNS; Neda Ratanawongsa[2], MD, MPH

[1]Zuckerberg San Francisco General Hospital and Trauma Center, University of California San Francisco, San Francisco, CA, United States

[2]Zuckerberg San Francisco General Hospital and Trauma Center, UCSF Center for Vulnerable Populations, University of California San Francisco, San Francisco, CA, United States

[3]Zuckerberg San Francisco General Hospital and Trauma Center, San Francisco Department of Public Health, San Francisco, CA, United States

Corresponding Author:
Shobha Sadasivaiah, MD, MPH
Zuckerberg San Francisco General Hospital and Trauma Center
University of California San Francisco
1001 Potrero Avenue
San Francisco, CA,
United States
Phone: 1 415 206 2951
Fax: 1 415 206 4882
Email: shobha.sadasivaiah@ucsf.edu

Abstract

Background: Offering hospitalized patients' enrollment into a health system's patient portal may improve patient experience and engagement throughout the care continuum, especially across care transitions, but this process is less studied than portal engagement in the ambulatory setting. Patient portal disparities exist and may lead to differences in access or outcomes. As such, it is important to study upstream factors in a typical hospital workflow that could lead to those disparities in safety-net settings.

Objective: The objective of this study was to evaluate sociodemographic characteristics associated with interest in a health care system's portal among hospitalized patients and reasons for no interest.

Methods: Nurses assessed interest in a Web-based patient portal, expressed by the patient as "yes" or "no," as part of the admission nursing assessment among patients at an academic urban safety-net hospital and recorded responses in the electronic health record (EHR), including reasons for no interest. We extracted patient responses from the EHR.

Results: Among 23,994 hospitalizations over a 2-year period, 35.90% (8614/ 23,994) reported an interest in a Web-based portal. Reasons for *no* interest included the following: *not interested/other reason* 41.68% (6410/15,380), *no ability to use/access computers/internet* 29.59% (4551/15,380), *doesn't speak English* 11.15% (1715/15,380), *physically or mentally unable* 8.70% (1338/15,380), *does not want to say* 8.70% (1338/15,380), *security concerns* 0.03% (4/15,380), and *not useful* 0.16% (24/15,380). Among the 16,507 unique patients included in this sample, portal interest was lower in older, African American, non-English speaking, and homeless patient populations.

Conclusions: In a safety-net system, patient interest at the time of hospitalization in a Web-based enterprise portal—a required step before enrollment—is low with significant disparities by sociodemographic characteristics. To avoid worsening the digital divide, new strategies are needed and should be embedded within routine workflows to engage vulnerable safety-net patients in the use of Web-based health technologies.

(J Med Internet Res 2019;21(3):e11421) doi:10.2196/11421

KEYWORDS

patient portals; vulnerable populations; health care disparities; safety net hospitals; health information technology

XSL•FO
RenderX

Introduction

Background

Patient portals are websites that offer access to personal health information on the Web, such as test results, medical histories, immunization histories, and secure messaging with providers. Health care systems—leveraging federal incentives for meaningful use of electronic health records (EHRs)—are promoting patient portals to improve quality and safety, engagement, and care coordination with a goal of improved patient outcomes. Although the evidence is mixed about the benefits to health outcomes [1], portals have been associated with improvement in patients' diabetes medication adherence [2], understanding of medical conditions [3], and retention in a health care system [4].

Demographic factors such as age and education affect patient interest and use of patient portals [5]. Ambulatory and inpatient research suggests racial and ethnic disparities in portal use [6,7]. However, early research shows that diverse patients from safety-net health care settings have high interest in portal features such as electronic communication with providers [8], suggesting an opportunity for intervention.

Safety-net health care systems, systems that care for vulnerable, uninsured, or predominantly Medicaid patient populations, may face additional challenges engaging patients in their portals. For patients in safety-net settings, health care systems must be able to assess interest in enrollment, including frequently explaining to patients what a portal is and how portals may be useful for their care across the care continuum. Although portals have traditionally been studied across ambulatory settings [6], enrollment in a health care organization's portal during a hospital admission represents an opportunity to engage high-risk patients in their health. By providing patients information about their hospital course and recommendations for treatment and follow-up care, a health care system portal may help improve a patient's postdischarge self-care, transitions to ambulatory care, and ultimately improve health outcomes.

Objectives

No study has systematically assessed interest in a health network's portal within a typical hospital workflow among hospitalized urban safety-net patients. This study examined the interest in an enterprise-wide Web-based portal reported by hospitalized patients, how interest varied by sociodemographic characteristics, and the barriers among those declining portal interest.

Methods

Overview

This cross-sectional study evaluated hospitalized patients' interest in enrolling in a health care system's portal at an urban academic safety-net hospital. The patient portal offered information about the hospitalization (hospital after-visit summary with discharge instructions, summary of hospital course, and test results), as well as ambulatory patient portal content (primary and specialty care after-visit summaries, medications, allergies, problem lists, test results, and secure messaging with primary care providers). Eligible patients included adults (aged 18 years and over) admitted to acute medical, surgical, or gynecologic obstetric units over a 2-year period (September 2015-September 2017). This sample excludes patients under 18 years of age, admitted to skilled nursing or psychiatric units or reporting existing portal accounts.

Patient Assessment

Nursing leadership provided universal education to bedside nurses about the content and value of the patient portal and how to assess interest using an EHR-built admission assessment question stem (*Are you interested in learning more about or signing up for our free, secure website to view your health record online? The website gives you access to your lab results, medication lists, discharge instructions, and other information from your stay*). Nurses could record patient interest as *yes* or *no*, followed by a structured drop-down menu based on existing literature about portal barriers [9,10]: *no ability to use/access computers/internet*, *doesn't speak English*, *physically or mentally unable*, *not interested/other reason*, *doesn't way to say*, or nurses could write over a free- text reason. By incorporating this assessment into routine clinical operations, nursing leadership standardized the verbiage of the assessment, but nurses could individualize any clarifying questions or follow-up counseling about the portal. Nurses did not offer specific technology or portal demonstrations.

The data for this analysis were abstracted from the EHR nursing assessments as well as the EHR sociodemographic section (age, gender, race/ethnicity, language, insurance type, and homelessness).

Analysis

The primary outcome was the proportion of patients with portal interest recorded as *yes* during at least one admission. As some patients may change their responses in subsequent admissions, we also calculated the proportion of portal interest expressed at the hospitalization level and calculated the proportion of reasons for *No* interest during each hospitalization, coding free-text reasons into the existing or separate categories.

We then used logistic regression with measured sociodemographic covariates to calculate the unadjusted and adjusted odds of portal interest during at least one admission. Statistical analysis was performed using Stata version 15.1 (StataCorp). The University of California San Francisco Institutional Review Board approved this study.

Results

Demographics

Among 16,507 unique patients, the average age was 53 years (SD 19.17): 42.16% (6960/16,507) were women, 26.43% (4362/16,507) white, 28.02% (4625/16,507) Hispanic, 18.56% (3063/16,507) African American, 19.06% (3147/16,507) Asian, 7.93% (1309/16,507) other. Furthermore, 49.35% (8146/16,507) preferred a non-English language--10.67% (1761/16,507) Spanish and 5.40% (892/16,507) Chinese. 80.50% (13,288/16,507) had Medicaid or Medicare, and 14.17%

XSL•FO
RenderX

(2339/16,507) reported homelessness. There was an average of 1.4 admissions per patient in this cohort during the period studied (see Table 1).

Reasons for Lack of Portal Interest

When analyzed across 23,995 admissions, patients reported interest in the portal during 35.90% (8614/ 23,994) of admissions. The most commonly documented reasons for *no* were the following: *not interested* 41.68% (6410/15,380), *no ability to use/access computers/internet* 29.59% (4551/15,380), *doesn't speak English* 11.15% (1715/15,380) , *physically or mentally unable* 8.70% (1338/15,380), *does not want to say* 8.70% (1338/15,380), *security concerns* 0.03% (4/15,380), *not useful* 0.16% (24/15,380).

Table 1. Demographic characteristics of adults asked about portal interest during admissions to an urban academic safety-net hospital between September 2015 and September 2017.

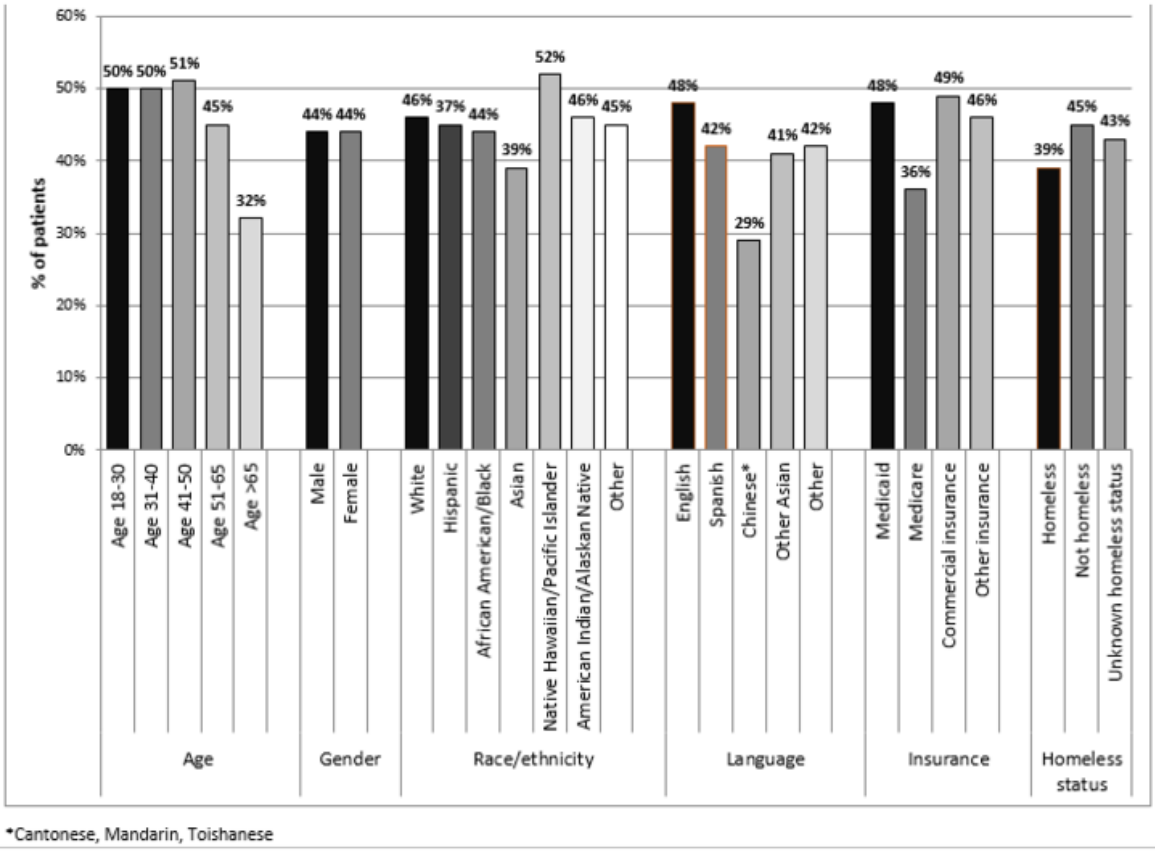
Patient Characteristics (N=16,507)	Statistics
Age (years), mean (SD)	53.06 (19.17)
Women, n (%)	6960 (42.16)
Self-reported race/ethnicity, n (%)	
White	4362 (26.43)
Hispanic	4625 (28.02)
African American/black	3063 (18.56)
Asian	3147 (19.06)
Native Hawaiian/Pacific Islander	222 (1.34)
American Indian/Alaskan Native	120 (0.73)
Other/mixed	967 (5.86)
Primary language, n (%)	
English	8361 (50.65)
Spanish	1761 (10.67)
Cantonese, Mandarin, or Toishanese	892 (5.40)
Other Asian languages, n (%)	208 (1.26)
Other, n (%)	5285 (32.02)
Insurance, n (%)	
Medicaid or county health plan for uninsured[a]	8278 (50.15)
Medicare	5010 (30.35)
Commercial insurance	1300 (7.88)
Other insurance	1919 (11.63)
Homeless status	
Yes[b], n (%)	2339 (14.17)
No, n (%)	12478 (75.59)
Unknown, n (%)	1690 (10.24)
Admissions, mean (range)	1.4 (1-33)

[a]These health care access programs (administered by the county Medicaid managed care organization) provide medical services for uninsured county residents (aged 18-64 years) or adults who provide county-supported in-home support services.

[b]Patients reported *yes* during at least one admission.

XSL•FO
RenderX

Figure 1. Patient interest in an online healthcare system portal during at least one admission to an academic safety net hospital (n=16507 patients, September 2015-September 2017).



Interest by Sociodemographic Characteristic

However, 44% of unique patients reported interest in the patient portal during at least one admission. Figure 1 shows unadjusted proportions of interest in the portal during at least one admission, stratified across sociodemographic characteristics. In unadjusted analyses (Table 2), the odds of expressing interest in the Web-based portal were significantly lower for patients in the 50 to 65 years age group (adjusted odds ratio [AOR] 0.82, 95% CI 0.74-0.90 vs 18-29 age group), the greater than 65 years of age group (AOR 0.47, 95% CI 0.42-0.52 vs 18-29 age group), Asians (AOR 0.76, 95% CI 0.69-0.84 vs white), Spanish-speaking (AOR 0.78, 95% CI 0.70-0.86 vs English), Chinese-speaking (AOR 0.46, 95% CI 0.39-0.53 vs English), patients on Medicare (AOR 0.60, 95% CI 0.56-0.65 vs Medicaid), and patients reporting homelessness (AOR 0.77 95% CI 0.70-0.84 vs not homeless).

In adjusted models (Table 2), older age (AOR 0.85, 95% CI 0.77-0.95 for >50 and AOR 0.50, 95% CI 0.44-0.57 for >65 vs 18-29), African American race (AOR 0.87, 95% CI 0.79-0.96 vs white), Spanish or Chinese languages (AOR 0.69, 95% CI 0.61-0.78 and AOR 0.54, 95% CI 0.46-0.65 vs English), and homelessness (AOR 0.60, 95% CI 0.54-0.66 vs housed) were significantly associated with *No* portal interest during any hospitalization.

XSL•FO
RenderX

Table 2. Odds of expressing interest in a patient portal among patients at an urban academic safety-net hospital (n=16,507 unique patients).

Sociodemographic characteristic	Unadjusted odds ratio	*P* value	Adjusted odds ratio[a]	*P* value
Gender				
Male	Ref	—[b]	Ref[c]	—
Female	1.01	0.85	1.04	0.25
Age (years)				
18-29	Ref	—	Ref	—
30-39	0.99	0.91	1.01	0.9
40-49	1.04	0.49	1.09	0.15
50-65	0.82	<.001	0.85	0.003
>65	0.47	<.001	0.5	<.001
Race or ethnicity				
White	Ref	—	Ref	—
Hispanic	0.98	0.65	0.93	0.15
African American	0.95	0.24	0.87	0.01
Asian	0.76	<.001	0.91	0.1
Native Hawaiian/Pacific Islander	1.31	0.05	1.09	0.52
American Indian/Alaskan Native	1.01	0.95	0.94	0.74
Other	1	0.96	0.92	0.24
Language				
English	Ref	—	Ref	—
Spanish	0.78	<.001	0.69	<.001
Cantonese, Mandarin, or Toishanese	0.46	<.001	0.54	<.001
Other Asian	0.77	0.07	0.9	0.47
Other	0.78	<.001	0.78	<.001
Insurance				
Medicaid	Ref	—	Ref	—
Medicare	0.6	<.001	0.91	0.07
Commercial insurance	1.03	0.64	0.94	0.34
Other insurance	0.94	0.2	0.98	0.71
Homeless status				
Not homeless	Ref	—	Ref	—
Homeless	0.77	<.001	0.6	<.001
Unknown homeless status	0.91	0.07	0.85	0.003

[a]Logistic regression adjusted for all the variable listed in the table.

[b]Not applicable.

[c]Ref: reference value.

Discussion

Conclusions

Fewer than half of the patients admitted to a safety-net hospital were documented by nurses as interested in the patient portal during at least one of their admissions to a safety-net hospital, with significantly lower odds of portal interest among older, nonwhite, non-English-speaking, and homeless patients. This is the first attempt at universally quantifying how many hospitalized patients are interested in a health care system's portal and highlighting the prevalence of patient barriers to portal engagement within a safety-net setting.

A notable finding from this analysis was that the proportion of hospitalizations during which patients expressed portal interest (31%) was lower than the proportion of patients expressing portal interest during at least one hospitalization (44%), a finding

XSL•FO
RenderX

that remained true across sociodemographic categories. This finding suggests that offering the patient portal to a patient who has previously declined may be useful for subsequent engagement. Organizational, technical, or workflow protocols that fail to offer portal enrollment to patients who previously declined may exacerbate disparities. In safety-net systems, the simple intervention of standardizing offers with every hospitalization may be an important tool for engaging patients in these technologies.

Most hospitals—particularly safety-net hospitals with limited resources—may rely on their existing staff to engage patients in their portal. This paper reveals what may naturally happen in a typical workflow where engagement happens universally as a required part of admission. A strength of this study is that the nursing staff were essential partners and allies for portal promotion given their role in providing direct clinical care; this study was driven strongly by nursing leadership. As portal interest screening was incorporated as a required component of the initial admission nursing assessment, this study allowed us to capture sufficient data to understand demographic characteristics associated with portal use.

Limitations

Nonetheless, there are limitations to nurses alone as the only source of assessing patient interest, with potential bias in interpreting patients' reasons for *no* interest. In addition, the time of admission may not be the best time during the course of a patient's hospitalization to make this type of assessment depending on the acuity of a patient's condition. Although this study did not assess portal enrollment or use, the results highlight potential disparities in a necessary and required upstream process.

Future Directions

Although Web-based portals have historically been associated with ambulatory care services, interest and research on their use for hospitalized patients are rising. Acute care portals that provide patients with health care information and communication with their teams during hospitalization have been studied for usability [11,12], implementation [13,14], and association with readmission rates [15]. Incorporating acute care portal functionality into a health care system portal may increase interest among hospitalized patients in using a platform that crosses a system's health care settings, and further research on acute care portal design and implementation is needed. However, our findings suggest that the period of hospitalization also offers an opportunity to engage patients in a health care system portal that lacks acute care functionality. Indeed, after discharge, patients and their loved ones may have greater energy and capacity for understanding the hospital course and discharge instructions; access to this information may facilitate their ability to communicate important information to their ambulatory care providers.

This study is the first to begin to understand issues around portal use among inpatients, specifically in a safety-net setting. Research suggests that vulnerable populations need significant support to register and use a portal website [16,17]. Safety-net hospitals may have significant workflow challenges with providing portal education and support tailored to patients' language, literacy, vision/hearing, and other needs. These findings triggered local quality improvement initiatives that could be replicated elsewhere: portal and internet education coordinated by the hospital library, in-person and phone technical support for patients, a Web-based curriculum to teach skills in accessing and navigating the portal, and promotional materials about designating a caregiver proxy in our 5 key languages. Future research and policies to incentivize health information technology should incorporate strategies for tailoring patient portal usability and implementation for diverse populations to avoid worsening the digital divide.

Acknowledgments

Each author has participated sufficiently in the work to take public responsibility for appropriate portions of the content. SS takes responsibility for the integrity of the work as a whole, from inception to publication. Authorship credit was based only on (1) substantial contributions to conception and design or acquisition of data or analysis and interpretation of data and (2) drafting the article or revising it critically for important intellectual content. This study was funded by AHRQ Grants R00HS022408, 1K08HS022561, and P30HS023558, National Library of Medicine Grant G08 NLM012166, and it received funding from the Gordon and Betty Moore Foundation. The funders did not play any role in the review or approval of the manuscript for publication. Publication made possible in part by support from the University of California San Francisco Open Access Publishing Fund.

Conflicts of Interest

None declared.

References

1. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012 Nov 26;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
2. Lyles CR, Sarkar U, Schillinger D, Ralston JD, Allen JY, Nguyen R, et al. Refilling medications through an online patient portal: consistent improvements in adherence across racial/ethnic groups. J Am Med Inform Assoc 2016 Apr;23(e1):e28-e33 [FREE Full text] [doi: 10.1093/jamia/ocv126] [Medline: 26335983]

XSL•FO
RenderX

3. Turvey C, Klein D, Fix G, Hogan TP, Woods S, Simon SR, et al. Blue Button use by patients to access and share health record information using the Department of Veterans Affairs' online patient portal. J Am Med Inform Assoc 2014;21(4):657-663. [doi: 10.1136/amiajnl-2014-002723] [Medline: 24740865]
4. Turley M, Garrido T, Lowenthal A, Zhou YY. Association between personal health record enrollment and patient loyalty. Am J Manag Care 2012 Jul 1;18(7):e248-e253 [FREE Full text] [Medline: 22823553]
5. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
6. Wallace LS, Angier H, Huguet N, Gaudino JA, Krist A, Dearing M, et al. Patterns of electronic portal use among vulnerable patients in a nationwide practice-based research network: from the OCHIN Practice-based Research Network (PBRN). J Am Board Fam Med 2016 Oct;29(5):592-603 [FREE Full text] [doi: 10.3122/jabfm.2016.05.160046] [Medline: 27613792]
7. Robinson JR, Davis SE, Cronin RM, Jackson GP. Use of a patient portal during hospital admissions to surgical services. AMIA Annu Symp Proc 2016;2016:1967-1976 [FREE Full text] [Medline: 28269956]
8. Schickedanz A, Huang D, Lopez A, Cheung E, Lyles CR, Bodenheimer T, et al. Access, interest, and attitudes toward electronic communication for health care among patients in the medical safety net. J Gen Intern Med 2013 Jul;28(7):914-920 [FREE Full text] [doi: 10.1007/s11606-012-2329-5] [Medline: 23423453]
9. Lyles CR, Harris LT, Jordan L, Grothaus L, Wehnes L, Reid RJ, et al. Patient race/ethnicity and shared medical record use among diabetes patients. Med Care 2012 May;50(5):434-440. [doi: 10.1097/MLR.0b013e318249d81b] [Medline: 22354209]
10. Lyles CR, Allen JY, Poole D, Tieu L, Kanter MH, Garrido T. "I Want to Keep the Personal Relationship With My Doctor": understanding barriers to portal use among African Americans and Latinos. J Med Internet Res 2016 Dec 3;18(10):e263 [FREE Full text] [doi: 10.2196/jmir.5910] [Medline: 27697748]
11. Walker DM, Menser T, Yen P, McAlearney AS. Optimizing the user experience: identifying opportunities to improve use of an inpatient portal. Appl Clin Inform 2018 Dec;9(1):105-113. [doi: 10.1055/s-0037-1621732] [Medline: 29444536]
12. Kelly MM, Coller RJ, Hoonakker PL. Inpatient portals for hospitalized patients and caregivers: a systematic review. J Hosp Med 2018 Dec 1;13(6):405-412. [doi: 10.12788/jhm.2894] [Medline: 29261819]
13. Grossman LV, Choi SW, Collins S, Dykes PC, O'Leary KJ, Rizer M, et al. Implementation of acute care patient portals: recommendations on utility and use from six early adopters. J Am Med Inform Assoc 2018 Apr 1;25(4):370-379. [doi: 10.1093/jamia/ocx074] [Medline: 29040634]
14. Kelly MM, Dean SM, Carayon P, Wetterneck TB, Hoonakker PL. Healthcare team perceptions of a portal for parents of hospitalized children before and after implementation. Appl Clin Inform 2017 Dec 15;8(1):265-278 [FREE Full text] [doi: 10.4338/ACI-2016-11-RA-0194] [Medline: 28293685]
15. Griffin A, Skinner A, Thornhill J, Weinberger M. Patient portals: who uses them? What features do they use? And do they reduce hospital readmissions? Appl Clin Inform 2016;7(2):489-501. [doi: 10.4338/ACI-2016-01-RA-0003] [Medline: 27437056]
16. Tieu L, Schillinger D, Sarkar U, Hoskote M, Hahn KJ, Ratanawongsa N, et al. Online patient websites for electronic health record access among vulnerable populations: portals to nowhere? J Am Med Inform Assoc 2016 Jul 8:e47-e54. [doi: 10.1093/jamia/ocw098] [Medline: 27402138]
17. Tieu L, Sarkar U, Schillinger D, Ralston JD, Ratanawongsa N, Pasick R, et al. Barriers and facilitators to online portal use among patients and caregivers in a safety net health care system: a qualitative study. J Med Internet Res 2015;17(12):e275 [FREE Full text] [doi: 10.2196/jmir.4847] [Medline: 26681155]

Abbreviations

AOR: adjusted odds ratio
EHR: electronic health record

Edited by G Eysenbach; submitted 11.09.18; peer-reviewed by K Blondon, J Hefner, S Prior; comments to author 12.10.18; revised version received 05.12.18; accepted 31.12.18; published 26.03.19.

Please cite as:
Sadasivaiah S, Lyles CR, Kiyoi S, Wong P, Ratanawongsa N
Disparities in Patient-Reported Interest in Web-Based Patient Portals: Survey at an Urban Academic Safety-Net Hospital
J Med Internet Res 2019;21(3):e11421
URL: http://www.jmir.org/2019/3/e11421/
doi:10.2196/11421
PMID:30912747

XSL•FO
RenderX

©Shobha Sadasivaiah, Courtney R Lyles, Stephen Kiyoi, Piera Wong, Neda Ratanawongsa. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Impact of a Web-Based Course Concerning Patient Education for Mental Health Care Professionals: Quasi-Experimental Study

Anna Laine[1], RN, MNSc; Maritta Välimäki[1,2], PhD; Eliisa Löyttyniemi[3], MSc; Virve Pekurinen[1], RN, PhD; Mauri Marttunen[4,5], MD, PhD; Minna Anttila[1], PhD

[1]Department of Nursing Science, University of Turku, Turku, Finland

[2]School of Nursing, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]Department of Biostatistics, University of Turku, Turku, Finland

[4]University of Helsinki, Helsinki, Finland

[5]Helsinki University Hospital, Helsinki, Finland

Corresponding Author:
Maritta Välimäki, PhD
Department of Nursing Science
University of Turku
20014 University of Turku
Turku,
Finland
Phone: 358 406890546
Email: mava@utu.fi

Abstract

Background: Continuing education has an important role in supporting the competence of health care professionals. Although Web-based education is a growing business in various health sectors, few studies have been conducted in psychiatric settings to show its suitability in demanding work environments.

Objective: We aimed to describe the impact of a Web-based educational course to increase self-efficacy, self-esteem, and team climate of health care professionals. Possible advantages and disadvantages of the Web-based course are also described.

Methods: The study used nonrandomized, pre-post intervention design in 1 psychiatric hospital (3 wards). Health care professionals (n=33) were recruited. Self-efficacy, self-esteem, and team climate were measured at 3 assessment points (baseline, 8 weeks, and 6 months). Possible advantages and disadvantages were gathered with open-ended questions at the end of the course.

Results: Our results of this nonrandomized, pre-post intervention study showed that health care professionals (n=33) had higher self-efficacy after the course, and the difference was statistically significant (mean 30.16, SD 3.31 vs mean 31.77, SD 3.35; *P*=.02). On the other hand, no differences were found in the self-esteem or team climate of the health care professionals before and after the course. Health care professionals found the Web-based course useful in supporting their work and relationships with patients. The tight schedule of the Web-based course and challenges in recruiting patients to use the patient education program with health care professionals were found to be the disadvantages.

Conclusions: Web-based education might be a useful tool to improve the self-efficacy of health care professionals even in demanding work environments such as psychiatric hospitals. However, more studies with robust and sufficiently powered data are still needed.

(J Med Internet Res 2019;21(3):e11198) doi:10.2196/11198

KEYWORDS

internet; online education; mental health; continuing education

XSL•FO
RenderX

Introduction

Background

Mental health settings are demanding work environments for health care professionals [1,2]. Mental health care professionals in particular face challenging and ethically burdening situations in psychiatric hospitals [3,4]. To deal with these demanding situations, health care professionals need to work effectively [5] and be confident in their professional role [6]. Maintaining competence is an ethical responsibility of health care professionals [7,8], and the capacity to use information technology has been included as a core capacity in health care [9].

Continuing education is an important part of employability and personal development [10], and the costs of continuing education are remarkable in Europe [11]. The previous studies have found that education has an important role in health care professionals' feelings of competence [12,13]. Other studies have also reported association between education and work-related self-efficacy [12,14,15], general self-efficacy [16], self-esteem [17,18], knowledge level, skills [19], and personal development [20]. Furthermore, high competence [12,21] and personal development [16,22] have been found to have a positive impact on health care professionals' self-efficacy. For example, Proudfoot et al [23] found in their study that having less nonclinical staff was associated with better team climate. On the contrary, low self-esteem seems to be an explanatory variable when it comes to burnout among health care professionals [24].

Web-based learning has become a rapidly growing educational method [25]. Besides catering to different learning styles [26], it offers flexible [27] and effective education for health care professionals [28-31]. Web-based education requires the capacity to use a computer [32] and a positive attitude toward computers [33]. The impact of Web-based learning methods is still less often assessed in mental health care settings than it is in other health care settings [29]. Web-based learning has been used in practice, for example, to implement suicide prevention guidelines [34], to decrease the number of patient aggressive events and physical restrictions used in psychiatric hospitals [35], and to support health care professionals' engagement in personality difficulties [36]. On the other hand, the use of computers in supporting the knowledge level of health care professionals, changes in attitudes, and work habits has been doubted, not only regarding education but in mental health services altogether [37-39]. It has been questioned, for example, whether real learning requires face-to-face contact [30]. At the same time, a wide range of Web-based interventions have been developed for persons with special mental health needs [40-43]. A lack of computer skills [44,45], a lack of confidence in computer use [46], or the wide range of diverse Web-based mental health interventions [47] might increase concerns of the benefit of Web-based e-learning methods in continuing education.

Objectives

Earlier studies have shown that higher education [16,48], theory-based training [49], and communication skills training [50] have positive impacts on the self-efficacy of health care professionals. In addition, Web-based education has been found to have a positive effect on health care professionals' knowledge, attitudes, and practices [51]. However, there is still the need to evaluate the impacts of Web-based education [29]. Therefore, the purpose of this study was to describe the impact of a Web-based educational course on the self-efficacy, self-esteem, and team climate of health care professionals. Possible advantages and disadvantages of the Web-based course are also described. In this quasi-experimental study, we hypothesize that the Web-based course can be a beneficial and usable learning method to support health care professionals' self-efficacy and self-esteem and support a positive team climate among health care professionals. In addition, the health care professionals' perception of advantages and disadvantages of the course are described.

Methods

Design

A nonrandomized, pre-post intervention design was used. This design is suitable for our purpose as it is used to evaluate the benefits of specific interventions as preintervention and postintervention measurements with nonrandomly selected control groups [52]. A qualitative approach was also used to describe the health care professionals' feedback on the Web-based course [53].

Setting

The study was conducted on 3 closed psychiatric inpatient wards (a total of 41 beds) at 1 psychiatric hospital in southern Finland. The wards treated adult patients with serious mental disorders such as schizophrenia, which is the most common diagnosis for patients treated in Finnish psychiatric hospitals [54]. Moreover, 2 wards were rehabilitation wards and 1 was an acute ward.

Participants

Participants were health care professionals from the 3 wards working in multi-professional teams (altogether 46 health care professionals: 40 nurses, 3 psychologists, 1 occupational therapist, and 2 not known), which is essential in ensuring good patient care [55]. Nurses make up the largest professional group in current psychiatric health services in Finland. In 2014, there were altogether 258,567 people working in municipal health and social care including 46,446 nurses and 89,800 practical nurses or people working in similar occupations [56].

Inclusion criteria for health care professionals were that they worked in the selected wards, participated in the Web-based course targeted for the health care professionals, were Finnish-speaking, and were willing and able to participate in the study. Temporary workers in the ward were excluded.

Recruitment and the Follow-Up

An information meeting was held at the hospital concerning the Web-based course and the study, and the course was offered to all health care professionals (N=46) from the 3 study wards. All health care professionals were recruited to participate in the study after the meeting.

XSL•FO
RenderX

Table 1. The schedule of recruitment, intervention, and assessments.

Schedule of the study	Study period					
	Recruitment (March 2015)	Baseline (April 2015)	Post intervention			Closeout (6 months)
			Month 1	Month 2	8 weeks	
Eligibility screen	X[a]	N/A[b]	N/A	N/A	N/A	N/A
Informed consent	X	N/A	N/A	N/A	N/A	N/A
Web-based course	N/A	N/A	X	X	N/A	N/A
Background information	N/A	X	N/A	N/A	N/A	N/A
Self-esteem measurement	N/A	X	N/A	N/A	X	X
Self-efficacy measurement	N/A	X	N/A	N/A	X	X
Team climate measurement	N/A	X	N/A	N/A	X	X
Feedback of Web-based course	N/A	N/A	N/A	X	N/A	N/A

[a]X: applicable.

[b]N/A: not applicable.

Baseline data were collected in April 2015, and follow-ups were conducted after 8 weeks and 6 months using an online questionnaire. A link to the online questionnaire was sent to the health care professionals before the 3 measurement points. One researcher (VP) monitored data collection and sent reminder messages to the health care professionals. The timeline of participation is presented in Table 1 (modified based on the study by Chan et al [57]).

Intervention

Previous studies have shown that health care professionals have difficulties in supporting patients in computer use, especially in using Web-based patient education [39,58]. We assumed that if nurses were more familiar with online websites and Web-based education, it might encourage nurses to support patients' use of computers [59]. Therefore, a Web-based course about patient education was organized for health care professionals to enhance their capacity to support patients in Web-based education.

The Web-based course comprised 4 modules: (1) patient education—methods, effectiveness, and ways to put it into practice, (2) initializing of information technology in mental health care, (3) patient education in practice, and (4) practical training in patient education with a Web-based program. As an example of patient education, we used the website MentalNet, whose quality has been approved by the Health On the Net Code of Conduct to certify that it is a trustworthy and reliable medical website [60]. The MentalNet website has been developed for professionals, patients with psychosis (International Statistical Classification of Diseases and Related Health Problems, tenth revision, codes F20-F29 [61]), and their family members. It includes information about mental illness, treatment, well-being, patients' rights, daily life, links to other websites, a discussion forum, and a question and answer column [62]. More detailed information about the course is provided in Table 2 (modified based on Template for Intervention Description and Replication [63]).

Outcomes

In this study, self-reported instruments were chosen to be used because we were interested in getting the health care professionals' own perspective [67]. Self-reported instruments are useful when measuring respondents' personal opinions and beliefs and delivering questionnaires electronically without interpersonal contact [68].

Primary Outcome

Self-efficacy was measured with the General Self-Efficacy Scale (GSE). This instrument is designed to measure general self-efficacy as a person's capacity to trust his or her own ability to survive in new or difficult situations in life rather than situation-specific self-efficacy [69]. The instrument includes 10 questions (4-point Likert scale). The sum score of the answers ranges from 10 to 40; a lower score represents poorer coping on a daily basis. The instrument is widely used in studies concerning health care professionals [16,70,71]. Psychometric properties of the GSE have been examined in 25 countries; the Cronbach alpha has varied from .75 to .91 [72]. In this study, the Cronbach alpha was .82.

Secondary Outcomes

Self-Esteem

Self-esteem was measured using the Rosenberg Self-Esteem Scale (SES) [73]. The instrument is designed to measure global self-esteem. SES is widely used in educational [18] and other types of studies concerning health care professionals [24,74-76]. It includes 10 items with a 4-point Likert scale. The sum score of the answers ranges from 10 to 40. Having a higher score indicates a higher level of self-esteem. In a review by Schmitt and Allik [77], the data of the SES from 53 countries were compared. Internal consistency was found to be good (Cronbach alpha .80, range .45 to .90). In these data, the Cronbach alpha was .86.

XSL•FO
RenderX

Table 2. Description of the intervention.

Categories	Description of the intervention
Name: patient education in mental health work	Educational intervention for health care professionals in psychiatric hospitals to deliver patient education
Rationale and theory [64-66]	Health care professionals' education provides competence and skills for professionals to deliver patient education. Patient education is an important component of care of people with severe mental health disorders. It enables patients to cope more effectively with the illness, which has a positive effect on their well-being. Educational intervention focuses on professional knowledge transfer and health care professionals' behavioral changes. It provides reasoning for patient education and Web-based services, knowledge of the effectiveness of patient education, and methods to deliver it
Materials	Information about evidence-based research, a written information package, and MentalNet website usable with user account and password (also accessible via an URL link)
Procedures	Individual and group studying and exercises based on health care professionals' own time schedules; individual and interactive exercises, including a reflective diary and discussions on the forums; and delivery of 5 practical patient education exercise sessions with a patient using the MentalNet website in which patients had their own user account and password
Providers	Trained teachers, tutors, and researchers with a background in mental health provided feedback to the health care professionals after the completed exercises and answered any emerging questions. A course coordinator answered emerging questions and informed about the phases of the course
How	PowerPoint presentations, reading material, and a manual on how to proceed with patient education in clinical practice
Where	On the Moodle learning portal hosted by the University of Turku, usable with personal user accounts and passwords
When and how much	The course consisted of 5 phases (3 European Credit Transfer and Accumulation System credits, 72 hours), and the length of the course was 2 months
Tailoring and modifications	The course was held over a 9-month period due to vacations and health care professionals working in 3 shifts in the hospital

Team Climate

Team climate was measured with the Team Climate Inventory (TCI) [78] Finnish version [79]. The TCI is based on West's theory of innovation [80], and it includes 4 subscales with a varying number of items. The 4 subscales are (1) participative safety (12 items), (2) support for innovation (8 items), (3) vision (10 items), and (4) task orientation (8 items). The first 2 subscales are evaluated with a 5-point Likert scale and the next 2 subscales with a 7-point scale. The TCI has been found to be a reliable measurement. In a previous study involving health care professionals in Finnish mental health care, the Cronbach alpha of the Finnish version of the TCI was as follows: (1) participative safety: .88, (2) support for innovation: .87, (3) vision: .95, and (4) task orientation: .91 [81]. In this study's data, the Cronbach alpha values were as follows: (1) participative safety: .91, (2) support for innovation: .89, (3) vision: .95, and (4) task orientation: .92.

Advantages and Disadvantages of the Web-Based Course for Health Care Professionals

Health care professionals' perceptions about advantages and disadvantages of the Web-based course were collected via a questionnaire in the course platform by using the following open questions: (1) how does the Web-based course support the professional skills of mental health care professionals? and (2) what were the advantages and disadvantages of the course?

Background Information

Information about the health care professionals' gender, age, level of work experience, and information about their internet use, purposes, skills, and attitudes were collected. Health care professionals' internet use was described using an adapted measure by Choi and DiNitto [82]: "Have you ever used the internet?" (1) "No, I have never used it" (never user), (2) "I have used it before but not currently" (previous user), and (3) "Yes, I am a current user" (current user). Furthermore, the purpose of the health care professionals' internet use was explored. Health care professionals were asked to select what kind of activities they conducted on the internet. Options included research health-related information; research information about other topics or issues of interest; send/receive email; buy products online; conduct banking online and/or pay bills; read news, papers, magazines, and books online; play games online; watch videos (including YouTube); use social networking or dating sites (eg, Facebook and Match.com); and other (they were asked to specify) [82]. Activities in which a health care worker would "communicate with health professionals about health-related issues" or "communicate with other users about health-related issues" were added to the original list [83]. Internet/computer skills were evaluated in an item with a 5-point Likert scale: "Your computer/internet-skills are" (1=very good to 5=very poor), whereas attitudes against the computer/internet were examined using an item with another 5-point Likert scale (1=very positive to 5=very negative).

XSL•FO
RenderX

Sampling and Recruitment of the Participants

Consecutive sampling was used to recruit all the health care professionals who worked in these 3 wards and who participated in the Web-based e-course targeted for the health care professionals. This sampling method was suitable for our quasi-experimental study, as we aimed to recruit all possible study participants accessible at the time of data collection [53].

Data Analysis

Numerical variables are summarized with a median, mean, and SD, whereas categorical variables are reported with counts and percentages. The total sums for scores were calculated. Self-efficacy, self-esteem, and factors of team climate were all analyzed with hierarchical linear mixed models for repeated measures for all 3 time points (baseline, 8 weeks, and 6 months). The model was adjusted for age, gender, and work experience. A compound symmetry covariance structure was used for repeated measures. All statistical tests were performed as two-sided with the statistical significance level set at .05. The analyses were performed using the SAS System, version 9.4 for Windows (SAS Institute Inc, Cary, NC, US).

Cohen *d* was assessed at baseline and 6 months to determine the effect size of the intervention on measured outcomes. An effect size of 0.2 is considered small, 0.5 is medium, and 0.8 is a great effect size [84]. An online calculator was used for the analyses [85].

A qualitative analysis was conducted by 2 authors (AL and MA). Writings of health care professionals were collected from the Moodle learning portal. Similarities in the answers were searched for based on the study questions and were classified into categories [86]. The texts answering the question asking how the Web-based course supports professional skills were combined with the answers about advantages of the course because the answers were very similar.

Ethics

Ethical permission for the research was obtained from the ethics committee of the Hospital District of Southwest Finland (ETMK:40/1801/2015). Permission for data collection was obtained from the research permission committee of the city (2015-002638).

After providing oral and written information about the study, an information letter and informed consent forms were sent via email to the health care professionals of the wards. The health care professionals were asked to print 2 copies and sign them, after having approximately 2 weeks to decide if they wanted to participate in the study or not. Of them, 1 researcher (AL) collected 1 signed consent form from each health care professional when visiting the wards and left the duplicate with the health care professional. Participation was voluntary, and withdrawal was allowed at any time without needing to give a reason. The data were handled by a research group. It used anonymous ID codes [87].

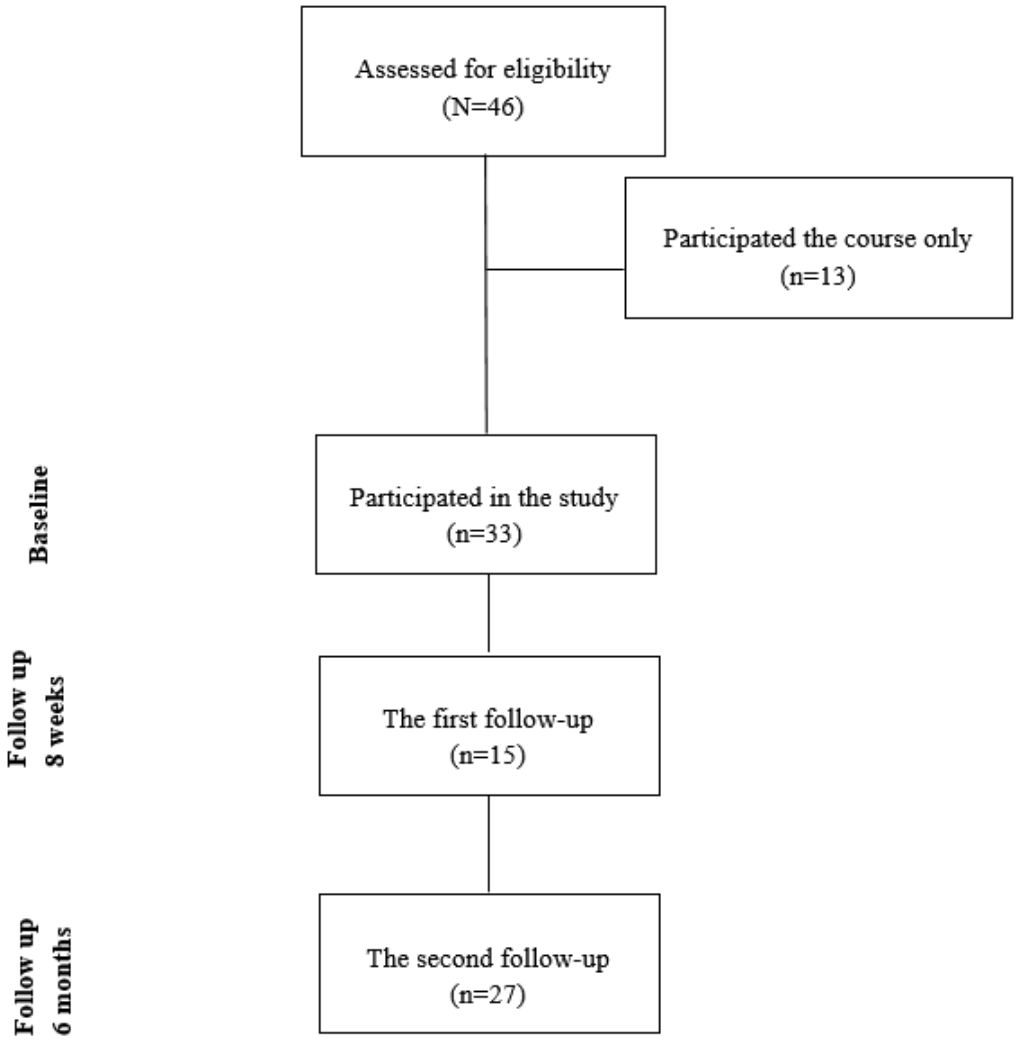
Results

Participant Flow

Out of 46 possible Web-based course participants, 13 (28%) health care professionals decided to participate only in the Web-based course and 33 (72%) participated both in the study and online course. Out of them, 15 (45%) participated in the first follow-up and 27 (82%) in the second follow-up. A description of the flow diagram of the participants is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Flow diagram of the participants.



Description of Participants

The study sample consisted of 33 health care professionals, 24 of them were females and 9 were males. Their mean age was 41 years (range 25-61), and their mean working experience in mental health care was 14 years (range 2-40). All health care professionals used the internet (33/33, 100%). Over half (21/33, 64%) estimated that they have very good or good computer skills and two-thirds (27/33, 82%) had very positive or positive attitudes toward the internet. Health care professionals chose all options for internet use that were related to them. They used the internet for banking, emailing, searching for health care knowledge, and reading news or books (Table 3). In addition, 4 health care professionals also used the internet for music streaming, searching for information on public transportation, discussion purposes, or for work.

The Impact of the Web-Based Course on Health Care Professionals' Self-Efficacy, Self-Esteem, and Team Climate

The impact of the Web-based course on health care professionals' self-efficacy, self-esteem, and team climate was measured at 3 time points. The analysis showed that health care professionals' self-efficacy scores increased between baseline (mean 30.16, SD 3.31) and 8 weeks (mean 31.53, SD 2.82), and even further at 6 months (mean 31.77, SD 3.35). The difference between baseline and 6 months measurements was statistically significant (P=.02) with an effect size of 0.48 (Cohen d).

On the contrary, no statistical differences were found in health care professionals' self-esteem during the 6-month study period. In addition, no significant changes were found in health care professionals' team climates or its subscales (participative safety: P=.82, support for innovation: P=.81, vision: P=.48, and task orientation: P=.85; see Table 4).

XSL•FO
RenderX

Table 3. Internet skills, attitudes toward the internet, and internet use of health care professionals.

Characteristic (n=33)	n (%)
Computer/internet skills	
Very good	7 (21)
Good	14 (42)
Neither good nor bad	9 (27)
Fairly poor	2 (6)
Poor	1 (3)
Attitudes toward computers/internet	
Very positive	13 (39)
Positive	14 (42)
Neither positive nor negative	5 (15)
Negative	1 (3)
Very negative	0 (0)
Purpose of internet use	
For banking online and/or to pay bills	33 (100)
Send/receive email	32 (97)
Research health-related information	30 (91)
Read news, papers, magazines, and books online	30 (91)
Research information about other topics of interest	28 (85)
Buy products online	24 (73)
Watch videos	22 (67)
Use social networking or dating sites	19 (58)
Play games online	7 (21)
Communicate with health professionals about health-related issues	5 (15)
Other	4 (12)
Communicate with other users about health-related issues	3 (7)

Table 4. Hierarchical linear mixed models of self-efficacy, self-esteem, and factors of team climate at baseline, 8 weeks, and 6 months.

Measurement	Time point						*F* (*df*)	*P* value	Cohen *d*
	Baseline		8 weeks		6 months				
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)			
Self-efficacy	32	30.16 (3.31)	15	31.53 (2.83)	26	31.77 (3.35)	4.59 (2)	.02[a]	0.48
Self-esteem	33	34.00 (4.62)	15	33.80 (4.93)	26	35.54 (4.01)	1.55 (2)	.23	0.36
Team climate									
Participative safety	33	3.99 (0.57)	15	3.97 (0.46)	27	3.96 (0.46)	0.19 (2)	.82	−0.06
Support for innovation	32	3.50 (0.59)	15	3.47 (0.66)	27	3.57 (0.57)	0.21 (2)	.81	0.12
Vision	33	5.41 (0.90)	15	5.10 (0.68)	26	5.38 (0.89)	0.74 (2)	.48	−0.03
Task orientation	33	5.06 (0.95)	15	4.87 (0.72)	26	4.96 (0.91)	0.16 (2)	.85	−0.01

[a]Statistically significant difference analyzed with hierarchical linear mixed models for repeated measures.

The demographic characteristics of nonparticipating and participating health care professionals in the first follow-up were similar. The mean age of the nonparticipating health care professionals was 41 years (range 30-61), and participating health care professionals had a mean age of 38 years (range 25-60). The mean number of working years was 15 (range 2-38) for nonparticipating health care professionals and 14 years (range 2-40) for participating health care professionals. Out of

XSL•FO
RenderX

the 18 nonparticipating health care professionals, 4 (22%) were men, and out of the 15 participating health care professionals, 5 (33%) were men.

Advantages and Disadvantages of the Web-Based Course

Advantages of the Web-Based Course

On the basis of the health care professionals' descriptions, the advantages of the course were divided into 5 categories: (1) a modern method for patient education, (2) support for health care professionals' competence, (3) support for discussions between a health care professional and a patient, (4) support for relations between a health care professional and a patient, and (5) support for structural nursing. The advantages and examples of phrases from health care professionals are presented in Table 5.

First, in general, the use of computers was described as a novel way to administer patient care in ward settings. Health care professionals found that the Web-based resources were a modern approach to carrying out patient education, and the website was a new tool they could use for it:

> *MentalNet opens a new and modern way to realize patient education.*
>
> *The staff have gained one new way for implementing psychoeducation.*
>
> *The staff have received a new tool for themselves.*

Second, health care professionals described that the Web-based course supported their competence level. They were able to find current information, recall old information, and also learn something new when they used the website together with patients. The course also supported professional skills when health care professionals were able to learn how to use a computer in patient care:

> *The latest versatile information needed in caring for a patient with psychosis can be found on the website.*
>
> *While going through MentalNet patient education, I got new information about the latest recommendations myself.*
>
> *The staff have benefited from the repetition of the psychoeducation content.*
>
> *Nurses also learn to utilize information technology*

Third, the Web-based course supported discussions between health care professionals and patients. Health care professionals considered the themes of the website to be starting points for discussions. The themes were also seen as a checklist when all important topics were discussed through them. Health care professionals described how using the website in discussions helped them to bring up different topics. Patient education meetings were found to be natural moments to also bring up those topics that were otherwise difficult to discuss with the patients:

> *The program would be a good foundation for psychoeducation. The topics are categorized, so you can address them according to the patient's interests.*
>
> *Themes are collected as a clear checklist, and it is always possible to go back to something that a patient is wondering about.*
>
> *There are topics that would have not necessarily been worked through with a patient otherwise (e.g. sexuality, patient's rights).*
>
> *Patient education sessions included many natural situations and moments to bring up things in different themes.*

Fourth, patient education helped to create, support, and strengthen relationships between patients and health care professionals. Health care professionals described that using the website together with the patient helped them to have good contact with the patients and create relationships with the patients. After using the website, it was also easier to recognize the patients' needs and understand their symptoms:

> *The use of MentalNet may make it easier to create a good patient-nurse relationship.*
>
> *There has been an opportunity to obtain good contact with the patient.*
>
> *MentalNet provides a foundation for staff to better understand patients' symptoms, and it makes patients' situations easier to identify with.*

Fifth, the course was found to support structural nursing and the planning aspect of patient education. Health care professionals found that using the website made patient education more systematic. It also strengthened the quality of the patient education when all patients got the same information regardless of which ward or health care professional provided it:

> *The adoption of patient education has brought more planning.*
>
> *In addition, it gives staff an opportunity to deal with aspects concerning patients' health and wellbeing in an even more systematic manner.*
>
> *All patients get similar information regardless of the ward or nurse, which strengthens the quality of the care.*

XSL•FO
RenderX

Table 5. Advantages of the Web-based course.

Categories and subcategories	Phrases
A modern method for patient education	
Modern way to realize patient education	"The staff have gained one new way for implementing psychoeducation."; "MentalNet opens a new and modern way to realize patient education."; and "MentalNet is a good new method that can be used to provide and maintain patient education to patients of the ward."
New tool to realize patient education	"MentalNet has provided a new instrument for implementing patient education to patients treated in the ward."; "The staff have gained a new tool for implementing psychoeducation."; and "The staff have received a new tool for themselves."
Support for health care professionals' competence	
Current information for health care professionals	"The site provides current information to staff in the same manner as it does to patients."; "The latest versatile information needed in caring for a patient with psychosis can be found on the website."; and "Information is collected into themes, which can be easily accessed with a link."
Develop and maintain health care professionals' knowledge	"The theme may even expose nurses to new information also."; "MentalNet also provides the staff with a lot of well-researched and up-to-date information on mental health issues, which will help to maintain good professional skills."; and "While going through MentalNet patient education, I got new information about the latest recommendations myself."
Repetition of knowledge for health care professionals	"The content resembles things that have already been learned."; "The staff have benefited from the repetition of the psychoeducation content."; and "By using MentalNet, it is also possible to recall up to date and trustworthy information, and to possibly gain some new information."
Support for health care professionals' skills in technology	"Nurses also learn to utilize information technology" and "It has given us experience in how computers can be used in patient care."
Support for discussions between a health care professional and a patient	
Base for discussions with a patient	"MentalNet can be used as a foundation for discussion between nurses and patients."; "The program would be a good foundation for psychoeducation. The topics are categorized, so you can address them according to the patient's interests."; and "I believe that MentalNet can be a useful tool for patient education and as a basis for discussions between nurses and patients."
Checklist for discussions with the patient	"Themes are collected as a clear checklist, and it is always possible to go back to something that a patient is wondering about."; "Providing psychoeducational information has become easier; all the details do not have to be memorized by heart and the information is structured in MentalNet."; and "Themes regarding psychosis are clearly grouped together, and they serve as a good checklist for the professional."
Helps to bring various themes into discussions	"There are topics that would have not necessarily been worked through with a patient otherwise (eg, sexuality and patient's rights)."; "I personally feel that, with the program, it was easier to bring up things with the patient."; and "It encourages the discussion of health and wellbeing-related topics with the patient: For example, the importance of sleep, nutrition and exercise to wellbeing are not usually covered in patient education as thoroughly as it is in medical treatment and psychosis. Also, it makes it easier to bring up topics and patient system dimensions that are not needed at the moment."
Natural moments to discuss different topics	"Patient education meetings were natural moments to bring up different topics with the patient."; "Patient education sessions included many natural situations and moments to bring up things in different themes."; and "During the patient education, discussing the most difficult things were easier and more natural for the patient."
Support for relations between a health care professional and a patient	
Help for a good relationship with a patient	"The use of MentalNet may make it easier to create a good patient-nurse relationship."; "It has helped to create a good nurse-patient relationship."; and "For example, creating a new framework for being connected with the patient/patients."

XSL•FO
RenderX

Categories and subcategories	Phrases
Strengthens interaction between a health care professional and a patient	"Using MentalNet created a good moment of interaction with a patient."; "Strengthens interaction with patients."; and "There has been an opportunity to obtain good contact with the patient."
Increase of understanding of patients' needs and symptoms	"Recognizing patients' needs is easier." and "MentalNet provides a foundation for staff to better understand patients' symptoms, and it makes patients' situations easier to identify with."
Support for structural nursing	
Planning and systematic increase of patient education	"The adoption of patient education has brought more planning."; "Patients are dealt with more systematically."; and "In addition, it gives staff an opportunity to deal with aspects concerning patients' health and wellbeing in an even more systematic manner."
Same information for all patients	"All patients get similar information regardless of the ward or nurse, which strengthens the quality of the care." and "The content stays the same even if the patient changes."

Disadvantages of the Web-Based Course

Disadvantages of the course consisted of 3 categories: (1) factors concerning schedule and working time, (2) factors concerning patients, and (3) factors concerning equipment and environment. Disadvantages and examples of phrases from health care professionals are presented in Table 6.

First, health care professionals described disadvantages concerning schedules and working times. Tight schedules and the hurried implementation of the course were found to be challenging and caused stress. Completing the course and scheduling the patient education meetings were challenging because of shift work and vacation days during the course:

> *The busy timetable for implementing the course was challenging and caused stress to at least some of the personnel.*
>
> *Scheduling patient education meetings was challenging because of shift work.*
>
> *The schedule has been tight, and it was during a period when some of the employees still had winter holidays and some were about to start their summer holidays. This has created some challenges in the implementation.*

Second, health care professionals described disadvantages concerning patients. Health care professionals had difficulties finding willing and suitable patients to participate in patient education. Some of the patients also had a lack of computer skills, which complicated the patient education:

> *There was pressure to hold education meetings and to find willing patients with the cognitive skills to understand the content.*
>
> *Also, however, using a computer is often difficult for patients. Therefore, we sometimes have needed to start with how to use the computer.*

Third, factors concerning the equipment and the environment were considered to be a disadvantage, as health care professionals described the fact that there were only a few computers that could be used for Web-based patient education. Moreover, when health care professionals needed to book a room for himself/herself and a patient, other health care professionals and patients could not use the room during that time:

> *There are only a few computers available for patient education.*
>
> *When one nurse stays an hour in a patient education meeting and a group room is occupied for that hour, others have to adapt their work and schedule around the meeting accordingly.*

XSL•FO
RenderX

Table 6. Disadvantages and examples of phrases from health care professionals.

Categories	Subcategories	Phrases
Factors concerning schedules and working times	Tight schedule of the course	"The busy timetable for the implementation of the course was challenging and caused stress to at least some of the personnel."; "The schedule of the course was pretty tight, which created challenges in the implementation."; and "The schedule of this course has added negative pressure to the work."
	Challenges concerning working times and the course	"Scheduling patient education meetings was challenging because of shift work."; "It was challenging to fit schedules to one's own working time."; and "The schedule has been tight, and it was during a period when some of the employees still had winter holidays and some were about to start their summer holidays. This has created some challenges in the implementation."
Factors concerning patients	Patients' unwillingness to participate in patient education	"There was pressure to hold education meetings and to find willing patients with the cognitive skills to understand the content."; "Some were stressed by patients' low interest in taking part in the MentalNet patient education."; and "The drawback for some staff has been patients who are not willing to take part in education/research."
	Patients' poor computer skills	"Also, however, using a computer is often difficult for patients. Therefore, we sometimes have needed to start with how to use the computer."
Factors concerning equipment and environment	Room booked during the patient education	"When one nurse stays an hour in patient education meeting and a group room is occupied for that hour, others have to adapt their work and schedule around the meeting accordingly."
	Too few computers available	"There are only a few computers available for patient education."

Discussion

Principal Findings

The aim of our small-scale study was to find out how a Web-based course impacts health care professionals' self-efficacy, self-esteem, and team climate. The results show that health care professionals who participated in the Web-based course focusing on patient education had an increased sense of self-efficacy at baseline and after the course, and again after 6 months. These findings support earlier studies that found that the self-efficacy of health care professionals could be supported by education. Other studies have shown that health care professionals who start their career with a higher educational level seem to have higher self-efficacy than health care professionals with lower education [16,48]. Similar to our intervention, some studies have shown that continuing education for health care professionals with theory-based training [49] and communication training [50] increases health care professionals' self-efficacy.

In contrast to previous findings, we did not find any significant difference between health care professionals' self-esteem at different time points. On the basis of a study by Van Eckert et al [18], nurses with an academic education have higher self-esteem than nurses without an academic education. Therefore, it is possible that this kind of short course might not be powerful enough to support self-esteem. Furthermore, no changes were found between health care professionals' sense of team climate during the course of follow-up. We assumed that team climate would be stronger after the course because being in close cooperation and reflecting on learning experiences together might positively affect team climate [88], and teamwork can increase a sense of belonging [89]. Having a sense of a good team climate is important for nurses' well-being [90], and this might, in turn, affect patients' satisfaction levels with care [23]. It is still unclear whether a good team climate can be developed by using technological solutions. Although team climate for health care professionals is an important goal in health services, Web-based learning might not be the prioritized method for this task. However, expectations toward the adoption of technology in health care are positive [9].

In our study, the health care professionals found the Web-based course to be diversely useful. According to their feedback, the Web-based course supported the structure of patient education. Health care professionals were able to utilize the content of the MentalNet website as *a checklist* in conversations with the patients. They also found the website helpful in supporting their professional skills and the structure of nursing activities. Some health care professionals thought that MentalNet supported their relationships with patients because it provided evidence-based knowledge that could be used during challenging discussion topics, such as sexuality. This feedback is useful information for improving relationships between health care professionals and patients on a larger scale—something that has been found to be crucial for treatment adherence, for example [91]. Patient education is an essential part of care for patients with schizophrenia [92]. Therefore, it is important to increase health care professionals' skills in delivering patient education.

On the other hand, health care professionals found some disadvantages of the patient education course. The disadvantages were mainly related to the scheduling problems of the course when most of the health care professionals worked in shifts and some of them were on vacation. In the future, it would be important to pay even more attention to planning the course to minimize these kinds of disadvantages.

Currently, continuing education is a vital part of the European Commission's lifelong learning policy [10], and the costs of continuing vocational education are significant in European Union countries [11]. In addition, Finnish health and social services put a strong emphasis on continuing education. For example, in 2015, 69% of employees in social and health care

XSL•FO
RenderX

(N=202,413) participated in continuing education with total costs of 48 million euros [93]. The costs of continuing education are based on, for example, travel costs and payment for substitutes for employees while they are studying. If at least some of the costs could be saved, by removing travel costs through Web-based education, for example, the savings could be millions per year for employers. In the future, Web-based courses will become increasingly important as younger health care professionals, who are accustomed to using computers and other information technology as a part of their daily activities, enter into the workforce.

Limitations

The results of this study should be considered in the context of its limitations and strengths. First, the number of health care professionals was small, and measurements were done with a pre-post design without a control group, which limits the generalizability of the results. The original plan was to have a control group of other health care professionals from similar wards, but the number of health care professionals in the control group would have been very low; therefore, we decided to use data from the intervention group only. However, a quasi-experimental study design without a control group has been found to be a suitable method for medical informatics studies. Moreover, the design, with its pretest and multiple posttests, increases the validity of the method [52]. Second, dropout rates were high throughout the study, especially for the first follow-up measurement (17/33, 52%). This first follow-up measurement was conducted 8 weeks after baseline when the course was supposed to be completed. However, health care professionals had difficulties in recruiting patients to use MentalNet with them for the course. Therefore, after 8 weeks, the length of the course was delayed, which might have had an impact on measurements. However, the demographic characteristics of nonparticipating and participating health care professionals in the first follow-ups were similar. Therefore, although the dropouts in follow-ups can be assumed to be random, we can also assume that missing values were completely random and the assumptions for linear mixed models for repeated measures analyses were met.

Third, self-reported measurements were used to assess the impact of the course. More objective measurements could offer a more valid perspective on the impact of the course. For example, only the health care professionals were asked about the advantages and disadvantages of the Web-based course. The perspective of other stakeholders, such as patients [34], could widen the understanding of the real impact of the course in clinical practice [94]. In addition, we may ask whether our outcome measures were specific enough to measure the impact of the Web-based course for health care professionals. More studies are needed in this area.

Despite the study design used and the small sample size, we might still assume that health care professionals' Web-based training has a positive impact on health care professionals' self-efficacy. This finding is significant for health care leaders when evaluating the importance of education in the professional performance of health care workers [12,13], especially when considering Web-based education or other learning methods involving technological devices.

Conclusions

Our study provides essential information on how a Web-based course can be useful to health care professionals as continuing education. The results of this study show that self-efficacy of health care professionals can be supported by Web-based continuing education. In addition, this kind of study, where the study task is for health care professionals to use a Web-based program with their patients, can improve the relationships between health care professionals and patients. The results of this study can be utilized when planning a Web-based course and trying to avoid difficulties with scheduling the course and recruiting the patients, as we experienced here. When offering a course in which health care professionals need to work with patients with schizophrenia, there should be enough time to recruit patients. In addition, shift work and vacations can cause challenges in maintaining the course. Therefore, an adequate amount of consideration should be spent on scheduling the course.

Acknowledgments

The authors would like to thank the study hospital, health care professionals of the wards, and the patients they educated.

Conflicts of Interest

None declared.

References

1. Jenkins R, Elliott P. Stressors, burnout and social support: nurses in acute mental health settings. J Adv Nurs 2004 Dec;48(6):622-631. [doi: 10.1111/j.1365-2648.2004.03240.x] [Medline: 15548253]
2. Dickinson T, Wright KM. Stress and burnout in forensic mental health nursing: a literature review. Br J Nurs 2008;17(2):82-87. [doi: 10.12968/bjon.2008.17.2.28133] [Medline: 18414278]
3. Foster C, Bowers L, Nijman H. Aggressive behaviour on acute psychiatric wards: prevalence, severity and management. J Adv Nurs 2007 Apr;58(2):140-149. [doi: 10.1111/j.1365-2648.2007.04169.x] [Medline: 17445017]
4. Lantta T, Anttila M, Kontio R, Adams CE, Välimäki M. Violent events, ward climate and ideas for violence prevention among nurses in psychiatric wards: a focus group study. Int J Ment Health Syst 2016;10:27 [FREE Full text] [doi: 10.1186/s13033-016-0059-5] [Medline: 27051463]

XSL•FO
RenderX

5. d'Ettorre G, Pellicani V. Workplace violence toward mental healthcare workers employed in psychiatric wards. Saf Health Work 2017 Dec;8(4):337-342 [FREE Full text] [doi: 10.1016/j.shaw.2017.01.004] [Medline: 29276631]
6. Kaddoura MA. New graduate nurses' perceptions of the effects of clinical simulation on their critical thinking, learning, and confidence. J Contin Educ Nurs 2010 Nov;41(11):506-516. [doi: 10.3928/00220124-20100701-02] [Medline: 20672760]
7. International Council of Nurses. The ICN Code of Ethics for Nurses. Geneva, Switzerland: Norsk Sykepleierforbund; 2000. URL: https://www.nsf.no/Content/113948/icncode.pdf [WebCite Cache ID 6zmjrbePn]
8. Finnish Nurses Association. Ethical Guidelines of Nursing.: Sairaanhoitajat; 2014. URL: https://sairaanhoitajat.fi/artikkeli/ethical-guidelines-nursing/ [WebCite Cache ID 6zm5AWUvX]
9. Shaw T, Blake R, Hübner U, Anderson A, Wangia-Anderson V, Elias B. Healthcare Information and Management Systems Society. 2017. The Evolution of TIGER Competencies and Informatics Resources Executive Supplemental Report URL: http://www.himss.org/library/evolution-tiger-competencies-and-informatics-resources [accessed 2018-05-29] [WebCite Cache ID 6zmFSmgUc]
10. European Commission. 2018. Strategic framework - Education & Training 2020 URL: http://ec.europa.eu/education/policy/strategic-framework_en [accessed 2018-05-29] [WebCite Cache ID 6zm4bFx1K]
11. Vocational education and training statistics. 2017. URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/Vocational_education_and_training_statistics [WebCite Cache ID 6zm4xIhiK]
12. Lee TW, Ko YK. Effects of self-efficacy, affectivity and collective efficacy on nursing performance of hospital nurses. J Adv Nurs 2010 Apr;66(4):839-848. [doi: 10.1111/j.1365-2648.2009.05244.x] [Medline: 20423371]
13. Zhu DQ, Norman IJ, While AE. Nurses' self-efficacy and practices relating to weight management of adult patients: a path analysis. Int J Behav Nutr Phys Act 2013 Dec 04;10:131 [FREE Full text] [doi: 10.1186/1479-5868-10-131] [Medline: 24304903]
14. Doyle D, Copeland HL, Bush D, Stein L, Thompson S. A course for nurses to handle difficult communication situations. A randomized controlled trial of impact on self-efficacy and performance. Patient Educ Couns 2011 Jan;82(1):100-109. [doi: 10.1016/j.pec.2010.02.013] [Medline: 20303230]
15. Hsu LL, Huang YH, Hsieh SI. The effects of scenario-based communication training on nurses' communication competence and self-efficacy and myocardial infarction knowledge. Patient Educ Couns 2014 Jun;95(3):356-364. [doi: 10.1016/j.pec.2014.03.010] [Medline: 24718019]
16. Soudagar S, Rambod M, Beheshtipour N. Factors associated with nurses' self-efficacy in clinical setting in Iran, 2013. Iran J Nurs Midwifery Res 2015;20(2):226-231 [FREE Full text] [Medline: 25878701]
17. Hagbaghery MA, Salsali M, Ahmadi F. The factors facilitating and inhibiting effective clinical decision-making in nursing: a qualitative study. BMC Nurs 2004 Apr 06;3(1):2 [FREE Full text] [doi: 10.1186/1472-6955-3-2] [Medline: 15068484]
18. van Eckert S, Gaidys U, Martin CR. Self-esteem among German nurses: does academic education make a difference? J Psychiatr Ment Health Nurs 2012 Dec;19(10):903-910. [doi: 10.1111/j.1365-2850.2011.01862.x] [Medline: 22296591]
19. Hart P, Eaton L, Buckner M, Morrow BN, Barrett DT, Fraser DD, et al. Effectiveness of a computer-based educational program on nurses' knowledge, attitude, and skill level related to evidence-based practice. Worldviews Evid Based Nurs 2008;5(2):75-84. [doi: 10.1111/j.1741-6787.2008.00123.x] [Medline: 18559020]
20. Cleary M, Horsfall J, O'Hara-Aarons M, Jackson D, Hunt GE. The views of mental health nurses on continuing professional development. J Clin Nurs 2011 Dec;20(23-24):3561-3566. [doi: 10.1111/j.1365-2702.2011.03745.x] [Medline: 21722221]
21. Tsai CW, Tsai SH, Chen YY, Lee WL. A study of nursing competency, career self-efficacy and professional commitment among nurses in Taiwan. Contemp Nurse 2014;49:96-102. [doi: 10.5172/conu.2014.49.96] [Medline: 25549750]
22. van Dyk J, Siedlecki SL, Fitzpatrick JJ. Frontline nurse managers' confidence and self-efficacy. J Nurs Manag 2016 May;24(4):533-539. [doi: 10.1111/jonm.12355] [Medline: 26762223]
23. Proudfoot J, Jayasinghe UW, Holton C, Grimm J, Bubner T, Amoroso C, et al. Team climate for innovation: what difference does it make in general practice? Int J Qual Health Care 2007 Jun;19(3):164-169. [doi: 10.1093/intqhc/mzm005] [Medline: 17337517]
24. Molero Jurado MD, Pérez-Fuentes MD, Gázquez Linares JJ, Barragán Martín AB. Burnout in health professionals according to their self-esteem, social support and empathy profile. Front Psychol 2018;9:424 [FREE Full text] [doi: 10.3389/fpsyg.2018.00424] [Medline: 29731725]
25. Docebo. 2014. E-Learning Market Trends & Forecast 2014 - 2016 Report URL: https://www.docebo.com/landing/contactform/elearning-market-trends-and-forecast-2014-2016-docebo-report.pdf [accessed 2018-05-29] [WebCite Cache ID 6zmjZAbVP]
26. Laschinger HK, Borgogni L, Consiglio C, Read E. The effects of authentic leadership, six areas of worklife, and occupational coping self-efficacy on new graduate nurses' burnout and mental health: a cross-sectional study. Int J Nurs Stud 2015 Jun;52(6):1080-1089. [doi: 10.1016/j.ijnurstu.2015.03.002] [Medline: 25801311]
27. Mayes P, Schott-Baer D. Professional development for night shift nurses. J Contin Educ Nurs 2010 Jan;41(1):17-22. [doi: 10.3928/00220124-20091222-05] [Medline: 20102138]
28. Khatony A, Nayery ND, Ahmadi F, Haghani H, Vehvilainen-Julkunen K. The effectiveness of web-based and face-to-face continuing education methods on nurses' knowledge about AIDS: a comparative study. BMC Med Educ 2009;9:41 [FREE Full text] [doi: 10.1186/1472-6920-9-41] [Medline: 19591678]

XSL•FO
RenderX

29. Lahti M, Hätönen H, Välimäki M. Impact of e-learning on nurses' and student nurses knowledge, skills, and satisfaction: a systematic review and meta-analysis. Int J Nurs Stud 2014 Jan;51(1):136-149. [doi: 10.1016/j.ijnurstu.2012.12.017] [Medline: 23384695]
30. McCutcheon K, Lohan M, Traynor M, Martin D. A systematic review evaluating the impact of online or blended learning vs. face-to-face learning of clinical skills in undergraduate nurse education. J Adv Nurs 2015 Feb;71(2):255-270. [doi: 10.1111/jan.12509] [Medline: 25134985]
31. Ramos-Morcillo AJ, Fernández-Salazar S, Ruzafa-Martínez M, Del-Pino-Casado R. Effectiveness of a brief, basic evidence-based practice course for clinical nurses. Worldviews Evid Based Nurs 2015 Aug;12(4):199-207. [doi: 10.1111/wvn.12103] [Medline: 26220505]
32. Moule P, Ward R, Lockyer L. Nursing and healthcare students' experiences and use of e-learning in higher education. J Adv Nurs 2010 Dec;66(12):2785-2795. [doi: 10.1111/j.1365-2648.2010.05453.x] [Medline: 20946565]
33. Chong MC, Francis K, Cooper S, Abdullah KL, Hmwe NT, Sohod S. Access to, interest in and attitude toward e-learning for continuous education among Malaysian nurses. Nurse Educ Today 2016 Jan;36:370-374. [doi: 10.1016/j.nedt.2015.09.011] [Medline: 26455411]
34. de Beurs DP, de Groot MH, de Keijser J, van Duijn E, de Winter RF, Kerkhof AJ. Evaluation of benefit to patients of training mental health professionals in suicide guidelines: cluster randomised trial. Br J Psychiatry 2016 Dec;208(5):477-483. [doi: 10.1192/bjp.bp.114.156208] [Medline: 26678866]
35. Kontio R, Pitkänen A, Joffe G, Katajisto J, Välimäki M. eLearning course may shorten the duration of mechanical restraint among psychiatric inpatients: a cluster-randomized trial. Nord J Psychiatry 2014 Oct;68(7):443-449. [doi: 10.3109/08039488.2013.855254] [Medline: 24274836]
36. Clarke M, Jinks M, McMurran M. Readiness Enhancement Management Strategies (REMS): a proof of concept and evaluation feasibility study of staff training to improve service engagement by people with personality difficulties. J Psychiatr Ment Health Nurs 2015 Sep;22(7):502-509. [doi: 10.1111/jpm.12214] [Medline: 25977077]
37. Kuosmanen L, Jakobsson T, Hyttinen J, Koivunen M, Välimäki M. Usability evaluation of a web-based patient information system for individuals with severe mental health problems. J Adv Nurs 2010 Dec;66(12):2701-2710. [doi: 10.1111/j.1365-2648.2010.05411.x] [Medline: 20722811]
38. Strudwick G, Eyasu T. Electronic health record use by nurses in mental health settings: a literature review. Arch Psychiatr Nurs 2015 Aug;29(4):238-241. [doi: 10.1016/j.apnu.2015.03.007] [Medline: 26165979]
39. Koivunen M, Hätönen H, Välimäki M. Barriers and facilitators influencing the implementation of an interactive internet-portal application for patient education in psychiatric hospitals. Patient Educ Couns 2008 Mar;70(3):412-419. [doi: 10.1016/j.pec.2007.11.002] [Medline: 18079085]
40. Välimäki M, Hätönen H, Lahti M, Kuosmanen L, Adams CE. Information and communication technology in patient education and support for people with schizophrenia. Cochrane Database Syst Rev 2012 Oct;10:CD007198. [doi: 10.1002/14651858.CD007198.pub2] [Medline: 23076932]
41. Välimäki M, Athanasopoulou C, Lahti M, Adams CE. Effectiveness of social media interventions for people with schizophrenia: a systematic review and meta-analysis. J Med Internet Res 2016;18(4):e92 [FREE Full text] [doi: 10.2196/jmir.5385] [Medline: 27105939]
42. Hidalgo-Mazzei D, Mateu A, Reinares M, Matic A, Vieta E, Colom F. Internet-based psychological interventions for bipolar disorder: review of the present and insights into the future. J Affect Disord 2015 Aug 28;188:1-13. [doi: 10.1016/j.jad.2015.08.005] [Medline: 26342885]
43. Topolovec-Vranic J, Mansoor Y, Ennis N, Lightfoot D. Technology-adaptable interventions for treating depression in adults with cognitive impairments: protocol for a systematic review. Syst Rev 2015 Apr 07;4:42 [FREE Full text] [doi: 10.1186/s13643-015-0032-4] [Medline: 25875655]
44. Koivunen M, Välimäki M, Hätönen H. Nurses' information retrieval skills in psychiatric hospitals - are the requirements for evidence-based practice fulfilled? Nurse Educ Pract 2010 Jan;10(1):27-31. [doi: 10.1016/j.nepr.2009.03.004] [Medline: 19386551]
45. Hwang JI, Park HA. Factors associated with nurses' informatics competency. Comput Inform Nurs 2011 Apr;29(4):256-262. [doi: 10.1097/NCN.0b013e3181fc3d24] [Medline: 21076284]
46. Eley R, Fallon T, Soar J, Buikstra E, Hegney D. Barriers to use of information and computer technology by Australia's nurses: a national survey. J Clin Nurs 2009 Apr;18(8):1151-1158. [doi: 10.1111/j.1365-2702.2008.02336.x] [Medline: 19320784]
47. Leigh S, Flatt S. App-based psychological interventions: friend or foe? Evid Based Ment Health 2015 Nov;18(4):97-99. [doi: 10.1136/eb-2015-102203] [Medline: 26459466]
48. Winslow S, DeGuzman P, Kulbok P, Jackson S. Nurses' self-efficacy and academic degree advancement. J Nurses Prof Dev 2014;30(3):110-116. [doi: 10.1097/NND.0000000000000038] [Medline: 24845088]
49. Barta SK, Stacy RD. The effects of a theory-based training program on nurses' self-efficacy and behavior for smoking cessation counseling. J Contin Educ Nurs 2005;36(3):117-123. [Medline: 16022031]
50. Nørgaard B, Ammentorp J, Ohm Kyvik K, Kofoed P. Communication skills training increases self-efficacy of health care professionals. J Contin Educ Health Prof 2012;32(2):90-97. [doi: 10.1002/chp.21131] [Medline: 22733636]

XSL•FO
RenderX

51. Roh KH, Park HA. A meta-analysis on the effectiveness of computer-based education in nursing. Healthc Inform Res 2010 Sep;16(3):149-157 [FREE Full text] [doi: 10.4258/hir.2010.16.3.149] [Medline: 21818434]
52. Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, et al. The use and interpretation of quasi-experimental studies in medical informatics. J Am Med Inform Assoc 2006;13(1):16-23 [FREE Full text] [doi: 10.1197/jamia.M1749] [Medline: 16221933]
53. Polit DF, Beck CT. Essentials of Nursing Research: Appraising Evidence for Nursing Practice. 7th edition. Philadelphia, PA: Lippincott Williams & Wilkins; 2010.
54. National Institute for Health and Welfare. 2017. Psychiatric specialist medical care 2015 URL: http://www.julkari.fi/bitstream/handle/10024/134729/Tr21_17.pdf?sequence=1 [accessed 2018-05-29] [WebCite Cache ID 6zmlQlbE4]
55. National Institute for Health and Welfare. 2011. Guide for patient security [Potilasturvallisuusopas] URL: https://thl.fi/documents/10531/104871/Opas%202011%2015.pdf [accessed 2018-05-29] [WebCite Cache ID 6zmkwqEh7]
56. National Institute For Health and Welfare. 2015. Personnel in local government health and social services 2014 URL: https://www.julkari.fi/bitstream/handle/10024/126388/Tr16_15_kokonaisraportti.pdf?sequence=4 [accessed 2018-05-29] [WebCite Cache ID 6zml1js0H]
57. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
58. Anttila M, Koivunen M, Välimäki M. Information technology-based standardized patient education in psychiatric inpatient care. J Adv Nurs 2008 Oct;64(2):147-156. [doi: 10.1111/j.1365-2648.2008.04770.x] [Medline: 18990095]
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. Health On the Net. 2017. HONcode URL: https://www.healthonnet.org/HONcode/Finnish/?HONConduct193826 [accessed 2018-05-29] [WebCite Cache ID 6zmBOfDLt]
61. World Health Organization. 2016. International Statistical Classification of Diseases and Related Health Problems 10th Revision URL: http://apps.who.int/classifications/icd10/browse/2016/en [accessed 2018-05-29] [WebCite Cache ID 6zmJS92Zf]
62. Valimaki M, Anttila M, Hatonen H, Koivunen M, Jakobsson T, Pitkanen A, et al. Design and development process of patient-centered computer-based support system for patients with schizophrenia spectrum psychosis. Inform Health Soc Care 2008 Jun;33(2):113-123. [doi: 10.1080/17538150802127207] [Medline: 18604755]
63. 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. Br Med J 2014;348:g1687 [FREE Full text] [Medline: 24609605]
64. Curran VR, Fleet L. A review of evaluation outcomes of web-based continuing medical education. Med Educ 2005 Jun;39(6):561-567. [doi: 10.1111/j.1365-2929.2005.02173.x] [Medline: 15910431]
65. Xia J, Merinder LB, Belgamwar MR. Psychoeducation for schizophrenia. Cochrane Database Syst Rev 2011 Jun 15(6):CD002831 [FREE Full text] [doi: 10.1002/14651858.CD002831.pub2] [Medline: 21678337]
66. Leung EY, Malick SM, Khan KS, EBM-CONNECT Collaboration. On-the-job evidence-based medicine training for clinician-scientists of the next generation. Clin Biochem Rev 2013 Aug;34(2):93-103 [FREE Full text] [Medline: 24151345]
67. Holbrook A. Self-Reported Measure. In: Lavrakas P, editor. Encyclopedia of Survey Research Methods. Thousand Oaks, CA: Sage Pubns; 2008.
68. Waltz C, Strickland OL, Lenz E. Measurement in Nursing and Health Research, Fourth Edition. New York: Springer; 2010.
69. Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Weinman J, Wright SC, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs. Windsor, UK: NFER-Nelson; 1995:35-37.
70. Lu JL. Multiple interactions of hazard exposures, role stressors and situational factors, and burnout among nurses. Int J Occup Saf Ergon 2007;13(1):73-82. [doi: 10.1080/10803548.2007.11076710] [Medline: 17362660]
71. Shoji K, Cieslak R, Smoktunowicz E, Rogala A, Benight CC, Luszczynska A. Associations between job burnout and self-efficacy: a meta-analysis. Anxiety Stress Coping 2016 Dec;29(4):367-386. [doi: 10.1080/10615806.2015.1058369] [Medline: 26080024]
72. Scholz U, Gutiérrez Doña B, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Eur J Psychol Assess 2002 Sep 1;18(3):242-251. [doi: 10.1027//1015-5759.18.3.242]
73. Rosenberg M. Society and the adolescent self-image. Princeton, NJ: Princeton University Press; 1965.
74. Leão ER, Dal Fabbro DR, Oliveira RB, Santos IR, Victor ED, Aquarone RL, et al. Stress, self-esteem and well-being among female health professionals: a randomized clinical trial on the impact of a self-care intervention mediated by the senses. PLoS One 2017;12(2):e0172455 [FREE Full text] [doi: 10.1371/journal.pone.0172455] [Medline: 28241070]
75. Losa-Iglesias ME, López López D, Rodriguez Vazquez R, de Bengoa-Vallejo RB. Relationships between social skills and self-esteem in nurses: a questionnaire study. Contemp Nurse 2017 Dec;53(6):681-690. [doi: 10.1080/10376178.2018.1441729] [Medline: 29451080]

XSL•FO
RenderX

76. Rahnama F, Mardani-Hamooleh M, Kouhnavard M. Correlation between moral sensitivity and self-esteem in nursing personnel. J Med Ethics Hist Med 2017;10:16 [FREE Full text] [Medline: 30258550]
77. Schmitt DP, Allik J. Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. J Pers Soc Psychol 2005 Oct;89(4):623-642. [doi: 10.1037/0022-3514.89.4.623] [Medline: 16287423]
78. Anderson NR, West MA. Measuring climate for work group innovation: development and validation of the team climate inventory. J Organ Behav 1998 May;19(3):235-258. [doi: 10.1002/(SICI)1099-1379(199805)19:3<235::AID-JOB837>3.0.CO;2-C]
79. Kivimäki M, Elovainio M. A short version of the Team Climate Inventory: development and psychometric properties. J Occup Organ Psycho 1999;72(2):241-246. [doi: 10.1348/096317999166644]
80. West MA. The social psychology of innovation in groups. In: West MA, Farr JL, editors. Innovation and creativity at work: Psychological and Organizational Strategies. Oxford, England: Wiley; 1990:309-333.
81. Koivunen M, Anttila M, Kuosmanen L, Katajisto J, Välimäki M. Team climate and attitudes toward information and communication technology among nurses on acute psychiatric wards. Inform Health Soc Care 2015 Jan;40(1):79-90. [doi: 10.3109/17538157.2013.872112] [Medline: 24393065]
82. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
83. Athanasopoulou C, Välimäki M, Koutra K, Löttyniemi E, Bertsias A, Basta M, et al. Internet use, eHealth literacy and attitudes toward computer/internet among people with schizophrenia spectrum disorders: a cross-sectional study in two distant European regions. BMC Med Inform Decis Mak 2017 Sep 20;17(1):136 [FREE Full text] [doi: 10.1186/s12911-017-0531-4] [Medline: 28931385]
84. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Lawrence Erlbaum; 1988.
85. EasyCalculation. Effect size calculator URL: https://www.easycalculation.com/statistics/effect-size.php [accessed 2018-11-14] [WebCite Cache ID 73vpWQ0Ys]
86. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
87. World Medical Association. 2013. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2018-05-29] [WebCite Cache ID 6zmJYR0iM]
88. Koohestani HR, Baghcheghi N. The effects of team-based learning techniques on nursing students' perception of the psycho-social climate of the classroom. Med J Islam Repub Iran 2016;30:437 [FREE Full text] [Medline: 28210602]
89. Mohamed Z, Newton JM, McKenna L. Belongingness in the workplace: a study of Malaysian nurses' experiences. Int Nurs Rev 2014 Mar;61(1):124-130. [doi: 10.1111/inr.12078] [Medline: 24512262]
90. Dackert I. The impact of team climate for innovation on well-being and stress in elderly care. J Nurs Manag 2010 Apr;18(3):302-310. [doi: 10.1111/j.1365-2834.2010.01079.x] [Medline: 20546470]
91. Kauppi K, Hätönen H, Adams CE, Välimäki M. Perceptions of treatment adherence among people with mental health problems and health care professionals. J Adv Nurs 2015 Apr;71(4):777-788. [doi: 10.1111/jan.12567] [Medline: 25394750]
92. National Institute for Health and Care Excellence. 2014. Psychosis and schizophrenia in adults: prevention and management URL: https://www.nice.org.uk/guidance/cg178/ [accessed 2018-05-30] [WebCite Cache ID 6zmmMFhrO]
93. Local Government Employers. 2016. Further training in municipal social and health care staff and training for the development of vocational competence of other personnel [Kunnallisen sosiaali- ja terveydenhuollon henkilöstön ammatillinen täydennyskoulutus sekä koko muun henkilöstön ammatillisen osaamisen kehittämistä koskeva koulutus vuonna 2015] URL: https://bit.ly/2GxjcBY [accessed 2018-05-29] [WebCite Cache ID 6zmkgCBnH]
94. Lasalvia A, Boggian I, Bonetto C, Saggioro V, Piccione G, Zanoni C, et al. Multiple perspectives on mental health outcome: needs for care and service satisfaction assessed by staff, patients and family members. Soc Psychiatry Psychiatr Epidemiol 2012 Jul;47(7):1035-1045. [doi: 10.1007/s00127-011-0418-0] [Medline: 21850522]

Abbreviations

GSE: General Self-Efficacy Scale
SES: Self-Esteem Scale
TCI: Team Climate Inventory

XSL•FO
RenderX

Edited by G Eysenbach; submitted 31.05.18; peer-reviewed by A Kerkhof, I Belan; comments to author 06.10.18; revised version received 15.11.18; accepted 25.11.18; published 01.03.19.

Please cite as:
Laine A, Välimäki M, Löyttyniemi E, Pekurinen V, Marttunen M, Anttila M
The Impact of a Web-Based Course Concerning Patient Education for Mental Health Care Professionals: Quasi-Experimental Study
J Med Internet Res 2019;21(3):e11198
URL: https://www.jmir.org/2019/3/e11198/
doi:10.2196/11198
PMID:30821697

©Anna Laine, Maritta Välimäki, Eliisa Löyttyniemi, Virve Pekurinen, Mauri Marttunen, Minna Anttila. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of an E-Learning Training Program to Support Implementation of a Group-Based, Theory-Driven, Self-Management Intervention For Osteoarthritis and Low-Back Pain: Pre-Post Study

Deirdre A Hurley[1*], BPhysio, MAppSc, PhD; Alison Keogh[1], BSc, MSc, PhD; Danielle Mc Ardle[1], BSc, MSc; Amanda M Hall[2], BKin (Hons), MPE, PhD; Helen Richmond[3], BSc, MSc, PhD; Suzanne Guerin[4], BA, PhD; Tara Magdalinski[5], BHMS (Hons), PhD; James Matthews[1*], BA, MA, PhD

[1]School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland

[2]Faculty of Medicine, Memorial University, St Johns, NL, Canada

[3]Centre for Rehabilitation Research in Oxford (RRiO), Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom

[4]School of Psychology, University College Dublin, Dublin, Ireland

[5]Faculty of Health, Arts and Design, Swinburne University of Technology, Melbourne, Australia

[*]these authors contributed equally

Corresponding Author:
Deirdre A Hurley, BPhysio, MAppSc, PhD
School of Public Health, Physiotherapy and Sports Science
University College Dublin
Belfield
Dublin, D4
Ireland
Phone: 353 17166524
Fax: 353 17166501
Email: deirdre.hurleyosing@ucd.ie

Abstract

Background: By adaptation of the face-to-face physiotherapist-training program previously used in the Self-management of Osteoarthritis and Low back pain through Activity and Skills (SOLAS) feasibility trial, an asynchronous, interactive, Web-based, e-learning training program (E-SOLAS) underpinned by behavior and learning theories was developed.

Objective: This study investigated the effect of the E-SOLAS training program on relevant outcomes of effective training and implementation.

Methods: Thirteen physiotherapists from across Ireland were trained via E-SOLAS by using mixed methods, and seven physiotherapists progressed to implementation of the 6-week group-based SOLAS intervention. The effectiveness of E-SOLAS was evaluated using the Kirkpatrick model at the levels of reaction (physiotherapist engagement and satisfaction with E-SOLAS training methods and content), learning (pre- to posttraining changes in physiotherapists' confidence and knowledge in delivering SOLAS content and self-determination theory-based communication strategies, administered via a SurveyMonkey questionnaire), and behavior (fidelity to delivery of SOLAS content using physiotherapist-completed weekly checklists). During implementation, five physiotherapists audio recorded delivery of one class, and the communication between physiotherapists and clients was assessed using the Health Care Climate Questionnaire (HCCQ), the Controlling Coach Behaviour Scale (CCBS), and an intervention-specific measure (ISM; 7-point Likert scale). A range of implementation outcomes were evaluated during training and delivery (ie, acceptability, appropriateness, feasibility, fidelity, and sustainability of E-SOLAS) using a posttraining feedback questionnaire and individual semistructured telephone interviews.

Results: With regard to their reaction, physiotherapists (n=13) were very satisfied with E-SOLAS posttraining (median 5.0; interquartile range 1.0; min-max 4.0-5.0) and completed training within 3-4 weeks. With regard to learning, there were significant increases in physiotherapists' confidence and knowledge in delivery of all SOLAS intervention components (P<.05). Physiotherapists' confidence in 7 of 10 self-determination theory-based communication strategies increased (P<.05), whereas

XSL•FO
RenderX

physiotherapists' knowledge of self-determination theory-based strategies remained high posttraining (P>.05). In terms of behavior, physiotherapists delivered SOLAS in a needs supportive manner (HCCQ: median 5.2, interquartile range 1.3, min-max 3.7-5.8; CCBS: median 6.6, interquartile range 1.0, min-max 5.6-7.0; ISM: median 4.5, interquartile range 1.2, min-max 2.8-4.8). Fidelity scores were high for SOLAS content delivery (total %mean fidelity score 93.5%; SD 4.9%). The posttraining questionnaire and postdelivery qualitative interviews showed that physiotherapists found E-SOLAS acceptable, appropriate, feasible, and sustainable within primary care services to support the implementation of the SOLAS intervention.

Conclusions: This study provides preliminary evidence of the effectiveness, acceptability, and feasibility of an e-learning program to train physiotherapists to deliver a group-based self-management complex intervention in primary care settings, which is equivalent to face-to-face training outcomes and would support inclusion of physiotherapists in a definitive trial of SOLAS.

(J Med Internet Res 2019;21(3):e11123) doi:10.2196/11123

KEYWORDS

technology-enhanced learning; evaluation; e-learning; digital learning; program evaluation; effectiveness; physiotherapy; implementation; osteoarthritis; low-back pain

Introduction

International clinical guidelines for osteoarthritis and low-back pain endorse self-management, exercise, and physical activity as key components of health care interventions [1-4], but the evidence for their effectiveness is weak and of low quality [5-7]. The Self-management of Osteoarthritis and Low back pain through Activity and Skills (SOLAS) intervention is an evidence-supported group treatment approach developed through intervention mapping [8], which is a logical six-step process for the development and evaluation of theory-driven and evidence-based interventions that takes into account stakeholder needs and the practicalities of implementation [9]. SOLAS was evaluated for its acceptability and preliminary effects in comparison with individual physiotherapy in a feasibility trial (trial registration: ISRCTN49875385) set in Dublin, Ireland, between September 2014 and June 2016 [10]. Intervention physiotherapists who participated in the trial were trained using brief interactive lectures, videos, role play, and practical skills to deliver the SOLAS intervention using communication skills underpinned by self-determination theory. This theory proposes that people have basic psychological needs for autonomy, competence, and relatedness, which if met, for example, by the needs supportive communication style of a health care practitioner (HCP), will increase an individual's autonomous motivation and engagement in health behaviors such as self-management [11]. The Medical Research Council guidelines recommend that complex behavior-change programs train their intervention deliverers to ensure implementation with high fidelity [12]. Hence, the Kirkpatrick model was used to evaluate training at the levels of Reaction, Learning, and Behavior [13], which showed that physiotherapists were satisfied with face-to-face training and their confidence in the self-determination theory-based communication strategies. Knowledge of the intervention content significantly increased, and the physiotherapists delivered SOLAS in a needs supportive manner with high fidelity to the intervention content [14,15]. Upscaling to a definitive national trial would render the face-to-face training impractical for physiotherapists due to significant time, travel, and costs constraints [16]. Therefore, we subsequently developed an asynchronous, interactive, Web-based, e-learning training program for SOLAS (E-SOLAS) to prepare physiotherapists to deliver the SOLAS intervention. If successful, the program would reduce the time needed to move to a definitive trial. Furthermore, E-SOLAS has the potential to increase the competencies of physiotherapists with regard to self-management behavior-change skills in line with the shared strategic priority of Ireland's public health service and higher education institutions to train and prepare future health care graduates with the skills necessary to support lifestyle behavior change in their patients [17,18], making the intervention more accessible to physiotherapists for long-term sustainability.

Despite the increased availability of e-learning training for HCPs internationally, there is limited formal evaluation of such training programs. Current evidence, which predominantly involves undergraduate HCP students [19,20], suggests that e-learning shows similar effectiveness to traditional methods for knowledge acquisition [21] and user satisfaction [22], but further research regarding the effectiveness of e-learning on HCP behavior change and the translation of learning to clinical practice has been advocated [23]. Hence, we evaluated E-SOLAS in the same way as our face-to-face training. In addition, a range of World Health Organization–recommended implementation outcomes were included for evaluation, including the acceptability, appropriateness, feasibility, fidelity, and sustainability of E-SOLAS, in a range of primary care physiotherapist settings across Ireland [24,25] in order to understand the contextual elements of e-learning [16].

The study objectives were to evaluate the effect of the E-SOLAS training program on physiotherapists' reaction, learning, and delivery of the SOLAS intervention as intended and to assess the acceptability, appropriateness, feasibility, fidelity, and sustainability of E-SOLAS to aid the implementation of the SOLAS intervention in primary care settings.

Methods

Study Design and Research Ethics

This was a single-group, pre-post study. Ethical approval was granted by the UCD Human Subject (Sciences) Ethics Committee in two phases: in Phase 1, for the E-SOLAS training program (September 30, 2016; LS-E-16-121-Hurley) and in Phase 2, for implementation of the SOLAS intervention (December 21, 2016; LS-16-97-Hurley). The study was also

XSL•FO
RenderX

approved for Phase 1 (November 17, 2016) and Phase 2 (January 17, 2017) by the Health Service Executive Primary Care Research Committee.

Participants and Procedure

Physiotherapy managers from 10 primary care areas across Ireland who had not participated in the SOLAS feasibility trial were sent a study information leaflet for screening based on their service facilities and staffing capabilities. Seven physiotherapy managers fulfilled the criteria for inclusion, provided letters of support, and nominated two staff members to undertake E-SOLAS training. Nominated physiotherapist staff were sent the study information leaflet and consent form. Consenting participants were required to possess a device that could connect to the internet and were given password-protected access to the social learning platform Curatr [26] that hosted the E-SOLAS training program. Participants were encouraged to complete the training over a 4-week period by working at their own pace and at times that were convenient for them. During training, they had access to ongoing technical support from the research team and were requested to keep a log of the time spent on each aspect of training.

At the end of the training period, participants were invited to set up and deliver the SOLAS intervention according to the treatment protocol [10] in each of their primary care areas. Physiotherapists had ongoing access to E-SOLAS during implementation and were provided with any additional intervention materials required to deliver the intervention by the research team (ie, intervention PowerPoint [Microsoft Corp, Redmond, WA] slide deck on a universal serial bus, pedometers, and relaxation CDs for each client). Following completion of the 6-week delivery phase, each physiotherapist was invited to participate in an individual semistructured telephone interview to explore their views of E-SOLAS as a tool to support implementation of the intervention.

E-SOLAS Training Program

E-SOLAS is a Web-based e-learning training program designed to train physiotherapists to deliver a group-based education and exercise intervention for patients with osteoarthritis and chronic low-back pain. The content is based on the face-to-face training program developed for the SOLAS feasibility trial [10,15]. The E-SOLAS program is hosted on Curatr [26], an online social learning platform that creates a collaborative learning environment and uses gamification principles. The development process for the E-SOLAS program is outlined in Multimedia Appendix 1 [27-31].

E-SOLAS Program Structure and Content

The E-SOLAS program contained six hierarchical linear levels, whereby the user was required to finish each level before progressing to the next level (Multimedia Appendix 2, Figure 1).

Briefly, the program begins in Level 1 with an overview of the training program and the SOLAS intervention. Level 2 describes the education content for each week of the SOLAS intervention (eg, the key learning points and the materials required; Figure 2). At Level 3, the self-determination theory-based communication strategies that physiotherapists use as part of the intervention are introduced (Figure 3), and in Level 4, they are given the opportunity to role play these strategies. Level 5 highlights the exercises and their mode of delivery, and finally, Level 6 concludes the program by highlighting the next steps for intervention delivery.

Figure 1. E-SOLAS home page screenshot. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.



XSL•FO
RenderX

Figure 2. E-SOLAS program content screenshot. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

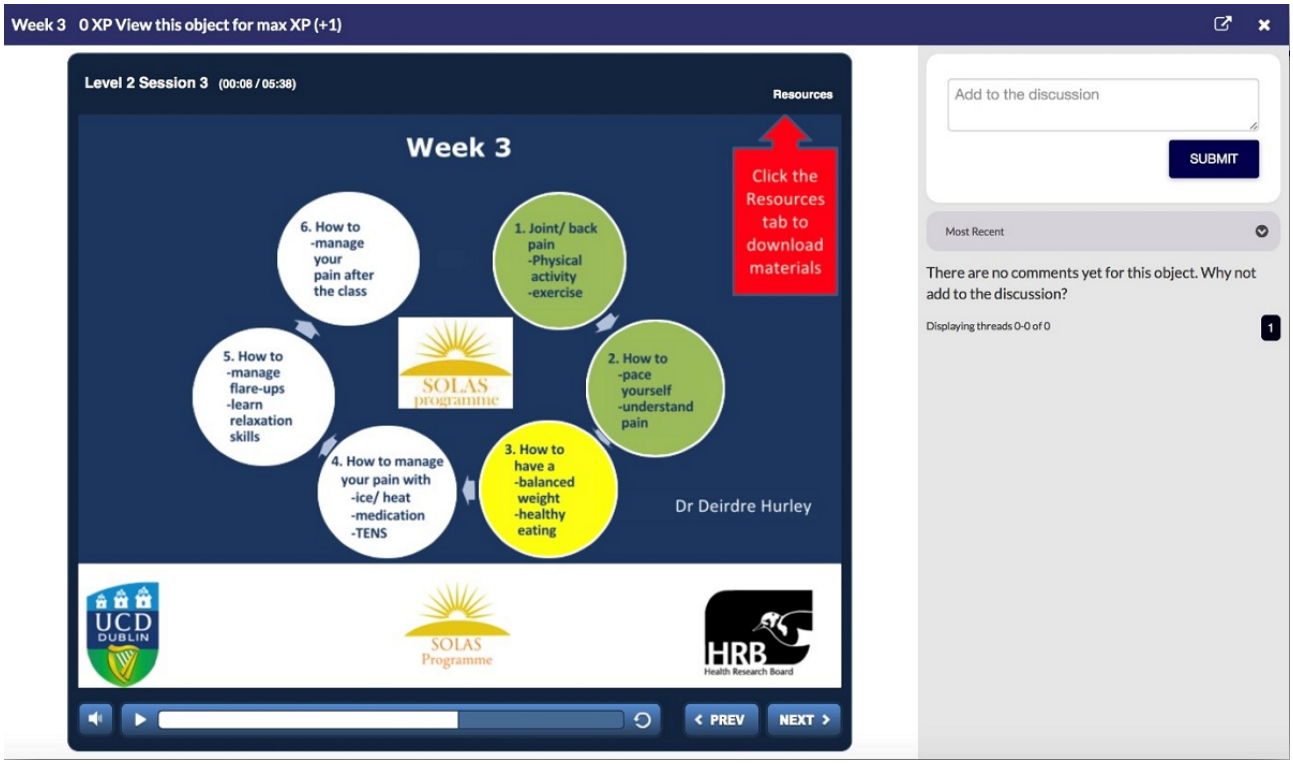
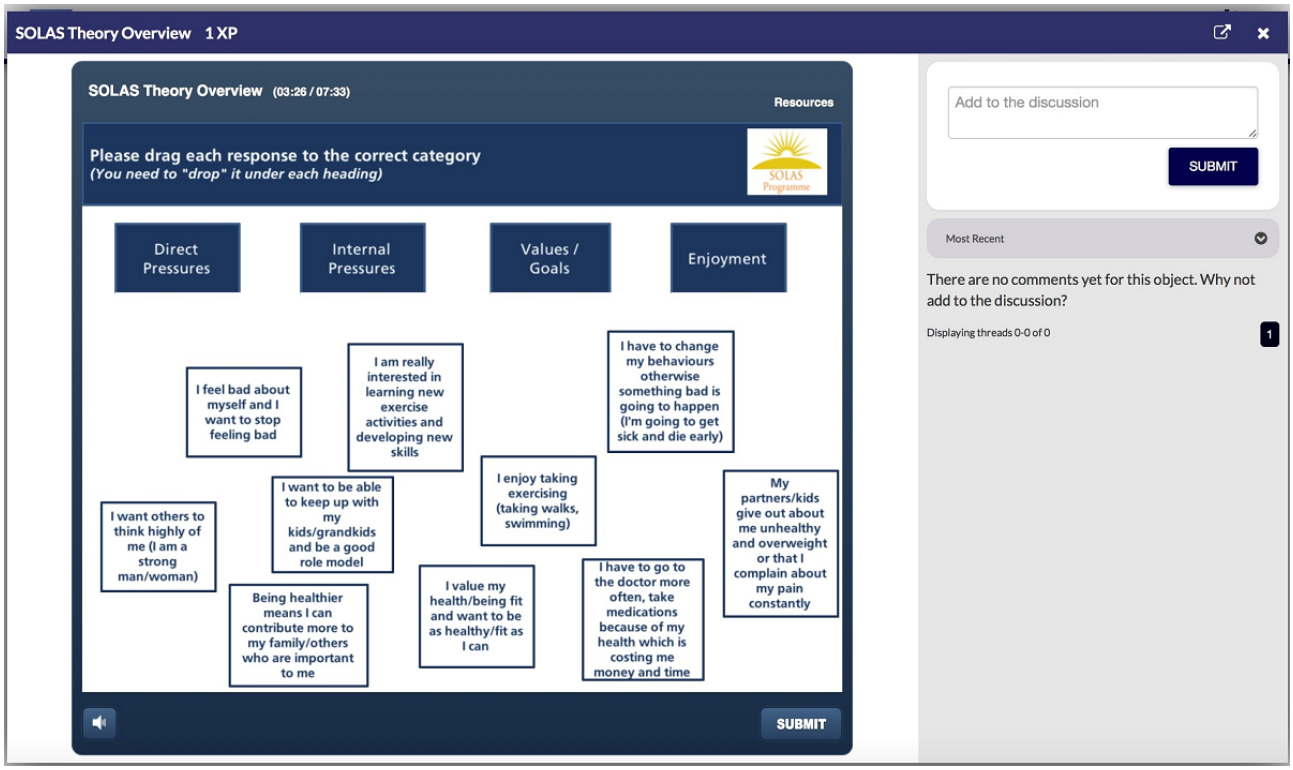
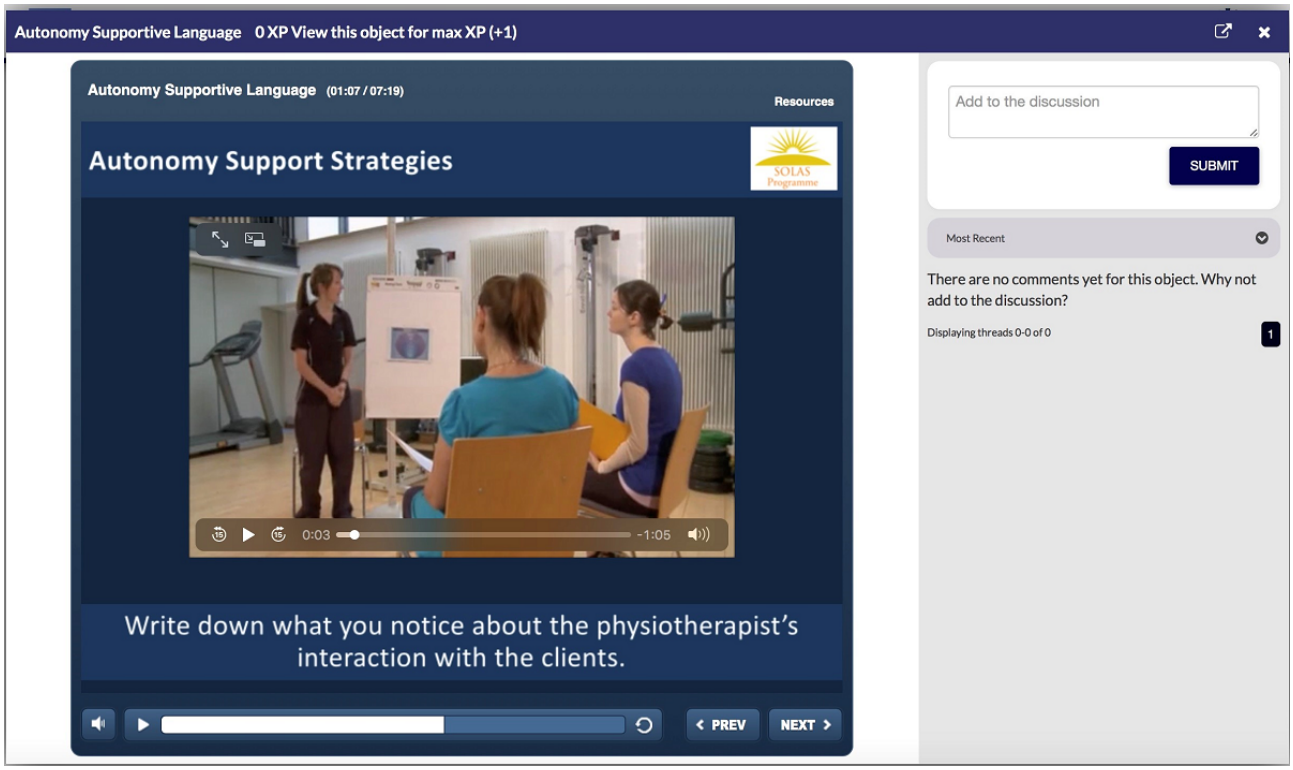


Figure 3. E-SOLAS theory screenshot. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.



XSL•FO
RenderX

Figure 4. E-SOLAS video-based program activity screenshot. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.



Throughout each level, there are lectures with voice-overs, video examples of good and poor practice (Figure 4), videos from the research team, and a peer role model explaining certain elements of the intervention; short "in level" activities and self-reflection opportunities; and end-of-level "gate" assessments with varying levels and modes of feedback depending on the activity.

Outcome Measures

The effectiveness of the E-SOLAS training program was assessed using the Kirkpatrick model of evaluation at the levels of reaction, learning, and behavior [13]. Furthermore, a range of implementation outcomes were evaluated during the training and intervention-delivery phases. The measurement tools used to assess learning and implementation outcomes are described in detail in Multimedia Appendix 3 and briefly outlined below.

Training Outcomes

Reaction

To measure physiotherapists' reaction to E-SOLAS, a researcher-devised feedback measure was developed by adapting the face-to-face training feedback measure [15] and incorporating factors related to the evaluation of technology enhanced learning [25]. It was administered following posttraining assessment via SurveyMonkey and included items related to participant satisfaction, engagement, accessibility, and quality of E-SOLAS as well as several implementation outcomes as detailed below (Multimedia Appendix 4). Physiotherapists' engagement with E-SOLAS was further evaluated using Curatr analytics [26] and a self-reported activity log completed by each physiotherapist during training.

Learning

Learning was assessed by evaluating physiotherapists' perceptions of self-reported knowledge and confidence pre- and posttraining (Multimedia Appendix 5) and their use of skills during training by using a range of measures.

Behavior

Physiotherapists' behavior was assessed during delivery of the SOLAS intervention to evaluate fidelity to the intervention content and self-determination theory-based communication strategies using previously validated checklists [14] and audio recordings [15]. In line with fidelity guidelines [32], each audio recording was coded by one blinded expert rater (AK) to assess physiotherapists' communication style [33], and three audio recordings were coded by a second expert rater (JM). The Health Care Climate Questionnaire (HCCQ) [34] was the primary measure to assess provider delivery of the self-determination theory-based communication style, with an adapted version of the Controlling Coach Behaviour Scale (CCBS) [35] and an intervention-specific SOLAS scale used as secondary measures [15].

Implementation Outcomes

Implementation outcomes were measured through specific items on the feedback measure related to the acceptability, appropriateness, feasibility, and sustainability of E-SOLAS (Multimedia Appendix 4) and an individual semistructured telephone interview of the physiotherapists conducted by an experienced qualitative researcher (SG) within 2 weeks of completing group class delivery. A topic guide was developed for the participant interviews with specific questions and probes related to their views of E-SOLAS as a model of training in order to support physiotherapists in delivering the SOLAS

XSL•FO
RenderX

intervention in primary care settings. All interviews were audio recorded.

Data Analysis

Data from all outcome measures were analyzed using Excel (version 14.2.3, Microsoft Corp) and a statistical software package (SPSS Statistics, version 20, IBM Corp, Armonk, NY) following checks for errors in data entry.

Training Outcomes

Reaction

In order to assess physiotherapists' views of their satisfaction, accessibility, and quality of the E-SOLAS program and their engagement with the e-learning training, descriptive statistics were used to analyze quantitative data, and thematic analysis was used to analyze free-text answers.

Learning

Analytical Methods

Descriptive statistics were used to calculate scores pre- and posttraining for overall confidence in delivering SOLAS content, the specific SOLAS intervention components, and the use of each self-determination theory-based communication strategy. Differences between pre- and posttraining were calculated using the Wilcoxon signed-rank tests and adjusted for multiplicity using Bonferroni corrections (0.05/n tests).

Knowledge

Descriptive statistics were used to calculate the level of SOLAS intervention knowledge, and pre- and posttraining differences were calculated using the Wilcoxon signed-rank tests. Following discussion between raters, there was excellent agreement (100%) in the coding of physiotherapists' narrative case studies. The number of self-determination theory-based communication strategies used by each physiotherapists and the percentage of physiotherapists who used each strategy was calculated, with differences in the rate of use of all strategies and each strategy pre- and posttraining determined using McNemar tests. All results were adjusted for multiplicity using Bonferroni corrections.

Skills

Each role-play audio recording was rated for the use of self-determination theory-based communications strategies on a 7-point Likert scale ranging from "1 - not at all well" to "7 - very well," with values at or above the mid-point of the Likert scale (4/7) defined as demonstrating skills that could be considered acceptable in terms of competence [15].

Behavior

The mean fidelity levels to SOLAS intervention content and fidelity levels according to physiotherapist, site, session, and session category were obtained by calculating total actual scores as a percentage of the total possible score using checklists. Fidelity of duration was established by calculating the difference between the actual and the intended session durations using a one-sample Wilcoxon test. Levels of fidelity were interpreted as previously reported in the literature [36]. A review of the raters' scores for the audio recordings of physiotherapists' delivery of SOLAS session 4 demonstrated excellent agreement (90%). To establish physiotherapists' competence in the self-determination theory-based communication style, a median result for each of the three outcome measures was calculated separately. For the SOLAS scale, a median score per construct subsection (eg, autonomy), subcomponent strategy (eg, positive feedback), and class component (eg, education) was also calculated.

Implementation Outcomes

Descriptive statistics were used to analyze quantitative data related to physiotherapists' views of the acceptability, appropriateness, feasibility, and sustainability of E-SOLAS. Qualitative data from the physiotherapists' interviews were transcribed verbatim and analyzed using inductive thematic analysis [37]. A coding frame was developed from a review of provisional themes, which were then reexamined and refined (DMcA). The reliability of the identified themes was established by a second researcher (DAH) who independently coded a random sample of 25% of each dataset using the coding frame, with 70% agreement taken as the minimum cut-off rate [10]. The level of agreement between raters was 85%.

Results

Principal Findings

Thirteen physiotherapists from seven primary care areas completed the E-SOLAS training, of which 12 were invited to participate in the implementation study (ie, delivery of the SOLAS intervention). Nine physiotherapists consented to participate, and seven progressed to deliver SOLAS. The profile of physiotherapists in each study phase is provided in Table 1, and the flow of participants through Phase 2 is outlined in Figure 5. The training and delivery groups were comparable for the majority of descriptive variables, apart from the median years qualified, which was lower in the delivery group.

Training Outcomes

Reaction

Physiotherapists (n=13) were very satisfied with E-SOLAS training posttraining and found it enjoyable and engaging, with all participants completing the program within the 4 weeks available (Multimedia Appendix 4). Physiotherapists reported that they spent a mean of 9.1 (SD 3.3) hours (min-max 4.1-16.1) over 16.3 (SD 6.0) days to complete E-SOLAS, which was not statistically different from the duration of training recorded by Curatr analytics (mean difference –1.69; SD 4.37; t=–1.397; df=12; P=.19; Multimedia Appendix 6). All physiotherapists successfully completed all-level gate assessments and the required three uploads and made at least one online posting to the group discussion. The majority of physiotherapists reported completing E-SOLAS outside work hours and spent 1-2 hours at any one time on training. The most commonly cited positive features of E-SOLAS were the range of brief video clips (46.2%; n=6) and focus on communication skills and client motivation (23.1%; n=3). Nine of the 13 participants experienced some difficulties during training; the most common difficulty was related to accessing online materials (46.2%, n=6), completing gate assessments (38.5%, n=5), and computer access at work

XSL•FO
RenderX

(30.8%, n=4). Although the median ratings for working independently and not having access to other therapists were very positive, four physiotherapists required support from the University College Dublin team during training for accessing resources (n=3), logging into E-SOLAS via work email (n=2), or uploading audio files (n=1). Nonetheless, the majority of physiotherapists highly rated the quality of the training program and format.

Table 1. Baseline characteristics of physiotherapists.

Demographic characteristics	Training group (n=13)	Delivery group (n=7)
Gender, n (%)		
Male	2.0 (15.4)	1.0 (14.3)
Female	11.0 (84.6)	6.0 (85.7)
Age (years), n (%)		
26-35	2.0 (15.4)	2.0 (28.6)
36-45	4.0 (30.8)	2.0 (28.6)
46-55	3.0 (23.0)	1.0 (14.2)
56-65	4.0 (30.8)	2.0 (28.6)
Number of physiotherapists in the primary care area (ID[a])		
Area 1	2.0 (4,7)	0
Area 2	1.0 (2)	0
Area 3	2.0 (6, 8)	1.0 (6; Site 2)[b]
Area 4	2.0 (5, 13)	2.0 (5 and 13; Site 4)[c]
Area 5	2.0 (1, 3)	2.0 (1, Site 1; 3, Site 6)[d]
Area 6	2.0 (10, 11)	2.0 (10, Sites 3 and 5; 11, Site 5)[e]
Area 7	2.0 (9, 12)	0
Clinical experience (years qualified), median (interquartile range), min-max	21.0 (15.5), 5.0-37.0	14.0 (18.0), 5.0-37.0
Delivered groups previously (yes), n (%)	12.0 (92.31)	6.0 (85.7)
Previous training, n (%)		
Communication skills (yes)	7.0 (53.8)	3.0 (42.9)]
E-Learning		
Yes	2.0 (15.4)	1.0 (14.3)
No	11.0 (84.6)	6.0 (85.7)
Preference for training, n (%)		
None	1.0 (7.7)	1.0 (14.3)
Face to face	2.0 (15.4)	1.0 (14.3)
E-learning and face to face	10.0 (76.9)	5.0 (71.4)

[a]ID: participant identification number.

[b]ID 6 delivered all 6 sessions at Site 2.

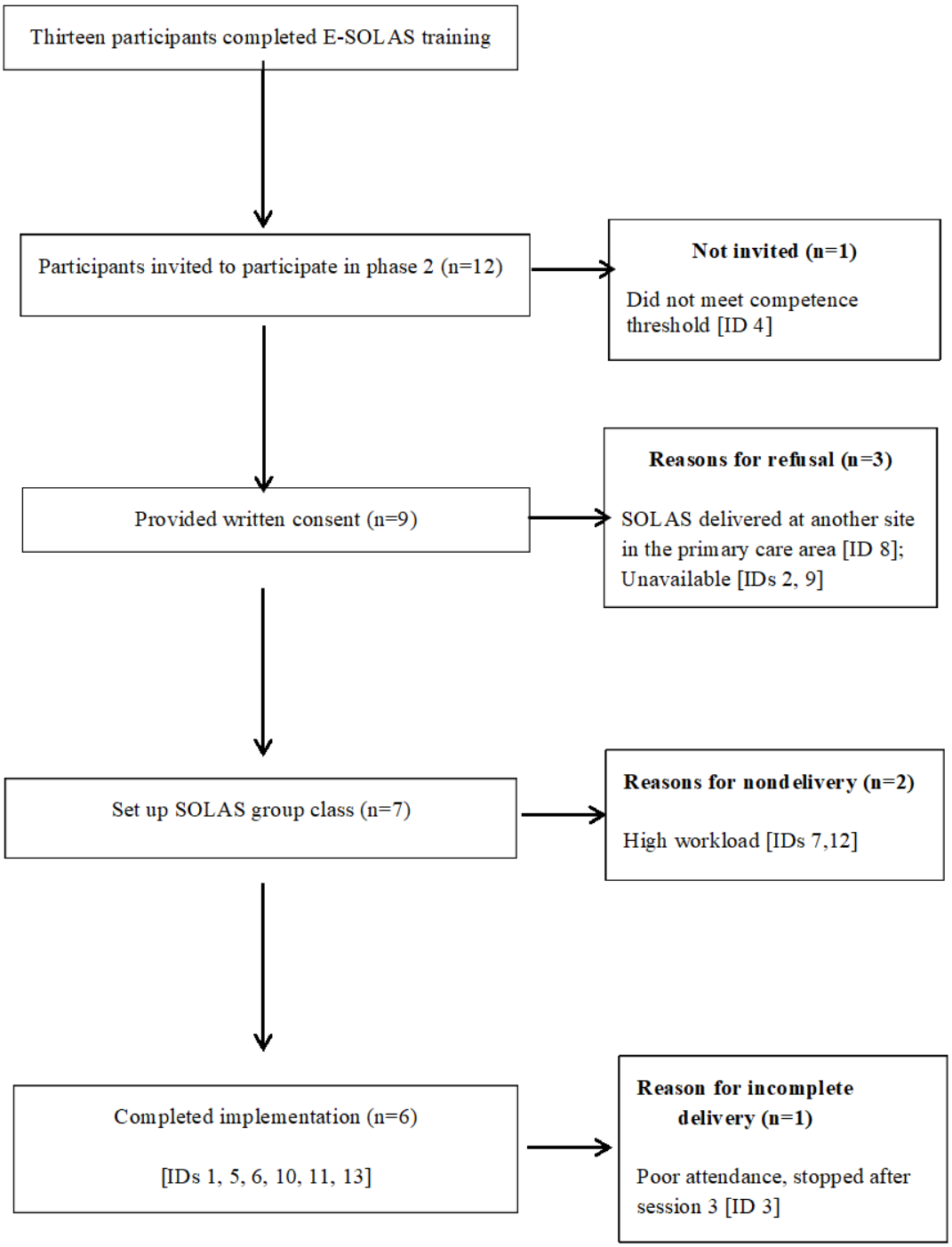
[c]IDs 5 and 13 delivered 3 sessions each at Site 4.

[d]ID 1 delivered all 6 sessions at Site 1; ID 3 ceased delivery in Site 6 after session 3.

[e]ID 10 delivered all 6 sessions at Site 3, ID 10 and 11 delivered 3 sessions each at Site 5.

XSL•FO
RenderX

Figure 5. Participant flow through the study. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills; ID: participant identification number.



Learning

Knowledge

Physiotherapists used all nine self-determination theory-based communication strategies in their responses pretraining; the most commonly used strategies were *collaborative goal setting and action planning* and *building relationships*, with no significant change in the rate of use of individual strategies posttraining (Table 2). Knowledge of the SOLAS intervention content and structure improved overall as well as in nine of the 10 intervention components. The use of pain modalities significantly increased posttraining, of which knowledge of content, structure, and group-based exercise programs remained significant following Bonferroni corrections (Table 3).

Confidence

Physiotherapists' confidence significantly increased posttraining overall and for 7 of the 10 individual self-determination theory strategies; *set clear expectations and provide direction* remained significant after Bonferroni correction (Table 2).

Similarly, physiotherapists' confidence in delivery of the SOLAS content overall and all 10 intervention components significantly increased posttraining; five components remained significant after Bonferroni correction (Table 3).

XSL•FO
RenderX

Table 2. Change in physiotherapists' confidence and knowledge of self-determination theory-based communication strategies.

SDT[a]-based communication strategies	Confidence[b]				Knowledge[c]		
	Median pretraining score (interquartile range), min-max	Median posttraining score (interquartile range), min-max	*Z* score	*P* value	Median pretraining scores (% of physiotherapists)	Median posttraining score (% of physiotherapists)	*P* value
Total	4.5 (1.0), 2.7-5.4	5.0 (0.8), 3.6-5.9	–2.08	.037	8.0 (2.0), 5.0-9.0[d]	7.5 (2.0), 6.0-9.0[d]	.70
Offer a meaningful rationale	5.0 (2.0), 3.0-7.0	7.0 (1.0), 0.0-7.0	–1.80	.07	11.0 (84.6)	8.0 (61.5)	.38
Provide opportunities for patient input and choice	4.0 (3.0), 2.0-7.0	6.0 (1.0), 4.0-7.0	–2.59	.009	6.0 (46.2)	10.0 (76.9)	.22
Use support and encouragement rather than pressurising behaviours	4.0 (2.0), 2.0-6.0	6.0 (2.0), 4.0-7.0	–2.45	.01	3.0 (23.1)	7.0 (53.8)	.29
Set clear expectations and provide direction[e]	5.00 (2.0), 3.0-6.0	6.00 (2.0), 5.0-7.0	–2.85	.004[f]	Not assessed	Not assessed	Not assessed
Collaborative goal setting and action planning[g]	5.0 (3.0), 2.0-7.0	6.0 (1.0), 5.0-7.0	–2.54	.01	12.0 (92.3)	12.0 (92.3)	1.00
Collaborative problem solving	N/A[h]	N/A	N/A	N/A	4.0 (30.8)	8.0 (61.5)	0.13
Provide positive information rich feedback	6.0 (1.0), 4.0-7.0	7.0 (1.0), 4.0-7.0	–1.84	.07	2.0 (15.4)	2.0 (15.4)	1.00
Provide opportunities to practice behaviors	6.0 (2.0), 2.0-6.0	6.0 (1.0), 4.0-7.0	–2.04	.04	3.0 (23.1)	1.0 (7.7)	.63
Acknowledge patients' feelings and perspectives	6.0 (2.0), 2.0-7.0	7.0 (1.0), 4.0-7.0	–2.09	.04	6.0 (46.2)	9.0 (69.2)	.38
Building relationships	6.0 (1.0), 4.0-7.0	7.0 (1.0), 6.0-7.0	–1.61	.11	12.0 (92.3)	12.0 (92.3)	1.00

[a]SDT: self-determination theory.

[b]Scale range: 1 (not at all good) to 7 (very good).

[c]Percentage of physiotherapists is calculated on the basis of the presence/absence of each SDT strategy in the narrative response.

[d]Values are presented as median (interquartile range), min-max

[e]"Setting clear expectations" was not included in the narrative component of the assessment, as it was not expected to be delivered within the context of the case study.

[f]Significant after Bonferroni adjustment for multiplicity.

[g]Problem solving was considered under the heading of goal setting within the confidence scale.

[h]N/A: not applicable.

XSL•FO
RenderX

Table 3. Change in physiotherapists' confidence and knowledge of the Self-management of Osteoarthritis and Low back pain through Activity and Skills intervention content.

Self-management of Osteoarthritis and Low back pain through Activity and Skills intervention content	Confidence[a]				Knowledge			
	Median pretraining score (interquartile range), min-max	Median posttraining score (interquartile range), min-max	Z score	*P* value	Median pretraining score interquartile range), min-max	Median posttraining score interquartile range), min-max	Z score[b]	*P* value
Total	4.9 (1.8), 3.0-6.0	6.3 (0.8), 4.0-7.0	–3.04	.002[c]	2.0 (2.0), 0.0-6.0	6.0 (1.5), 4.0-7.0	3.09	<.001[c]
Disease mechanisms	5.0 (1.0), 4.0-6.0	7.0 (1.0), 4.0-7.0	–2.49	.01	6.0 (1.0), 4.0-6.0	6.0 (1.0), 5.0-7.0	–2.46	.01
Exercise	6.0 (1.0), 5.0-7.0	7.0 (1), 4.0-7.0	–2.33	.02	6.0 (0.0), 5.0-7.0	7.0 (1.0), 5.0-7.0	–2.82	.005
Physical activity promotion	6.0 (1.0), 2.0-7.0	7.0 (1.0), 4.0-7.0	–2.04	.04	6.0 (1.0), 4.0-7.0	7.0 (1.0), 5.0-7.0	–1.81	.07
Healthy eating and diet	4.0 (2.0), 2.0-6.0	6.0 (1.5), 4.0-7.0	–2.97	.003[c]	4.0 (2.0), 3.0-6.0	6.0 (1.0), 4.0-7.0	–2.72	.006[c]
Relaxation	5.0 (3.0), 1.0-6.0	6.0 (1.5), 4.0-7.0	–2.86	.004[c]	4.0 (3.0), 1.0-7.0	6.0 (1.0), 4.0-7.0	–2.14	.03
Pain-relief techniques	5.0 (2.0), 4.0-6.0	7.0 (1.0), 4.0-7.0	–2.88	.004[c]	6.0 (2.0), 3.0-7.0	6.0 (1.0), 5.0-7.0	–2.48	.01
Medication	4.0 (2.0), 1.0-6.0	6.0 (0.5), 4.0-7.0	–2.88	.004[c]	4.00 (3), 1.0-7.0	6.00 (0.5), 5.0-7.0	–2.73	.006
Pacing	5.0 (2.0), 1.0-7.0	7.0 (1.0), 4.0-7.0	–2.82	.005	6.0 (3.0), 1.0-7.0	6.0 (1.0), 4.0-7.0	–2.50	.01
Mood regulation	4.0 (2.0), 1.0-6.0	6.0 (2.0), 4.0-7.0	–2.83	.005	4.0 (2.0), 1.0-7.0	6.00 (2.0), 4.0-7.0	–2.62	.009
Group-based exercise for osteoarthritis and chronic low-back pain	5.0 (2.0), 2.0-6.0	6.0 (1.0), 4.0-7.0	–3.02	.003[c]	5.0 (1.0), 2.0-6.0	6.0 (1.0), 5.0-7.0	–2.99	.003[c]
Cycle of change[d]	N/A[e]	N/A	N/A	N/A	57.0 (87.7)[f]	61.0 (93.8)[f]	–0.647	.52
Advice to patients[d]	N/A	N/A	N/A	N/A	33.5 (85.9)[f]	36.5 (93.5)[f]	–1.56	.12
Use of pain modalities[d]	N/A	N/A	N/A	N/A	83.0 (79.8)[f]	93.0 (89.4)[f]	–2.33	.02

[a]Scale range: 1 (not at all good) to 7 (very good).

[b]Z score from Wilcoxon signed-rank test.

[c]Significant after Bonferroni adjustment for multiplicity.

[d]Reported as a percentage of the total possible score for the knowledge category.

[e]N/A: not applicable.

[f]Values presented as percentage of the total possible score for the knowledge category and percentage of physiotherapists providing correct responses.

Skills

The majority of physiotherapists demonstrated acceptable use of the self-determination theory skill scores during training (median 5.0; interquartile range 1.3; min-max 2.0-6.0), with only two physiotherapists scoring <4.

Behavior

Of the six primary care sites that agreed to implement the intervention, five completed delivery and one site ceased delivery after Session 3 due to poor client attendance. Physiotherapists delivered SOLAS to a median of 4.0 (interquartile range 4; range 3-8) participants per class. The total mean %fidelity score (93.5%; SD 4.9%) and the overall fidelity scores were high (≥80%) (Multimedia Appendix 7). The difference between the actual and intended duration of all sessions was not statistically significant, apart from the education component of Session 1, which was significantly longer than the protocol (*P*=.03, Z=–2.23).

Physiotherapists delivered SOLAS in a needs supportive manner consistent with a self-determination theory-based communication style (HCCQ: median 5.2, interquartile range 1.3, min-max 3.7-5.8; CCBS: median 6.6, interquartile range 1.0, min-max 5.6-7.0; Table 4). The SOLAS scale results demonstrated that physiotherapists implemented the intervention overall with acceptable competence (median 4.5, interquartile range 1.2, min-max 2.8-4.8; Table 4). The median scores of

XSL•FO
RenderX

only 2 of the 15 self-determination theory strategies were delivered below the competence level during both the education and exercise components of the intervention (ie, *use support and encouragement rather than pressurising behaviors* and *acknowledge patient's feelings and perspectives*).

Implementation Outcomes

Posttraining Feedback Questionnaire

The median scores for physiotherapists' ratings of the acceptability, appropriateness, and sustainability of E-SOLAS training to support delivery of the SOLAS intervention were high (Multimedia Appendix 4). All physiotherapists reported that E-SOLAS could be used as a training method in primary care, with 100% of respondents (n=13) recommending it to other primary care physiotherapists and the majority expressing a preference for e-learning alone (69.2%, n=9) over blended learning (30.8%, n=4).

Postdelivery Qualitative Interviews

Five of the seven physiotherapists who delivered the SOLAS intervention were interviewed within 2 weeks of program completion. Ten themes were identified from the analysis of participant interview data (Multimedia Appendix 8).

Table 4. Physiotherapists' use of the self-determination theory-based communication strategies during implementation of Session 4 of the Self-management of Osteoarthritis and Low back pain through Activity and Skills.

SDT[a]–based communication strategies[b]	Overall class median score (interquartile range), min-max	Education component median score (interquartile range), min-max	Exercise component median score (interquartile range), min-max	Z score[c]	*P* value
Total	4.5 (1.2), 2.8-4.8	4.3 (1.3), 3.2-5.2	4.5 (2.1), 2.3-5.2	–0.13	.89
Autonomy support					
Offer a meaningful rationale	4.5 (1.8), 3.5-6.5	5.0 (1.5), 4.0-6.0	5.0 (2.5), 3.0-7.0	0	1.00
Provide opportunities for patient input and choice	4.0 (1.3), 3.5-5.5	4.0 (1.0), 3.0-4.0	5.0 (2.0), 4.0-7.0	–1.84	.07
Use support and encouragement rather than pressurising behaviours	3.0 (0.8), 2.0-3.0	3.0 (1.0), 3.0-4.0	2.0 (1.0), 1.0-3.0	–1.89	.06
Structure					
Set clear expectations and provide direction	3.0 (3.5), 1.0-6.5	5.0 (2.0), 4.0-6.0	4.0 (4.0), 1.0-6.0	–1.08	.27
Goal setting					
Review goal setting	3.5 (1.5), 3.0-5.0	5.0 (2.0), 4.0-6.0	3.0 (3.5), 1.0-5.0	–1.62	.10
Collaborative goal setting	5.0 (2.3), 2.5-5.5	4.0 (2.0), 3.0-6.0	4.0 (3.5), 1.0-6.0	–0.27	.78
Collaborative action planning	2.5 (2.5), 1.0-5.5	4.0 (4.0), 1.0-5.0	1.0 (4.0), 1.0-6.0	–0.55	.58
Collaborative barrier identification	4.5 (2.5), 2.5-5.5	4.0 (3.5), 1.0-7.0	5.0 (3.0), 1.0-6.0	–0.13	.89
Collaborative problem solving	4.5 (2.5), 2.0-5.5	4.0 (3.5), 1.0-7.0	5.0 (3.5), 1.0-6.0	–0.27	.78
Provide positive encouragement	5.0 (0.8), 3.5-5.0	5.0 (1.5), 3.0-5.0	5.0 (1.0), 4.0-6.0	–1.34	.18
Provide positive, information-rich feedback	5.5 (2.3), 3.5-6.5	6.0 (2.5), 3.0-6.0	5.0 (2.0), 4.0-7.0	–1.00	.31
Provide opportunities for patient practice	6.0 (1.0), 4.0-6.0	Not applicable	6.0 (1.0), 4.0-6.0	Not tested	Not tested
Interpersonal involvement					
Acknowledge patients' feelings and perspectives	2.5 (2.5), 1.0-5.0	3.0 (3.5), 1.0-5.0	1.0 (4.0), 1.0-6.0	–0.18	.85
Build relationships					
Active listening	5.5 (3.0), 1.5-6.0	6.0 (2.5), 2.0-6.0	5.0 (3.5), 1.0-6.0	–1.73	.08
Interest in patients	5.5 (1.0), 4.0-6.0	5.0 (1.0), 4.0-6.0	6.0 (1.0), 4.0-6.0	–1.73	.08

[a]SDT: self-determination theory.

[b]Scale range: 1 (not at all good) to 7 (very good).

[c]Z score from Wilcoxon signed-ranks test to assess differences between the education and exercise components across all classes.

XSL•FO
RenderX

Acceptability of E-SOLAS

Physiotherapists reported that they had a very positive experience with E-SOLAS training and felt that it was an acceptable and valuable method of training. A number of physiotherapists emphasized the convenience and flexibility of e-learning as a method of training as compared to face-to-face training. The format of training with gate-level assessments and the resource materials contained in E-SOLAS were also viewed positively.

Appropriateness of E-SOLAS

All five physiotherapists were positive about the appropriateness of the E-SOLAS content and resources in meeting their practical needs and the needs of their clients in preparing them to deliver SOLAS using needs supportive communication. One physiotherapist also reported gaining greater confidence in managing clients beyond the class setting, whereas another physiotherapist added that the e-learning format had the advantage of allowing her to reflect on learning new skills in relation to the autonomy-supporting style of delivery.

Feasibility of E-SOLAS

Demand

Although all physiotherapists reported spending additional time reviewing the E-SOLAS content and resources in preparation for delivery of the intervention, they felt the additional time was important for the first delivery of any new program and would reduce with subsequent deliveries (Multimedia Appendix 9).

Adaptation

None of the physiotherapists reported deviating from the training specifications; however, all physiotherapists made recommendations for future adaptations to either the E-SOLAS content or training format. These included providing additional resources to guide physiotherapists in educating clients about the health risks associated with the overuse of pain medications, healthy eating guidelines, a wider range of exercise options, and the provision of outcome measures for clinicians to evaluate the intervention independently. Proposed adaptations to the training format included giving participants an estimate of the time required to complete each level and additional e-learning training and blended learning (ie, small-group face-to-face coaching alongside E-SOLAS) to support delivery of the self-determination theory-based communication strategies during goal setting and action-planning activities.

Fidelity to E-SOLAS

All physiotherapists aimed to deliver the intervention content and self-determination theory-based needs supportive communication with high fidelity.

Sustainability of E-SOLAS

Overall, physiotherapists were positive about the potential for integration of E-SOLAS into existing primary care settings to support the sustained use of the SOLAS intervention as a treatment and reported plans to continue implementation in their service area. One physiotherapist proposed training a designated clinician in each primary care area through face-to-face training, who would act as a peer mentor to support colleagues who completed E-SOLAS to specifically deliver the self-determination theory-based communication strategies.

Discussion

Overview

The overall aim of this study was to develop and evaluate an e-learning training program to support physiotherapists to deliver the SOLAS intervention in a primary care setting. The effectiveness of E-SOLAS on physiotherapists' knowledge, skills, and delivery of the SOLAS intervention was assessed alongside the acceptability and feasibility of the training program. Specifically, results indicated that physiotherapists' knowledge and confidence increased from pretraining to posttraining assessment and physiotherapists' behavior was positively influenced by E-SOLAS training, as the SOLAS intervention content and theory-based communication style were delivered with high fidelity. Finally, implementation outcomes posttraining and from the qualitative interviews were overtly positive with regard to the acceptability, appropriateness, feasibility, and potential for future integration of E-SOLAS into existing primary care health services.

Effectiveness of the E-SOLAS Training Program

Physiotherapists' confidence in the SOLAS intervention content and self-determination theory-based communication style increased posttraining, which is important because it can indicate how likely a learner (in this case, physiotherapist) is to engage in the required behavior [38,39]. Knowledge also increased for the SOLAS intervention content, but there were limited changes in knowledge of the self-determination theory-based communication strategies. This may be explained by physiotherapists' high pretraining knowledge levels, suggesting a ceiling effect, as they were highly experienced and the majority had undertaken communication-based training previously, thus limiting their potential for future improvement [40]. These findings mirror our previously published evaluation of SOLAS face-to-face training [15]. The majority of physiotherapists also demonstrated acceptable competence in relation to their skills; however, two physiotherapists were rated below the competence level. Review of the audio recordings revealed that one recording was very short (<2 minutes), and thus, it was difficult to assess it in a meaningful way. Although guidelines were provided on how to conduct the role play, no guidance was given on its duration. Despite the difficulty in prescribing a set amount of time, a minimum time period could have been set to ensure a meaningful assessment, which could be applied for future iterations of E-SOLAS.

In terms of behavior, physiotherapists delivered the intervention as intended, adhering to the intervention content and delivery in a manner consistent with the self-determination theory-based communication style. The mean high fidelity to intervention content based on physiotherapists' self-reported checklists was 93%, which aligns with the findings of the previous feasibility trial [14]. For assessment of the self-determination theory-based communication style, the physiotherapists' scores on the two global measures aligned closely with face-to-face training. More specifically, for the HCCQ, the median score was 5.2 in this

XSL•FO
RenderX

study and 5.3 (on a 7-point Likert scale) in the face-to-face training study [15]. Scores on the CCBS were consistent for both studies. However, an intervention-specific measure of needs support enables a more focused look at contextual elements [41]. Here, there was some divergence between e-training and face-to-face training, with median scores of 4.5 and 4.0 on a 7-point Likert scale, respectively, favoring E-SOLAS [15].

In this study, two self-determination theory-based communication strategies (*use support and encouragement rather than pressurising behaviours* and *acknowledge patients' feelings and perspectives*) were delivered with low competence across both the education and exercise components of the class, highlighting the need for further training or adaption to E-SOLAS to further support these strategies. Interestingly, the communication strategies related to goal setting, action planning, and problem solving were delivered to a higher level of competence than the face-to-face training [15]. This may have been due to the additional interactive elements added to E-SOLAS to address the concerns identified by the physiotherapists during the development phase. Furthermore, this improvement in goal setting-related strategies may have inadvertently reduced competence in the communication strategy *use support and encouragement rather than pressurising behaviours*, as emphasis was placed on physiotherapists being more directive with clients regarding goal setting in situations where clients were unable to articulate or formulate a goal themselves. Recent research has highlighted the difficulty in applying effective goal setting in clinical settings [42], and future training programs need to consider these strategies carefully. Overall, E-SOLAS training seems at least as effective as face-to-face training in developing physiotherapists' knowledge, confidence, and ability to deliver the intervention as intended [22].

Implementation Outcomes for the E-SOLAS Training Program

Physiotherapists were very positive about E-SOLAS following training and delivery and believed it was an acceptable, appropriate, feasible, and sustainable method of training in primary care. Participants spent a mean of 9 hours completing the training over 16 days while working at their own pace and predominantly in their own time, which has clear advantages over the 12-hour face-to-face training time in addition to travel, cost, and time off work experienced by physiotherapists in our previous feasibility trial. E-SOLAS participants demonstrated high levels of engagement with training, including a 100% completion rate within the specified timeframe. This may reflect the physiotherapist-recognized importance of group-based self-management programs for busy primary care settings as well as the emphasis HCPs now place on a client-centered communication style and the acquisition of behavior-change skills [17]. Furthermore, these high levels of physiotherapists' satisfaction and engagement could also reflect the systematic and inclusive process used to develop the E-SOLAS training program according to the recommendations of the Medical Research Council [12,43].

In terms of feasibility, technical difficulties can sometimes hamper the success of e-learning with HCPs [44]. Six of the 13 physiotherapists reported difficulty accessing online materials. Therefore, it is important to ensure that technical support is in place to maintain user engagement. One of the main advantages of e-learning is flexibility and control of the time and location for program completion [45], as demonstrated in this study, wherein the majority of participants completed E-SOLAS outside work.

Although physiotherapists were satisfied with the program overall, there were some adaptations suggested, including provision of further information to support the delivery of certain education components, inclusion of details of the estimated time to complete training, and the use of blended learning. These suggestions are in line with the general recommendations for e-learning programs should be tailored to HCPs' particular knowledge and experience [46]. For example, in the context of E-SOLAS, one physiotherapist may want more information on pain medication, whereas another might like additional videos of communication strategies [45]. Such individualized learning pathways may lead to not only a more engaged learner with enhanced knowledge but also more effective delivery of the intervention.

Despite the high rate of planned implementation of the SOLAS intervention posttraining, the program was fully delivered by six physiotherapists at five sites across four primary care areas. The main reasons for nonimplementation were beyond the scope of the study and were related to the nonavailability of staff. Of the five sites with full implementation, there was an equal mix of sole and shared delivery, in contrast to the previous feasibility trial where all physiotherapists delivered the intervention independently [15]. Physiotherapists who delivered the intervention implemented it with high fidelity, apart from the education component of Session 1, which is consistent with the findings of face-to-face training [14]. Although the qualitative interview findings did not suggest any significant barriers to future implementation by a sole practitioner following training, the suggestion of a local peer mentor and the development of blended learning may be warranted to overcome this potential obstacle.

Strengths and Limitations

The major strengths of this study are its focus on program development and evaluation within a group of experienced physiotherapists who received e-learning training while working within their primary care setting. Specifically, E-SOLAS was developed and underpinned by theory, with a clear rationale about how the intervention components were developed and adapted. The use of a formal evaluation model [13] allowed for a more comprehensive understanding of the effectiveness of E-SOLAS training, including the objective evaluation of physiotherapists' behavior during training, which is frequently absent from assessments [23,47,48]. Furthermore, the application of the World Health Organization's implementation outcomes using mixed methods enabled a comprehensive assessment of the feasibility of implementation of this e-learning training program and required adaptations to increase acceptability [24]. Finally, the assessment of fidelity of

XSL•FO
RenderX

intervention delivery using validated measures following e-learning has been rarely reported in the literature and is one of the novel aspects of this study.

A few limitations of this study should be acknowledged. Owing to the relatively small sample size, particularly for the delivery phase of the study, further investigation in a larger sample is warranted. Although a nonvalidated feedback measure was used to evaluate some training and implementation outcomes, its components were informed by a framework for the evaluation of technology-enhanced learning [25] and our face-to-face training feedback measure [15]. Although physiotherapists' competence to deliver the SOLAS intervention was assessed posttraining, there was no pretraining assessment of their skills, which should be included in future studies [46]. Future studies should also incorporate some form of client measurement to more clearly understand the efficacy of this training approach. Self-report checklists were used to assess the fidelity to intervention content, which is less robust than other methods such as independently rated audio recordings [14]. Any future research evaluating a new program should apply robust fidelity-assessment methods to all parts of the intervention [36,49]. Finally, the role-play activities were an important part of E-SOLAS training; however, they were designed as one-on-one interactions (ie, between the physiotherapist and one client). Therefore, physiotherapists did not get an opportunity to practice their delivery of the intervention in a group setting prior to implementation. Future programs should try to ensure that all elements of the intervention are accurately reflected in the training program.

Conclusions

The comprehensive evaluation reported in this study provides preliminary evidence of the effectiveness, acceptability, and feasibility of an e-learning program to train physiotherapists to deliver a group-based self-management intervention in a primary care setting that is equivalent to face-to-face training. These findings will inform the development and implementation of a definitive trial and support its scalability to the wider primary care system.

Acknowledgments

The authors wish to thank the primary care physiotherapists and patients who participated in the study; William Fox and Isabelle Jeffares who supported the training phase of the study; and University College Dublin BSc Physiotherapy students Aine Moynan, Anna Scott, and Jessica Gayer for analysis of intervention content fidelity data. This article presents independent research funded by the Health Research Board (Knowledge Exchange and Dissemination Scheme 2015-1637).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of the process used to develop the E-SOLAS program. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 23KB - jmir_v21i3e11123_app1.pdf]

Multimedia Appendix 2

E-SOLAS training program structure and content. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 40KB - jmir_v21i3e11123_app2.pdf]

Multimedia Appendix 3

Outcome measures for evaluation of E-SOLAS training and implementation of the SOLAS intervention. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 21KB - jmir_v21i3e11123_app3.pdf]

Multimedia Appendix 4

Feedback measures used to record physiotherapists' reaction to E-SOLAS training and implementation outcomes posttraining. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 57KB - jmir_v21i3e11123_app4.pdf]

XSL•FO
RenderX

Multimedia Appendix 5

Pre- and posttraining questionnaire for evaluation of physiotherapists' self-reported knowledge and perceived confidence of the SOLAS intervention and self-determination theory-based communication strategies. SOLAS: Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 55KB - jmir_v21i3e11123_app5.pdf]

Multimedia Appendix 6

Curatr analytics results.

[PDF File (Adobe PDF File), 23KB - jmir_v21i3e11123_app6.pdf]

Multimedia Appendix 7

Fidelity to delivery of SOLAS intervention content. SOLAS: Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 32KB - jmir_v21i3e11123_app7.pdf]

Multimedia Appendix 8

Acceptability and appropriateness of E-SOLAS training: Themes and theme examples from participant interviews. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 53KB - jmir_v21i3e11123_app8.pdf]

Multimedia Appendix 9

Feasibility of E-SOLAS training: Themes and theme examples from participant interviews. E-SOLAS: E-learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills.

[PDF File (Adobe PDF File), 62KB - jmir_v21i3e11123_app9.pdf]

References

1. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, European League Against Rheumatism (EULAR). EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013 Jul;72(7):1125-1135 [FREE Full text] [doi: 10.1136/annrheumdis-2012-202745] [Medline: 23595142]
2. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012 Apr;64(4):465-474. [Medline: 22563589]
3. National Institute for Health and Care Excellence. Osteoarthritis: care and management in adults URL: https://www.nice.org.uk/guidance/cg177 [accessed 2018-05-23] [WebCite Cache ID 6zdCUXCYH]
4. National Institute for Health and Care Excellence. Low back pain and sciatica in over 16s: assessment and management URL: https://www.nice.org.uk/guidance/ng59 [accessed 2018-05-23] [WebCite Cache ID 6zdCClV9T]
5. Du S, Yuan C, Xiao X, Chu J, Qiu Y, Qian H. Self-management programs for chronic musculoskeletal pain conditions: a systematic review and meta-analysis. Patient Educ Couns 2011 Dec;85(3):e299-e310. [doi: 10.1016/j.pec.2011.02.021] [Medline: 21458196]
6. Oliveira VC, Ferreira PH, Maher CG, Pinto RZ, Refshauge KM, Ferreira ML. Effectiveness of self-management of low back pain: systematic review with meta-analysis. Arthritis Care Res (Hoboken) 2012 Nov;64(11):1739-1748 [FREE Full text] [doi: 10.1002/acr.21737] [Medline: 22623349]
7. Toomey E, Currie-Murphy L, Matthews J, Hurley DA. The effectiveness of physiotherapist-delivered group education and exercise interventions to promote self-management for people with osteoarthritis and chronic low back pain: a rapid review part I. Man Ther 2015 Apr;20(2):265-286. [doi: 10.1016/j.math.2014.10.013] [Medline: 25466293]
8. Hurley DA, Murphy LC, Hayes D, Hall AM, Toomey E, McDonough SM, et al. Using intervention mapping to develop a theory-driven, group-based complex intervention to support self-management of osteoarthritis and low back pain (SOLAS). Implement Sci 2016 Apr 26;11:56 [FREE Full text] [doi: 10.1186/s13012-016-0418-2] [Medline: 27113575]
9. Eldredge LKB, Parcel GS, Kok G, Gottlieb NH, FernÃ¡ndez ME. Planning Health Promotion Programs: An Intervention Mapping Approach. San Francisco: Jossey-bass; 2019.
10. Hurley DA, Hall AM, Currie-Murphy L, Pincus T, Kamper S, Maher C, SOLAS Trial Team. Theory-driven group-based complex intervention to support self-management of osteoarthritis and low back pain in primary care physiotherapy: protocol

XSL•FO
RenderX

for a cluster randomised controlled feasibility trial (SOLAS). BMJ Open 2016 Jan 21;6(1):e010728 [FREE Full text] [doi: 10.1136/bmjopen-2015-010728] [Medline: 26801470]
11. Ng JYY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-Determination Theory Applied to Health Contexts: A Meta-Analysis. Perspect Psychol Sci 2012 Jul;7(4):325-340. [doi: 10.1177/1745691612447309] [Medline: 26168470]
12. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]
13. Kirkpatrick DKJ, Kirkpatrick JD. Evaluating Training Programs: The Four Levels (3rd Edition). San Francisco, CA: Berrett-koehler Publishers; 2019.
14. Toomey E, Matthews J, Hurley DA. Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain. BMJ Open 2017 Aug 04;7(8):e015452 [FREE Full text] [doi: 10.1136/bmjopen-2016-015452] [Medline: 28780544]
15. Keogh A, Matthews J, Segurado R, Hurley DA. Feasibility of Training Physical Therapists to Deliver the Theory-Based Self-Management of Osteoarthritis and Low Back Pain Through Activity and Skills (SOLAS) Intervention Within a Trial. Phys Ther 2018 Feb 01;98(2):95-107. [doi: 10.1093/ptj/pzx105] [Medline: 29088437]
16. Richmond H, Hall AM, Hansen Z, Williamson E, Davies D, Lamb SE. Using mixed methods evaluation to assess the feasibility of online clinical training in evidence based interventions: a case study of cognitive behavioural treatment for low back pain. BMC Med Educ 2016 Jun 18;16:163 [FREE Full text] [doi: 10.1186/s12909-016-0683-4] [Medline: 27316705]
17. Health Service Executive. Making Every Contact Count: A Health Behaviour Change Framework and Implementation Plan for Health Professionals in the Irish Health Service. In: Health Service Executive. Dublin, Ireland: Health Service Executive; 2016.
18. National Undergraduate Curriculum Working Group. National Undergraduate Curriculum for Chronic Disease Prevention and Management Part 1: Making Every Contact Count for Health Behaviour Change. In: Health Service Executive. Dublin, Ireland: Health Service Executive; 2017.
19. Preston E, Ada L, Dean CM, Stanton R, Waddington G, Canning C. The Physiotherapy eSkills Training Online resource improves performance of practical skills: a controlled trial. BMC Med Educ 2012 Nov 26;12:119 [FREE Full text] [doi: 10.1186/1472-6920-12-119] [Medline: 23176318]
20. Gardner P, Slater H, Jordan JE, Fary RE, Chua J, Briggs AM. Physiotherapy students' perspectives of online e-learning for interdisciplinary management of chronic health conditions: a qualitative study. BMC Med Educ 2016 Feb 16;16:62 [FREE Full text] [doi: 10.1186/s12909-016-0593-5] [Medline: 26879982]
21. Richmond H, Copsey B, Hall AM, Davies D, Lamb SE. A systematic review and meta-analysis of online versus alternative methods for training licensed health care professionals to deliver clinical interventions. BMC Med Educ 2017 Nov 23;17(1):227 [FREE Full text] [doi: 10.1186/s12909-017-1047-4] [Medline: 29169393]
22. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. JAMA 2008 Sep 10;300(10):1181-1196. [doi: 10.1001/jama.300.10.1181] [Medline: 18780847]
23. Sinclair PM, Kable A, Levett-Jones T, Booth D. The effectiveness of Internet-based e-learning on clinician behaviour and patient outcomes: A systematic review. Int J Nurs Stud 2016 May;57:70-81. [doi: 10.1016/j.ijnurstu.2016.01.011] [Medline: 27045566]
24. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. BMJ 2013 Nov 20;347:f6753. [Medline: 24259324]
25. Cook DA, Ellaway RH. Evaluating technology-enhanced learning: A comprehensive framework. Med Teach 2015;37(10):961-970. [doi: 10.3109/0142159X.2015.1009024] [Medline: 25782599]
26. Curatr. HT2 Labs URL: https://www.ht2labs.com/curatr/ [accessed 2019-01-03] [WebCite Cache ID 759e4xH0h]
27. Adams P. Exploring social constructivism: theories and practicalities. Education 3-13 2006 Oct;34(3):243-257. [doi: 10.1080/03004270600898893]
28. Zimmerman BJ. Self-Regulated Learning and Academic Achievement: An Overview. Educational Psychologist 1990 Jan;25(1):3-17. [doi: 10.1207/s15326985ep2501_2]
29. Ryan D, Deci E. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
30. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
31. 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]
32. Borrelli B. The Assessment, Monitoring, and Enhancement of Treatment Fidelity In Public Health Clinical Trials. J Public Health Dent 2011;71(s1):S52-S63 [FREE Full text] [doi: 10.1111/j.1752-7325.2011.00233.x] [Medline: 21499543]
33. Taylor IM, Ntoumanis N, Standage M. A self-determination theory approach to understanding the antecedents of teachers' motivational strategies in physical education. J Sport Exerc Psychol 2008 Feb;30(1):75-94. [Medline: 18369244]

XSL•FO
RenderX

34. Williams GC, Deci EL. Internalization of biopsychosocial values by medical students: a test of self-determination theory. J Pers Soc Psychol 1996 Apr;70(4):767-779. [Medline: 8636897]
35. Bartholomew KJ, Ntoumanis N, Thøgersen-Ntoumani C. The controlling interpersonal style in a coaching context: development and initial validation of a psychometric scale. J Sport Exerc Psychol 2010 Apr;32(2):193-216. [Medline: 20479478]
36. Toomey E, Matthews J, Guerin S, Hurley DA. Development of a Feasible Implementation Fidelity Protocol Within a Complex Physical Therapy-Led Self-Management Intervention. Phys Ther 2016 Aug;96(8):1287-1298. [doi: 10.2522/ptj.20150446] [Medline: 26939605]
37. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
38. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
39. Ajzen I. The theory of planned behaviour: reactions and reflections. Psychol Health 2011 Sep;26(9):1113-1127. [doi: 10.1080/08870446.2011.613995] [Medline: 21929476]
40. Su Y, Reeve J. A Meta-analysis of the Effectiveness of Intervention Programs Designed to Support Autonomy. Educ Psychol Rev 2010 Sep 25;23(1):159-188. [doi: 10.1007/s10648-010-9142-7]
41. Fortier MS, Duda JL, Guerin E, Teixeira PJ. Promoting physical activity: development and testing of self-determination theory-based interventions. Int J Behav Nutr Phys Act 2012 Mar 02;9:20 [FREE Full text] [doi: 10.1186/1479-5868-9-20] [Medline: 22385751]
42. Rose A, Rosewilliam S, Soundy A. Shared decision making within goal setting in rehabilitation settings: A systematic review. Patient Educ Couns 2017 Jan;100(1):65-75. [doi: 10.1016/j.pec.2016.07.030] [Medline: 27486052]
43. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015;350:h1258 [FREE Full text] [Medline: 25791983]
44. Eley R, Fallon T, Soar J, Buikstra E, Hegney D. The status of training and education in information and computer technology of Australian nurses: a national survey. J Clin Nurs 2008 Oct;17(20):2758-2767. [doi: 10.1111/j.1365-2702.2008.02285.x] [Medline: 18808645]
45. Fontaine G, Cossette S, Heppell S, Boyer L, Mailhot T, Simard M, et al. Evaluation of a Web-Based E-Learning Platform for Brief Motivational Interviewing by Nurses in Cardiovascular Care: A Pilot Study. J Med Internet Res 2016 Aug 18;18(8):e224 [FREE Full text] [doi: 10.2196/jmir.6298] [Medline: 27539960]
46. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. Acad Med 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]
47. Nicoll P, MacRury S, van Woerden HC, Smyth K. Evaluation of Technology-Enhanced Learning Programs for Health Care Professionals: Systematic Review. J Med Internet Res 2018 Apr 11;20(4):e131 [FREE Full text] [doi: 10.2196/jmir.9085] [Medline: 29643049]
48. Jacobs ZG, Elnicki DM, Perera S, Weiner DK. An E-learning Module on Chronic Low Back Pain in Older Adults: Effect on Medical Resident Attitudes, Confidence, Knowledge, and Clinical Skills. Pain Med 2018 Jun 1;19(6):1112-1120. [doi: 10.1093/pm/pnx333] [Medline: 29315426]
49. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004 Sep;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]

Abbreviations

CCBS: Controlling Coach Behaviour Scale
E-SOLAS: E-Learning training program for Self-management of Osteoarthritis and Low back pain through Activity and Skills
HCCQ: Health Care Climate Questionnaire
HCP: health care practitioner
ID: participant identification number
SOLAS: Self-management of Osteoarthritis and Low back pain through Activity and Skills

XSL•FO
RenderX

Edited by G Eysenbach; submitted 24.05.18; peer-reviewed by K Knittle, M El Tantawi; comments to author 30.09.18; revised version received 20.11.18; accepted 10.12.18; published 07.03.19.

Please cite as:
Hurley DA, Keogh A, Mc Ardle D, Hall AM, Richmond H, Guerin S, Magdalinski T, Matthews J
Evaluation of an E-Learning Training Program to Support Implementation of a Group-Based, Theory-Driven, Self-Management Intervention For Osteoarthritis and Low-Back Pain: Pre-Post Study
J Med Internet Res 2019;21(3):e11123
URL: https://www.jmir.org/2019/3/e11123/
doi:10.2196/11123
PMID:30843863

©Deirdre A Hurley, Alison Keogh, Danielle Mc Ardle, Amanda M Hall, Helen Richmond, Suzanne Guerin, Tara Magdalinski, James Matthews. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Clinical Virtual Simulation in Nursing Education: Randomized Controlled Trial

José Miguel Padilha[1], PhD; Paulo Puga Machado[2], PhD; Ana Ribeiro[2], PhD; José Ramos[3], MSc; Patrício Costa[4], PhD

[1]Nursing School of Porto; CINTESIS – Tech4edusim, Porto, Portugal

[2]Nursing School of Porto; CINTESIS – NursID, Porto, Portugal

[3]Nursing School of Porto, Porto, Portugal

[4]Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, ICVS / 3B's–PT Government Associate Laboratory, Braga / Guimarães, Portugal, Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal

Corresponding Author:
José Miguel Padilha, PhD
Nursing School of Porto; CINTESIS – Tech4edusim
Street Dr António Bernardino de Almeida
Porto, 4200-072
Portugal
Phone: 351 225 073 500
Fax: 351 225 096 337
Email: miguelpadilha@esenf.pt

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2019/6/e14155/

Abstract

Background: In the field of health care, knowledge and clinical reasoning are key with regard to quality and confidence in decision making. The development of knowledge and clinical reasoning is influenced not only by students' intrinsic factors but also by extrinsic factors such as satisfaction with taught content, pedagogic resources and pedagogic methods, and the nature of the objectives and challenges proposed. Nowadays, professors play the role of learning facilitators rather than simple "lecturers" and face students as active learners who are capable of attributing individual meanings to their personal goals, challenges, and experiences to build their own knowledge over time. Innovations in health simulation technologies have led to clinical virtual simulation. Clinical virtual simulation is the recreation of reality depicted on a computer screen and involves real people operating simulated systems. It is a type of simulation that places people in a central role through their exercising of motor control skills, decision skills, and communication skills using virtual patients in a variety of clinical settings. Clinical virtual simulation can provide a pedagogical strategy and can act as a facilitator of knowledge retention, clinical reasoning, improved satisfaction with learning, and finally, improved self-efficacy. However, little is known about its effectiveness with regard to satisfaction, self-efficacy, knowledge retention, and clinical reasoning.

Objective: This study aimed to evaluate the effect of clinical virtual simulation with regard to knowledge retention, clinical reasoning, self-efficacy, and satisfaction with the learning experience among nursing students.

Methods: A randomized controlled trial with a pretest and 2 posttests was carried out with Portuguese nursing students (N=42). The participants, split into 2 groups, had a lesson with the same objectives and timing. The experimental group (n=21) used a case-based learning approach, with clinical virtual simulator as a resource, whereas the control group (n=21) used the same case-based learning approach, with recourse to a low-fidelity simulator and a realistic environment. The classes were conducted by the usual course lecturers. We assessed knowledge and clinical reasoning before the intervention, after the intervention, and 2 months later, with a true or false and multiple-choice knowledge test. The students' levels of learning satisfaction and self-efficacy were assessed with a Likert scale after the intervention.

Results: The experimental group made more significant improvements in knowledge after the intervention (P=.001; d=1.13) and 2 months later (P=.02; d=0.75), and it also showed higher levels of learning satisfaction (P<.001; d=1.33). We did not find statistical differences in self-efficacy perceptions (P=.9; d=0.054).

XSL•FO
RenderX

Conclusions: The introduction of clinical virtual simulation in nursing education has the potential to improve knowledge retention and clinical reasoning in an initial stage and over time, and it increases the satisfaction with the learning experience among nursing students.

(J Med Internet Res 2019;21(3):e11529) doi:10.2196/11529

KEYWORDS

clinical virtual simulation; nursing education; virtual patient; user-computer interface

Introduction

Nursing education

The education of nursing students has always been a challenge for governments, health educators, health managers, and the students themselves to ensure the quality and safety of learning and clinical practice.

Twenty-first century students have grown up using information and communications technologies (ICT) on a day-to-day basis. The use of ICT leads to different learning processes and information structuring processes [1].

Professors and managers should bear in mind that these students are able to access information in real time, to use parallel processes and multitask; in addition, they prefer graphics to text, they function best when networked, and they need instant gratification and frequent rewards [2].

These students' ICT skills call for innovation in the pedagogical strategies in health education underpinned by a constructivist paradigm of health education [3]. Nowadays, professors play the role of learning facilitators rather than simple "lecturers" and face students as active learners who are capable of attributing individual meanings to their personal experiences and building their own knowledge over time. An active and constructive educational environment based on challenges and learning objectives will promote deeper learning, emphasizing understanding and the application of knowledge over memorization and recall [4-8].

Innovation in simulation technologies has made available high-fidelity simulators that have supported the change in the health education paradigm. The use of high-fidelity simulators has improved the acquisition of knowledge and skills and strengthened quality and safety in clinical practice [3,9-15]. However, we have been facing challenges with the increasing cost of simulators, the difficulties of space management, and the low number of clinical scenarios available.

Clinical Virtual Simulation

Developments in digital and virtual technology have eased the way to recreating reality using virtual patients [16] depicted on a computer touchscreen (clinical virtual simulation). Clinical virtual simulation is the recreation of reality depicted on a computer screen, and it involves real people operating simulated systems. It is a type of simulation that places people in a central role through the exercising of their decision-making, motor control, and communication skills [11]. Clinical virtual simulation uses virtual patients in dynamic and immersive clinical environments ranging from prehospital environments to environments in the community (Figures 1 and 2). The concept is based on the virtual patient being accessed through a variety of multimedia, screen-based interactive [17] and dynamic patient scenarios, which are supported by physiological algorithms. Clinical virtual simulation increases interaction and feedback [18] and raises both the perception of self-efficacy and the user's satisfaction levels [19]. The use of clinical virtual simulation in the development of nursing competences improves performance [20] and competences related to psychomotor skills [21], critical thinking [22], clinical skills [23], and decision making [17].

The latest technological advances in clinical virtual simulation have improved realism and dynamic interaction, with the possibility of thousands of clinical scenarios depicted on a touchscreen table or on the Web. However, nowadays, little is known about its effectiveness with regard to students' learning satisfaction, self-efficacy, knowledge retention, and clinical reasoning, especially when using the latest advances in clinical virtual simulation.

As professors in the field of health, we are concerned about students' learning satisfaction and effective learning outcomes [13]. This study intended to assess the effectiveness of clinical virtual simulation in raising levels of learning satisfaction, self-efficacy, knowledge retention, and clinical reasoning among nursing students.

Methods

A randomized controlled trial and a prospective and analytical study was conducted between March and May 2017 with a pretest and 2 posttests.

Participants and Allocation Process

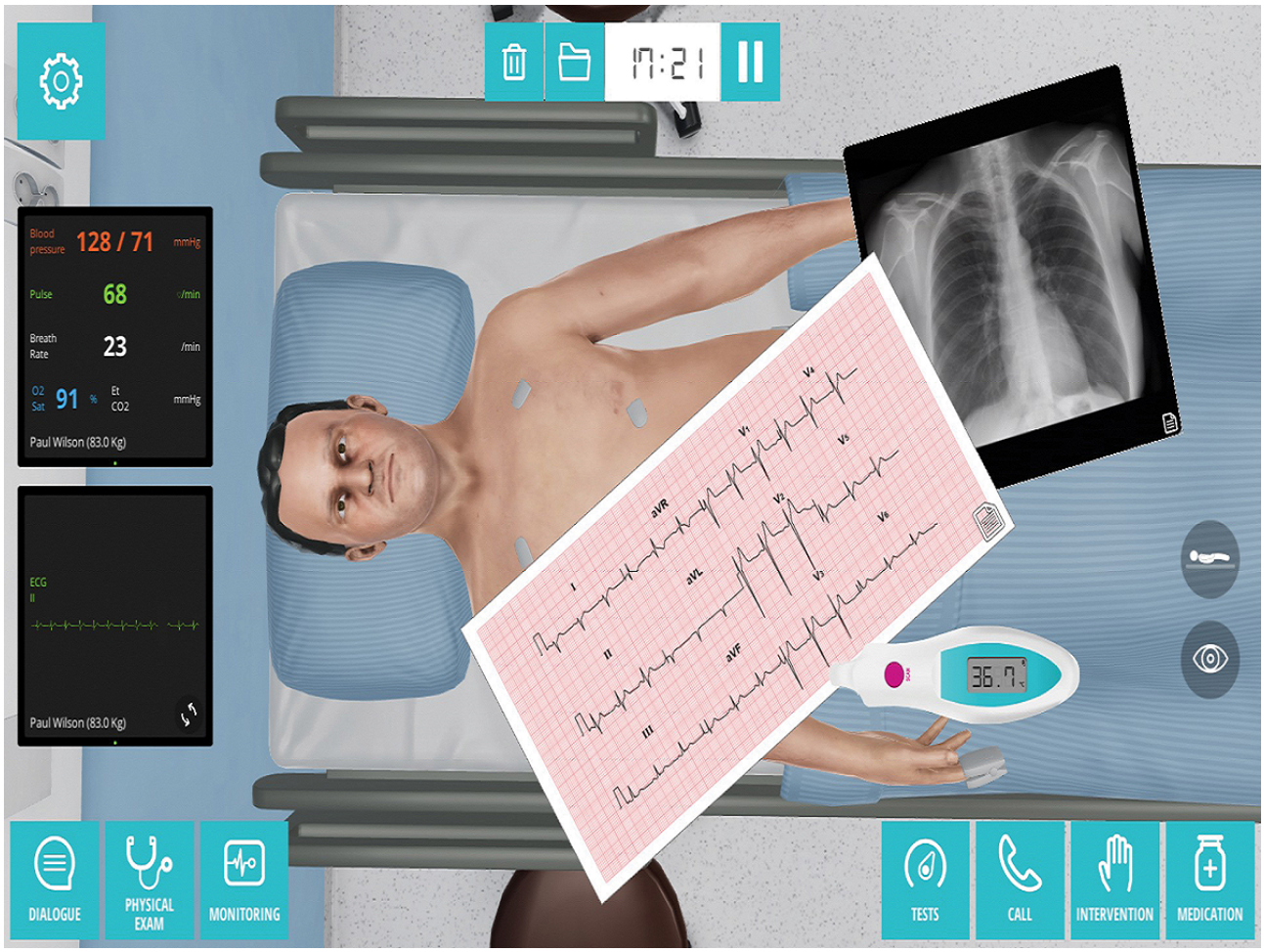
The participants were volunteer graduation students in the second year at the Nursing School of Porto in Portugal, enrolled in the course "Corporal Body Responses 1" (respiratory, cardiac, and urinary systems). This study was accomplished through an elective curriculum made available to all students. All the students enrolled in the course (N=128) were invited by email to be volunteers in the study. Those who did volunteer were invited to an initial meeting at which 56 student volunteers were present, all of whom accepted the invitation and gave informed consent. The volunteers filled out a questionnaire with sociodemographic and student data (average current course grade, number of European Credit Transfer System credits achieved as part of the nursing degree, and average grade required for admission into the degree course); these data were used in the randomization process. The anonymization of students was performed by the assignment of a number with 6 digits chosen by the student, with no possibility of the students being identified by the researchers.

XSL•FO
RenderX

The study sample size was determined considering a 1-tailed, unpaired *t* test, a type I error of 0.05, a statistical power of 0.80, and an effect size of *d*=0.80. Using G*Power3 [24] this study required a total of 42 students, 21 per group.

Students were allocated to each group through a simple random allocation using IBM SPSS Statistics version 24.

Figure 1. Clinical virtual simulation in hospital environment.



One week after the initial meeting (and after the randomization process), all the 56 volunteer students were invited to another meeting, which took place immediately before the intervention. At this second meeting, students were invited to do the first knowledge and clinical reasoning test (assessment before intervention—A0). Immediately after this, the students were directed, according to identification number (which only they were able to identify), to the classroom where they were informed about which group they had been allocated to.

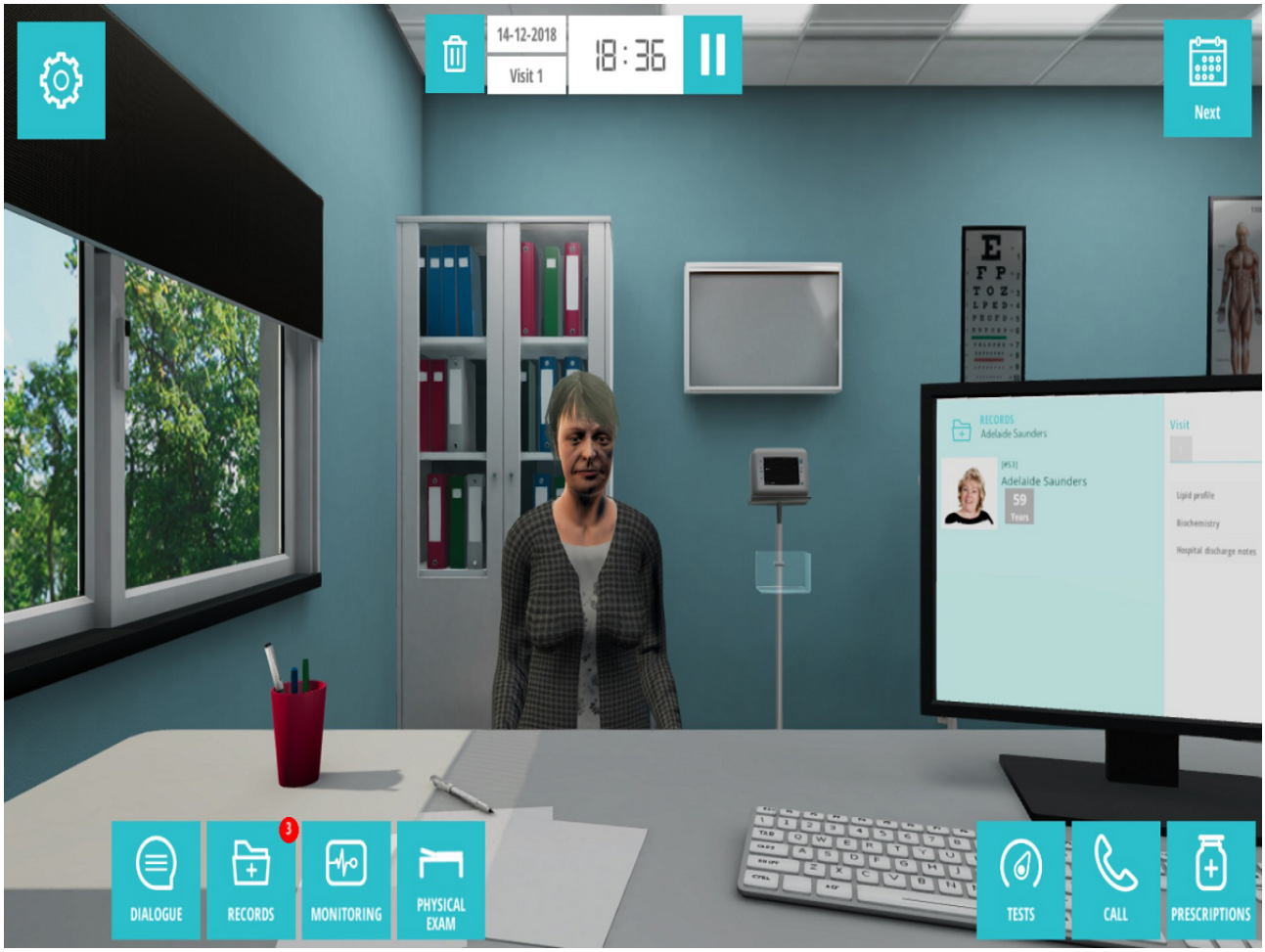
Both groups received a laboratory class of 45 min, with the aim of activating knowledge and developing clinical reasoning skills in the field of the respiratory process in relation to ineffective airway clearance and hypoxia. With the experimental group, a case-based learning approach was used, with recourse to a clinical virtual simulator scenario (Body Interact) facilitated by the regular subject teacher.

The clinical virtual simulator (Body Interact) presents virtual patients backed up by a physiological algorithm that recreates a dynamic health condition that responds to user interventions. The clinical scenario is initiated by a briefing; subsequently, the user can interact with the virtual patient through dialogues, monitoring the physiological parameters, observation and physical examination, the prescription and/or analysis of complementary examinations, and the prescription of intervention and/or pharmacological treatment. The responses to and the development of the clinical case are dynamic and conditional on the decisions taken. The closure of the clinical case is determined either by the successful resolution of the scenario or by the amount of time that has elapsed (as defined by the user). Immediately after the simulation ends, a differential diagnosis interface is presented. After the simulation has concluded, the simulator provides a debriefing tool whereby 3 categories of information can be analyzed: the simulation report, the simulation timeline, and the performance report. In the simulation report, the correct differential diagnosis and the option chosen by the user are presented. All the actions carried out and the hemodynamic consequences are presented on the timeline together with all the complementary diagnosis examinations that were requested. In the simulation report, performance scores are given for 3 categories of information: physical examination, diagnosis, and therapeutic activities. In each 1 of these categories, the decisions made and their appropriacy are presented, as well as the best decision, on the basis of the evidence. The debriefing tool also provides the

XSL•FO
RenderX

scientific references that support the clinical scenario and its optimal resolution.

Figure 2. Clinical virtual simulation in environments in the community.



The control group received a laboratory class of 45 min, with the same aim, using the same case-based learning approach but making use of a low-fidelity simulator and a realistic environment (pedagogical strategies that were already used in the nursing school), guided throughout by the regular subject teacher. For both groups, there was a simulation pedagogical strategy of briefing (5 min), simulation (20 min), and debriefing (20 min), with the same structure and contents.

Immediately after the end of the intervention (the laboratory class), all the students were invited to a second test (assessment after intervention—A1), and 2 months later, they were invited again to a third test (assessment follow-up—A2).

In all the knowledge assessments, we used the same true or false and multiple-choice test, which had been developed by the usual course lecturers. These knowledge assessments were based on features intrinsically related to the clinical reasoning applied within the specific scenario. In the assessment immediately after the intervention with both groups, we also assessed the students' satisfaction levels with the simulation, and their general perception of self-efficacy.

The assessment of student satisfaction was conducted using a Portuguese version [25] of the Learner Satisfaction with Simulation Tool [19], a 10-point Likert scale. The assessment of their perception of self-efficacy was conducted with a Portuguese version [26] of the General Self-efficacy Scale [27], a 5-point Likert scale. The Cronbach alpha coefficients of the scales have been illustrated in Table 1.

Data Analysis

We performed the Kolmogorov-Smirnov test with the Lilliefors correction to check for the normality assumption. We obtained statistically nonsignificant results for both groups in the 3 variables under study, meaning that the normality assumption was met.

The main variable under study (the development of knowledge and clinical reasoning) was obtained by the difference between the assessment before and after the intervention. Positive values reveal improvement between the 2 assessments.

To compare both groups in the relevant variables under study, we used an unpaired *t* student to compare averages.

When the homogeneity of variances assumption was violated, the Welch correction was used.

A multivariate analysis of variance (MANOVA) was performed to compare the 2 groups across the 3 measurement points.

The results were considered statistically significant for $P<.05$, and regarding effect size measures, Cohen criteria (1988) [29] were considered to rank the size of the magnitude effect (Cohen

XSL•FO
RenderX

d: 0.2—small, 0.5—medium, and 0.8—large; partial Eta-squared: 0.02—small, 0.13—medium, and 0.26—large).

Table 1. Cronbach alpha coefficients for the original, for the Portuguese versions, for this study's sample of the Learner Satisfaction with Simulation Tool, and for the General Self-efficacy Scale.

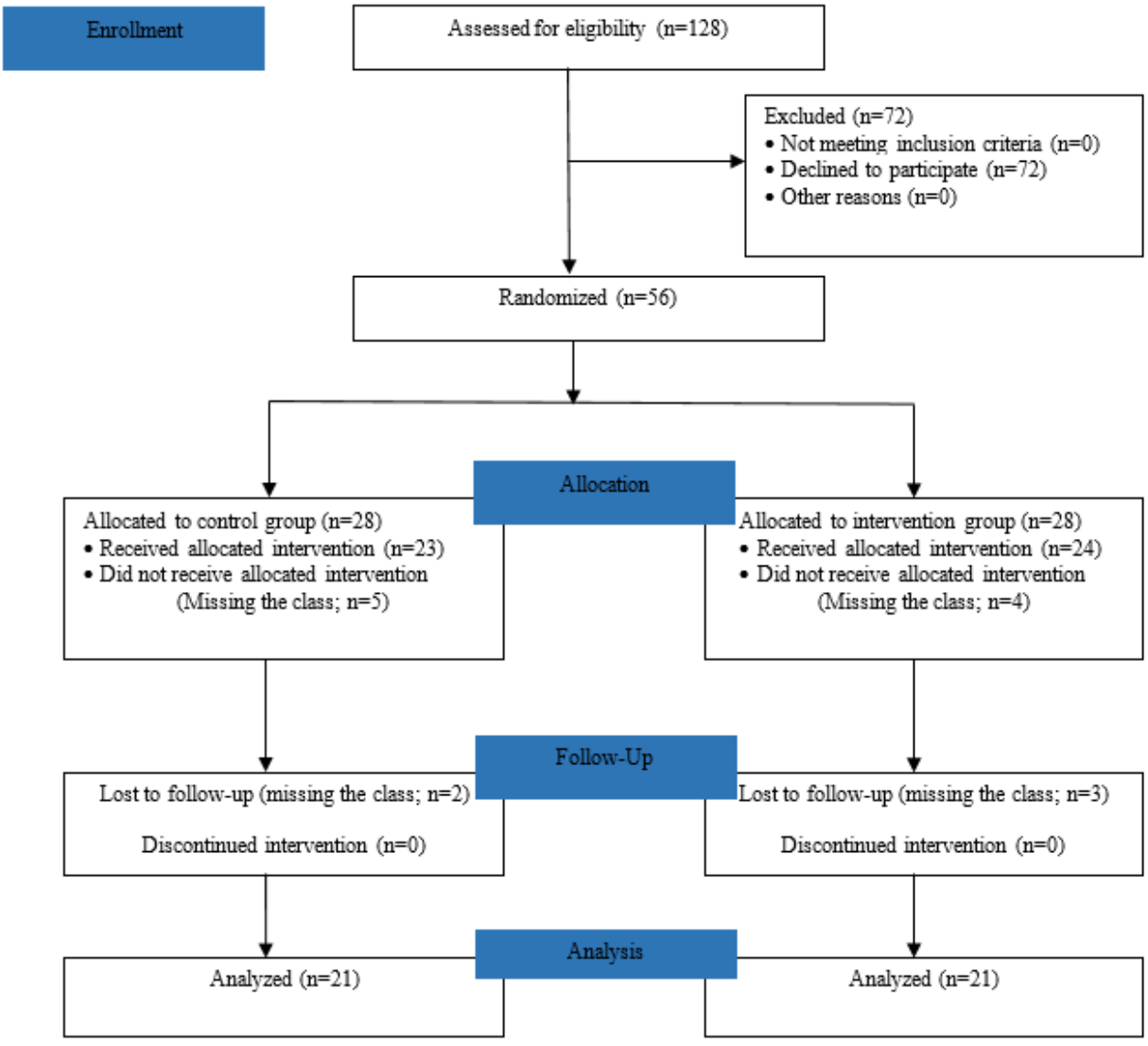
Scales	Original version, Cronbach alpha	Portuguese version		Study sample	
		Cronbach alpha	Correlation item-item total	Cronbach alpha	Correlation item-item total
Learner Satisfaction with Simulation Tool	.952	.969	.633-.823	.970	.660-.910
The General Self-Efficacy Scale (average for 25 language versions) [28]	.860	.760	.290-.530	.882	.527-.726

This study was approved by the ethics committee of the Nursing School of Porto with the number 2017/1. This randomized controlled trial does not possess a trial identifier as it is not legally required in the context of the study.

Results

A total of 42 students from the second year of a degree course participated in this study (n=21 in the experimental group and n=21 in the control group). The average age of the students was 19.9 (SD 1.99) years, and 95% (40/42) of the students were females. The flow diagram (Figure 3) represents the randomization and allocation process. Table 2 shows the results of the variables under analysis.

Figure 3. Flow diagram of sample randomization and allocation process.



XSL•FO
RenderX

Table 2. Means of sample characteristics and study variables and SDs.

Study variables	Control group	Experimental group
Sex, n		
Female	19	21
Male	2	0
Age, mean (SD)	20.29 (2.19)	19.29 (0.46)
Mean entry grade to the degree course, mean (SD)	15.54 (1.46)	15.97 (0.85)
European Credit Transfer System credits on the degree course, mean (SD)	87.29 (6.90)	86.86 (5.41)
Degree course mean grade so far, mean (SD)	13.21 (0.67)	13.42 (0.99)
Self-efficacy perception, mean (SD)	30.14 (4.29)	30.38 (4.57)
Learning Satisfaction, mean (SD)	7.47 (1.58)	9.04 (0.55)
Knowledge assessment before intervention (A0), mean (SD)	9.87 (2.24)	10.15 (1.27)
Knowledge assessment after intervention (A1), mean (SD)	10.51 (1.89)	12.47 (1.57)
Knowledge assessment follow-up (2 months; A2), mean (SD)	10.55 (1.81)	11.93 (1.84)

Knowledge Retention and Learning Satisfaction

The results of the students' *t* tests showed the existence of statistically significant differences in knowledge retention after the intervention (t_{40}=−3.656; *P*=.001; *d*=1.13), knowledge retention 2 months later (t_{40}=−2.439; *P*=.02; *d*=0.75), and in learning satisfaction (t_{40}=−4.309; *P*<.001; *d*=1.33). The students in the experimental group presented better outcomes in knowledge retention and learning satisfaction than students in the control group. The values of the Cohen *d* reinforce the magnitude effect of the intervention.

The MANOVA result was significant for time (Pillai Trace; $F_{2,39}$=13.4, *P*<.001, partial eta squared=.407) and for the interaction term time x group ($F_{2,39}$=4.45, *P*=.02, partial eta squared=.186), indicating that there are differences in the students' levels of knowledge across time and that those differences are group dependent. Differences among moments were tested through a Bonferroni test, and significant results were observed for A0-A1 (*P*<.001), for A0-A2 (*P*=.02) but not for A1-A2 (*P*>.99; A0—assessment before intervention, A1—assessment after intervention, A2—assessment follow-up). Regarding comparisons of the different groups across time, significant differences were observed for A0-A1 (*P*<.001), for A0-A2 (*P*=.01), but not for A1-A2 (*P*=.75). No significant differences were obtained for the control group (A0-A1: *P*=.44, A0-A2: *P*=.99, A1-A2: *P*>.99).

Self-Efficacy Perception

In self-efficacy perception, the results did not show statistical differences between the groups: t_{40}=−0.174, *P*=.9, *d*=0.054.

Statistically significant results were also found for the overall effect of the group at the 3 measurement points: $F_{1,40}$=10.2, *P*=.003, partial eta squared=.204. These results indicate that 20.4% of students' scores across the 3 measurement points are explained by the group to which the students were assigned.

Discussion

Principal Findings

This paper indicates that clinical virtual simulation improves knowledge retention and initial clinical reasoning over time (2 months) and improves student satisfaction with learning, without influencing the perception of general efficiency. Clinical virtual simulation enabled a 20.4% improvement in students' knowledge retention and clinical reasoning in the context of the study. This study showed that clinical virtual simulation is a pedagogical strategy that, combined with other strategies such as briefing, simulation, and debriefing, improves both initial knowledge retention and knowledge retention over time. Clinical virtual simulation also raises the level of satisfaction with the learning experience among nursing students. These results reveal the fit of clinical virtual simulation with the new generation's expectations and ways of learning. The effect of the use of clinical virtual simulation as a pedagogical strategy in improving knowledge retention and clinical reasoning and students' satisfaction levels showed a match with the features of twenty-first century nursing students. The twenty-first century nursing students had already shown high levels of usefulness, ease, and intention to use clinical virtual simulation [30]. In addition, this paper now indicates that the use of clinical virtual simulation can improve knowledge retention, clinical reasoning, and satisfaction with learning.

These results are in line with the results of other studies, where the authors found that levels of knowledge [31-33] and satisfaction [14] with the learning process improve with the use of virtual simulation.

Clinical virtual simulation brings together such strategies as gaming and problem-based learning, using an interactive and dynamic 3-dimensional technology that encourages active and critical action-based learning.

We did not find any differences in the self-efficacy perception of the students using this strategy. This is in line with the theoretical construct of Bandura's [34] self-efficacy theory, in

XSL•FO
RenderX

which the self-efficacy perception results from the interaction of different variables over time, and in this study, there was only 1 intervention with 1 class.

Clinical Virtual Simulation in Nursing Education

Clinical virtual simulation is a complementary pedagogical strategy that provides the opportunity to improve clinical reasoning skills in students through exposure to a large number of clinical scenarios. The use of clinical virtual simulation as a pedagogical strategy should be integrated and coordinated with other pedagogical strategies in classes [35,36] and with other resources, such as high-, medium-, and low-tech simulators in use in our simulation labs to maximize the development of cognitive, affective, and psychomotor skills in the students.

This study is in line with the writings of Berman and colleagues [17]. Clinical virtual simulation is an interactive learning strategy that captures students' intrinsic motivations and satisfaction, and it is focused on the application of foundational knowledge oriented toward a clinical learning challenge that recreates clinical scenarios with which students will be confronted in future clinical contexts. It allows a competency-based education and assessment that consequently enables a deep level of learning and the development of clinical expertise. Clinical virtual simulation can contribute toward reducing clinical error and improving the safety and quality of health care.

Clinical virtual simulation responds to the difficulties of managing laboratorial space, enabling teaching institutions to expand the number of clinical scenarios available for student training. Clinical virtual simulation makes training in the classroom context feasible and broadens the availability of scenarios in the Web environment, a feature that, in our experience, enables a tremendous increase in the number of students receiving individual training and a significant reduction in the costs of simulation use per student.

As limitations of this study, we identified the fact that it was only carried out in a single context, with second-year nursing students, and on a single course with content related only to the respiratory process. We also judge that the follow-up time was too short to fully evaluate the knowledge retention over time.

In light of these promising results, we suggest the replication of this study with a multicentric and prospective design on different health science courses.

Conclusions

Clinical virtual simulation is a pedagogical strategy that contributes to the improvement of knowledge retention initially and over time and increases the students' satisfaction.

This paper reveals the impact of clinical virtual simulation use in nursing education and helps professors in the field of health to be aware of its pedagogical utility and appropriacy.

These results show the potential of clinical virtual simulation to be an effective pedagogical strategy to build an educational environment that supports the development of clinical competences in the next generation of care providers, contributing toward improvements in the safety and quality of health care.

Acknowledgments

This study was supported by European Regional Development Fund through the operation POCI-01-0145-FEDER-023342 funded by the Programa Operacional Competitividade e Internacionalização-COMPETE2020 and by Portuguese National Funds through Fundação para a Ciência e a Tecnologia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 24MB - jmir_v21i3e11529_app1.pdf]

References

1. Prensky M. Digital natives, digital immigrants part 2: do they really think differently? On the Horizon 2001 Nov;9(6):1-6 [FREE Full text] [doi: 10.1108/10748120110424843]
2. Prensky M. Digital natives, digital immigrants part 1. On the Horizon 2001 Sep;9(5):1-6. [doi: 10.1108/10748120110424816]
3. Heimann C, Prado C, de Moraes RR, Vidal G, Liberal D, Oliveira G, et al. [Acquiring nursing knowledge through the constructivist method]. Rev Esc Enferm USP 2013 Aug;47(4):997-1000 [FREE Full text] [doi: 10.1590/S0080-623420130000400032] [Medline: 24310702]
4. Chi MT. Active-constructive-interactive: a conceptual framework for differentiating learning activities. Top Cogn Sci 2009 Jan;1(1):73-105 [FREE Full text] [doi: 10.1111/j.1756-8765.2008.01005.x] [Medline: 25164801]
5. Georg C, Zary N. Web-based virtual patients in nursing education: development and validation of theory-anchored design and activity models. J Med Internet Res 2014 Apr 10;16(4):e105 [FREE Full text] [doi: 10.2196/jmir.2556] [Medline: 24727709]

XSL•FO
RenderX

6. Pennaforte T, Moussa A, Loye N, Charlin B, Audétat MC. Exploring a new simulation approach to improve clinical reasoning teaching and assessment: randomized trial protocol. JMIR Res Protoc 2016 Feb 17;5(1):e26 [FREE Full text] [doi: 10.2196/resprot.4938] [Medline: 26888076]
7. 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 Sep 17;16(9):e214 [FREE Full text] [doi: 10.2196/jmir.3322] [Medline: 25230684]
8. Hege I, Kononowicz AA, Adler M. A clinical reasoning tool for virtual patients: design-based research study. JMIR Med Educ 2017 Nov 02;3(2):e21 [FREE Full text] [doi: 10.2196/mededu.8100] [Medline: 29097355]
9. Kneebone R. Evaluating clinical simulations for learning procedural skills: a theory-based approach. Acad Med 2005 Jun;80(6):549-553. [doi: 10.1097/00001888-200506000-00006] [Medline: 15917357]
10. Lapkin S, Fernandez R, Levett-Jones T, Bellchambers H. The effectiveness of using human patient simulation manikins in the teaching of clinical reasoning skills to undergraduate nursing students: a systematic review. JBI Libr Syst Rev 2010;8(16):661-694. [doi: 10.1016/j.ecns.2010.05.005] [Medline: 27820553]
11. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]
12. Yuan HB, Williams BA, Fang JB. The contribution of high-fidelity simulation to nursing students' confidence and competence: a systematic review. Int Nurs Rev 2012 Mar;59(1):26-33 [FREE Full text] [doi: 10.1111/j.1466-7657.2011.00964.x]
13. Meakim C, Boese T, Decker S, Franklin AE, Gloe D, Lioce L, et al. Standards of Best Practice: Simulation Standard I: Terminology. Clin Simul Nurs 2013 Jun;9(6):S3-S11. [doi: 10.1016/j.ecns.2013.04.001]
14. Batista R, Martins J, Pereira M, Mazzo A. [High-Fidelity Simulation in the Nursing Degree: gains perceived by students]. Rev Enf Ref 2014 Mar 15;4(1):135-144 [FREE Full text] [doi: 10.12707/RIII13169] [Medline: 22153054]
15. Pinar G, Knight C, Gaioso V, Watts P, Dailey K, Britt S, et al. The effects of high fidelity simulation on nursing students' perceptions and self-efficacy of obstetric skills. Int Arch Nurs Health Care 2015 Dec 31;1(1):1-7. [doi: 10.23937/2469-5823/1510008]
16. Lopreiato JO. Agency for Healthcare Research and Quality. Rockville, MD; 2016. Healthcare Simulation Dictionary URL: https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/quality-patient-safety/patient-safety-resources/research/simulation_dictionary/sim-dictionary.pdf [accessed 2019-02-11] [WebCite Cache ID 766yeNPXG]
17. Berman NB, Durning SJ, Fischer MR, Huwendiek S, Triola MM. The role for virtual patients in the future of medical education. Acad Med 2016 Dec;91(9):1217-1222. [doi: 10.1097/ACM.0000000000001146] [Medline: 26959224]
18. Chang KK, Chung J, Wong T. Learning intravenous cannulation: a comparison of the conventional method and the CathSim Intravenous Training System. J Clin Nurs 2002 Jan;11(1):73-78. [Medline: 11845758]
19. Roh YS, Lee W, Chung H, Park Y. The effects of simulation-based resuscitation training on nurses' self-efficacy and satisfaction. Nurse Educ Today 2013 Feb;33(2):123-128. [doi: 10.1016/j.nedt.2011.11.008] [Medline: 22153054]
20. Tsai SL, Chai S, Hsieh L, Lin S, Taur F, Sung W, et al. The use of virtual reality computer simulation in learning Port-A cath injection. Adv Health Sci Educ Theory Pract 2008 Mar;13(1):71-87. [doi: 10.1007/s10459-006-9025-3] [Medline: 16944000]
21. Wilson-Sands C, Brahn P, Graves K. The effect of instructional method on cardiopulmonary resuscitation skill performance: a comparison between instructor-led basic life support and computer-based basic life support with voice-activated manikin. J Nurses Prof Dev 2015;31(5):E1-E7. [doi: 10.1097/NND.0000000000000203] [Medline: 26381346]
22. Johnson D, Flagg A, Dremsa T. Effects of using human patient simulator versus a CD-ROM on learning the management of patients exposed to chemical agents. US Army Med Dep J 2010:9-16. [Medline: 21181669]
23. Liaw SY, Wong L, Chan S, Ho J, Mordiffi S, Ang S, et al. Designing and evaluating an interactive multimedia web-based simulation for developing nurses' competencies in acute nursing care: randomized controlled trial. J Med Internet Res 2015 Jan 12;17(1):e5-10 [FREE Full text] [doi: 10.2196/jmir.3853] [Medline: 25583029]
24. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [Medline: 17695343]
25. Duarte H. Perception of learning, satisfaction and self-efficacy of nursing students on high fidelity simulation. In: Leiria School of Health. Leiria: Polytechnic Institute of Leiria; 2015.
26. Nunes R, Schwarzer R, Jerusalém M. Perceived General Self-Efficacy Scale. Berlin University: Health Psychology; 1999. URL: http://userpage.fu-berlin.de/~health/selfscal.htm
27. Schwarzer R, Jerusalém M. Generalized self-eEfficacy scale. In: Measures in Health Psychology: A User's Portfolio. Casual and Control Beliefs. Windsor, England: NFER-NELSON; 1995:A.
28. Scholz U, Gutiérrez Doña B, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Eur J Psychol Assess 2002 Sep;18(3):242-251. [doi: 10.1027//1015-5759.18.3.242]
29. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
30. Padilha JM, Machado P, Ribeiro A, Ramos J. Clinical virtual simulation in nursing education. Clin Simul Nurs 2018 Feb;15:13-18 [FREE Full text] [doi: 10.1016/j.ecns.2017.09.005]
31. Tschannen D, Aebersold M, McLaughlin E, Bowen J, Fairchild J. Use of virtual simulations for improving knowledge transfer among baccalaureate nursing students. J Nurs Educ Pract 2012 May 13;2(3):15-24. [doi: 10.5430/jnep.v2n3p15]

XSL•FO
RenderX

32. Sperl-Hillen J, O'Connor PJ, Ekstrom HL, Rush WA, Asche SE, Fernandes OD, et al. Educating resident physicians using virtual case-based simulation improves diabetes management: a randomized controlled trial. Acad Med 2014 Dec;89(12):1664-1673 [FREE Full text] [doi: 10.1097/ACM.0000000000000406] [Medline: 25006707]
33. Tiffany JM, Hoglund B. Using virtual simulation to teach inclusivity: a case study. Clin Simul Nurs 2016 Apr;12(4):115-122. [doi: 10.1016/j.ecns.2015.11.003]
34. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliff, NJ: Prentice-Hall; 1986:A.
35. Berman N, Fall LH, Smith S, Levine D, Maloney CG, Potts M, et al. Integration Strategies for Using Virtual Patients in Clinical Clerkships. Acad Med 2009;84(7):942-949. [doi: 10.1097/ACM.0b013e3181a8c668] [Medline: 19550193]
36. Huwendiek S, Duncker C, Reichert F, De Leng BA, Dolmans D, van der Vleuten CP, et al. Learner preferences regarding integrating, sequencing and aligning virtual patients with other activities in the undergraduate medical curriculum: A focus group study. Med Teach 2013 Nov;35(11):920-929. [doi: 10.3109/0142159X.2013.826790] [Medline: 24003850]

Abbreviations

ICT: information and communications technologies
MANOVA: multivariate analysis of variance

Edited by G Eysenbach; submitted 10.07.18; peer-reviewed by P Sousa, M Triola; comments to author 14.10.18; revised version received 21.12.18; accepted 03.01.19; published 18.03.19.

Please cite as:
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
URL: http://www.jmir.org/2019/3/e11529/
doi:10.2196/11529
PMID:30882355

©José Miguel Padilha, Paulo Puga Machado, Ana Ribeiro, José Ramos, Patrício Costa. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

WeChat as a Platform for Problem-Based Learning in a Dental Practical Clerkship: Feasibility Study

Wei Zhang[1,2], DDS, MD, PhD; Zheng-Rong Li[1], DDS; Zhi Li[1,3], DDS, MD, PhD

[1]State Key Laboratory Breeding Base of Basic Science of Stomatology (Hubei-MOST) and Key Lab for Oral Biomedical Engineering of the Ministry of Education, School and Hospital of Stomatology, Wuhan University, Wuhan, China

[2]Department of Endodontics, School and Hospital of Stomatology, Wuhan University, Wuhan, China

[3]Department of Oral and Maxillofacial Surgery, School and Hospital of Stomatology, Wuhan University, Wuhan, China

Corresponding Author:
Zhi Li, DDS, MD, PhD
State Key Laboratory Breeding Base of Basic Science of Stomatology (Hubei-MOST) and Key Lab for Oral Biomedical Engineering of the Ministry of Education
School and Hospital of Stomatology
Wuhan University
237 Luoyu Road
Wuhan,
China
Phone: 86 02787686217
Fax: 86 02787873260
Email: zhili@whu.edu.cn

Abstract

Background: Problem-Based-Learning (PBL) has been widely accepted in student-centered medical education. Since WeChat is the most popular communication app in China, we have chosen to use WeChat as new platform for online PBL in order to reduce the limitations of traditional PBL in dental practical clerkships.

Objective: This study aims to demonstrate the feasibility and acceptability of online PBL using WeChat (WeChat-PBL) in a dental practical clerkship.

Methods: A total of 72 students in a dental practical clerkship and 10 tutors participated in this study from June to August 2017. We created 10 WeChat groups to provide a communication platform for the PBL teaching, in which the students selected the PBL cases themselves from their practical clerkship. After each individual PBL case, group members were required to complete an evaluation on the PBL process itself. A final questionnaire survey was completed by the participants to summarize the long-term evaluation of the whole WeChat-PBL experience after the 3-month clerkship. Data from the PBL cases, WeChat messages, periodic evaluations, and long-term evaluations were collected for analysis.

Results: There were 45 cases presented in the WeChat-PBL within the 3-month clerkship. All students had positive reactions to the communication within the PBL groups. The results of the periodic evaluation showed that the students and tutors were quite satisfied with the process of WeChat-PBL and appreciated the group members' contributions and performance. The final questionnaire results indicated that the WeChat-PBL had achieved positive effects.

Conclusions: The results of this study indicate the feasibility and acceptability of the app, WeChat, for problem-based learning in a dental practical clerkship.

(J Med Internet Res 2019;21(3):e12127) doi:10.2196/12127

KEYWORDS

PBL; app; WeChat; clerkship; dental medicine

Introduction

Dental medicine is a practical discipline, and the clerkship is a key component of dental medicine education. The quality of the clerkship directly affects the students' future career. In China, the dental practical clerkship is still a teacher-centered process rather than a student-centered one, leading to students who struggle to think independently and solve clinical problems by themselves. The traditional pedagogy of teacher-centered,

XSL•FO
RenderX

class-oriented didactic lectures and an examination-driven curriculum put the students in a passive state of "reception" [1], which results in a lack of initiative among students in clerkships. Problem-based learning (PBL) is a widely accepted student-centered educational method focused on the discussion and learning that emerge from focusing on a clinically based problem [2,3]. As an innovative approach in medical education, the role of PBL has been well documented since the pioneering use of PBL in medical education at McMaster University (Ontario, Canada) in the late 1960s [4]. There are significant contributions from PBL in medical education, and we wish to test whether PBL can improve the quality of dental practical clerkships.

However, traditional PBL has its own limitations in a dental practical clerkship. Traditional PBL usually involves a group of 6-8 students with a tutor in a face-to-face PBL session, where the participants get together in a classroom to discuss and share opinions. The face-to-face session is essential in traditional PBL. However, due to the geographic dispersion of students, it is difficult to perform traditional PBL in a dental practical clerkship. Generally, students are assigned to different clinical departments for their clerkship rotation, which makes it hard to schedule regular meetings of the students in order to launch a PBL session. Moreover, the students and tutors lack the free time needed for PBL during busy clerkships. Therefore, modifications to traditional PBL are needed in order to get around the physical and temporal restrictions in dental clerkships.

To address these problems, we are designing a novel method to transfer traditional PBL to an online PBL setting in a virtual environment. The new mode of PBL would be built on modern digital technology. Therefore, the physical and time restrictions of traditional PBL education are eliminated, stimulating students' autonomous learning. Mobile phones, specifically smartphones, are one of the fastest growing sectors in the Internet technology industry, and their impact in medicine has already been significant [5,6]. Many educators have used different mobile phone apps with positive results: in particular, students' self-efficacy, self-confidence, and self-management have shown improvement [7,8]. Because mobile phone apps are easy and convenient to operate, they are becoming more and more popular in education. Social media platforms such as Facebook and Twitter have been introduced into medical education in western countries with great success [9-11].

After reflecting on the situation in China, we decided on the multipurpose messaging, social media, and mobile payment app, WeChat, as a new platform to deliver PBL to students in a dental practical clerkship. WeChat, released in 2011 by Tencent, is one of the largest standalone mobile phone messaging communication apps in China, with over 938 million global active users by 2017. In addition, WeChat has already been used as a mobile and interactive communication tool in medical education [12,13]. WeChat provides text, hold-to-talk voice and broadcast (one-to-many) messaging, instant video conferencing, and photograph or video sharing. With those functions included in WeChat, it is a suitable tool for online PBL, eliminating the physical limitations of traditional PBL.

In this study, we explored WeChat-based PBL for students in a dental practical clerkship and aimed to demonstrate its feasibility and acceptability.

Methods

Participants

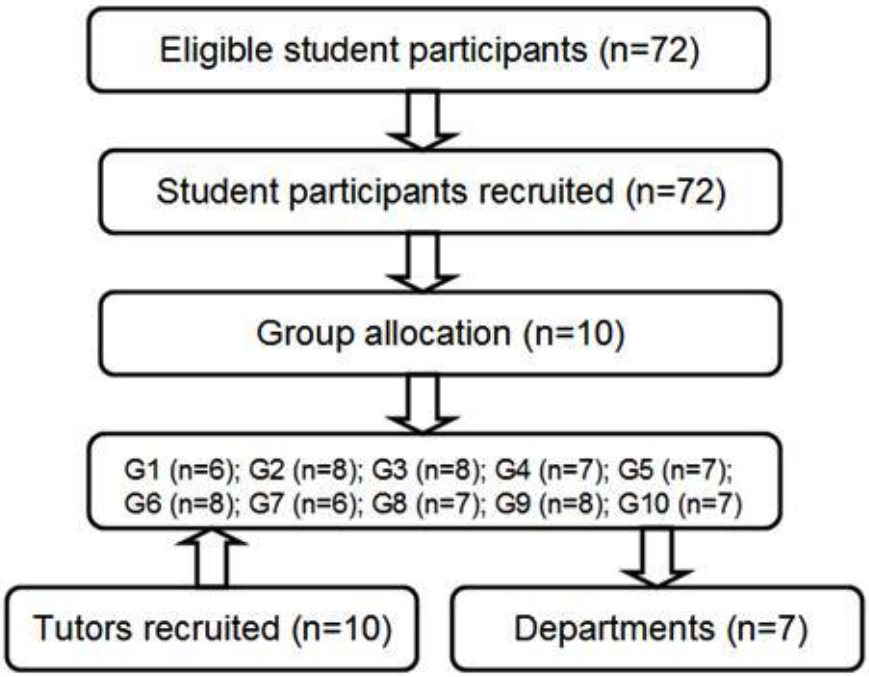
In this study, use of the WeChat-PBL took place from June to August 2017 at the Hospital of Stomatology, Wuhan University, in China. Clerkship students practicing dental medicine in different departments (including the Departments of Oral and Maxillofacial Surgery, Endodontics, Prosthodontics, Periodontology, Orthodontics, Pediatric Dentistry, and Dental Implantology) participated in this study. During their clerkships, students are required to complete clinical rotations in different departments and perform treatments under the supervision of senior doctors. The time spent per rotation in each department is 6 weeks.

The study was performed with dental students participating in a 3-month clerkship. Students who had traditional PBL experience in the preclinical curriculum were selected. Clinical doctors were assigned as tutors in the WeChat-PBL groups. They were required to have at least one year or more of experience in traditional PBL teaching. In total, 72 students (34 males, 38 females; mean age 23.6 years [SD 2]) and 10 tutors were included in this study (Figure 1). Informed consent was given to participants and signed by them prior to beginning the study.

All participants had their own mobile phones and were required to install WeChat. They were trained to use the practical aspects of WeChat in the PBL context. A total of 10 WeChat groups was created to provide the communication platform for the PBL teaching. The 72 students and 10 tutors from different departments were randomly assigned to each group. Six to eight students and the assigned tutor were asked to join the same WeChat group. Each group consisted of students working in different departments.

XSL•FO
RenderX

Figure 1. Flowchart showing the sampling process and conceptual framework.



WeChat Problem-Based Learning Mode

WeChat is an app available on Android, iPhone, BlackBerry, Windows, and Symbian platforms and is supported by Wi-Fi, 3G, and 4G data networks. Students can communicate with each other on WeChat at any time anywhere. This new mode of PBL was designed based on the WeChat app. The WeChat group worked like the traditional PBL group, comprising 6-8 students and a tutor (Figure 2). The PBL group members performed the WeChat-PBL by sending and receiving texts, images, voice messages, videos, and documents (eg, MS PowerPoint, MS Word, and Adobe PDF) (Figure 3). In addition, the PBL group members had real-time discussions using the video conferencing function in WeChat (Figure 4).

The dental students themselves selected the PBL cases. There were no definitive criteria to assess the PBL cases, and the number of the cases was not prespecified. Students selected the PBL cases, based on the clerkship intent and initially guided by the tutor. When a student saw a potentially suitable case in their clerkship, they would upload the clinical data of the case (including basic information, chief complaint, symptoms, signs, and laboratory examinations) to the WeChat-PBL group (Figure 5). The group members would discuss and decide whether the case was chosen for the PBL. The tutor then approved the case put forward by the students, aiming to ensure the quality of case.

Due to WeChat's powerful functions, the group members could upload texts, images, voice messages, and videos easily and instantly (Figure 6). The group members asked questions about the selected case in the WeChat-PBL group, then discussed the various issues and summarized the questions within several days (Figure 7). The questions were then classified and assigned to the group members. Group members would then find the information, search for possible answers, and post their opinions by text, image, voice, video, or other forms (ie, documents) (Figure 8). Furthermore, the PBL discussion aligned with the clinical treatment of the selected case (Figure 9).

The tutor joined in the WeChat-PBL discussion and occasionally gave guidance to students. The tutor's aim was to observe the performance and the number of active students in this new mode of PBL. The tutor was not supposed to answer the questions summarized by the students. The main duties of the tutor were as follows: (1) guide the students for deeper and wider thinking, and (2) ensure that every group member is participating in the discussion.

XSL•FO
RenderX

Figure 2. One of the problem-based-learning groups (7 students and a tutor).

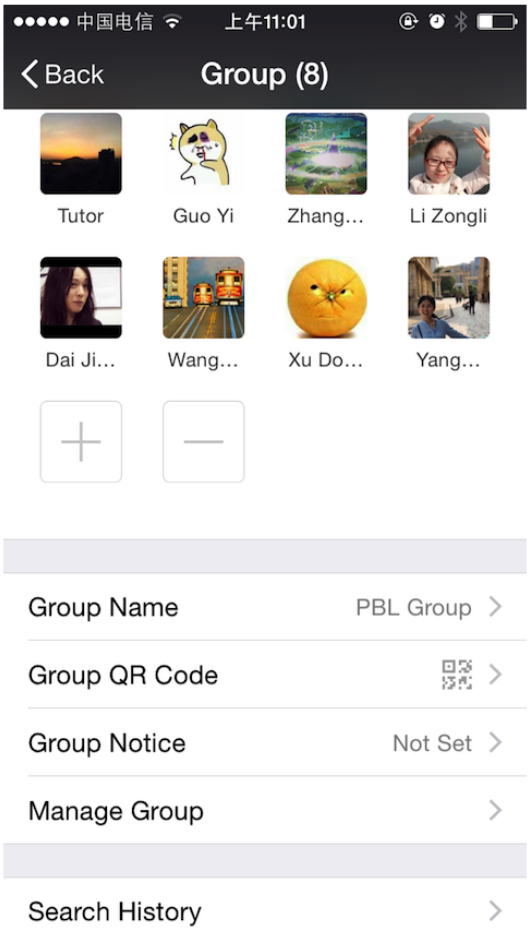
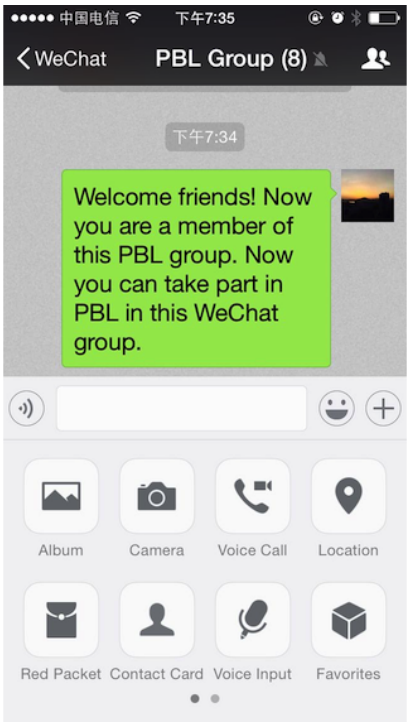


Figure 3. User interface for chatting, which was used to send and receive texts, images, voice messages, videos, and documents.

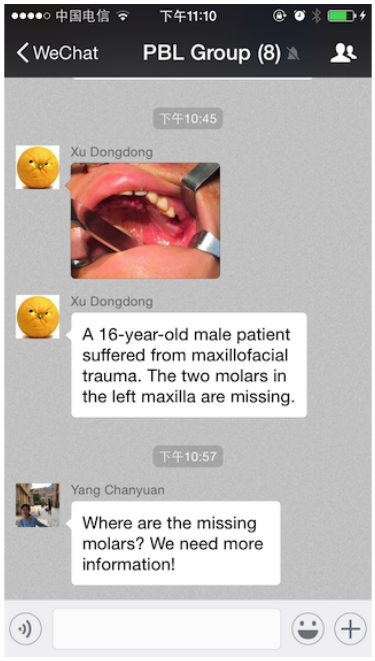


XSL•FO
RenderX

Figure 4. Video conferencing in WeChat, which was used for real-time problem-based-learning discussion.



Figure 5. A real case was selected as the problem-based-learning (PBL) case by students, and the clinical data were uploaded to the WeChat-PBL group.



XSL•FO
RenderX

Figure 6. In the WeChat-PBL (problem-based-learning) group, the group members carry out the discussion by sending texts, images, voice messages, and videos.

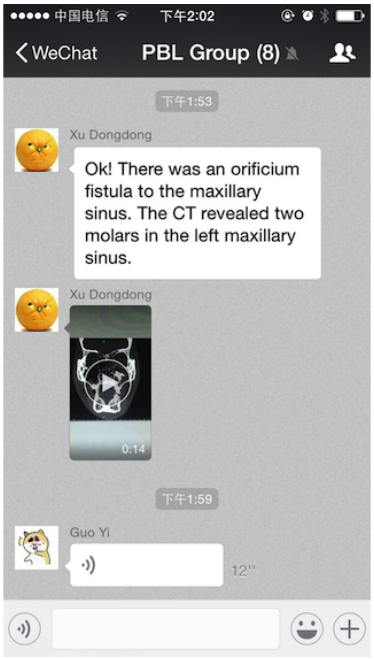
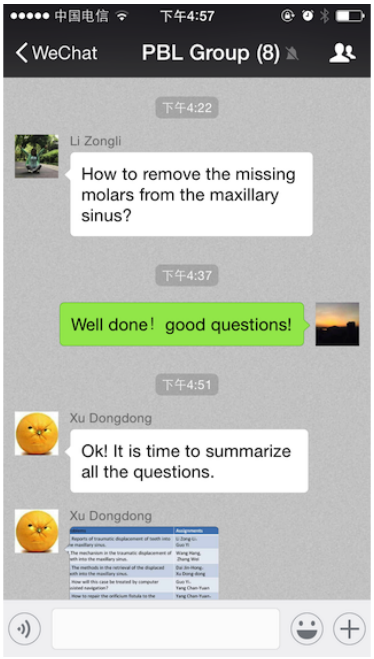


Figure 7. In the problem-based-learning discussion, questions were summarized and assigned to group members.

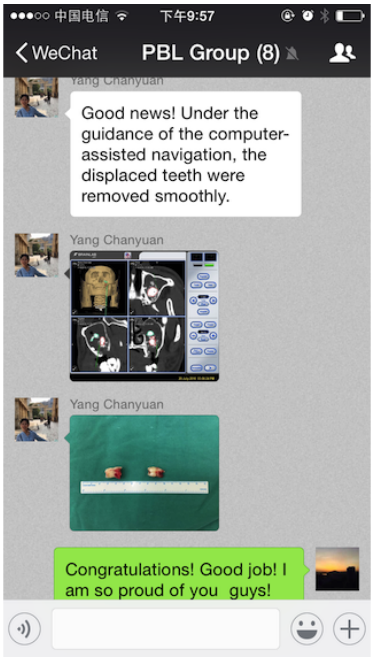


XSL•FO
RenderX

Figure 8. A group member expressed opinions on a question after literature review and learning, and then shared the relevant documents with group members.



Figure 9. Problem-based-learning discussions align with the clinical treatment of the selected case.



Evaluation

The key to evaluating the WeChat-PBL was to infer whether the students were effectively gaining knowledge and improving skills, as well as developing independent clinical thinking and team skills. The evaluations included two sections: (1) periodic evaluation of the PBL process after each individual PBL case, and (2) long-term evaluation of the whole WeChat-PBL experience after the 3-month clerkship.

After each individual PBL case, the students and tutor were required to finish the periodic evaluation, including self-evaluation and peer evaluation, which comprised discussions in the WeChat group. The tutor also evaluated each student's behavior and performance in the WeChat-PBL.

The long-term evaluation of the whole WeChat-PBL experience after the 3-month clerkship was summarized by a questionnaire survey to each student (Table 1). The questionnaire was designed and modified based on previous studies [14-16] and the objectives of this study. The questionnaire was validated by a pilot study with a smaller sample.

XSL•FO
RenderX

Table 1. Questionnaire results after 3 months of clerkship.

Questions	Yes, n (%)	No, n (%)
1. Did you enjoy discussing the PBL cases from the clerkship?	70 (97.22)	2 (2.78)
2. Did you enjoy communicating by WeChat in the online PBL?	72 (100)	0 (0.00)
3. Did you contribute to the group discussion?	60 (83.33)	12 (16.67)
4. Did you complete the tasks assigned to you on time?	58 (80.56)	14 (19.44)
5. Did you apply prior knowledge to solve problems?	62 (86.11)	10 (13.89)
6. Were you able to provide new information?	65 (90.28)	7 (9.72)
7. Did you feel free to propose questions?	70 (97.22)	2 (2.78)
8. Did you pay attention to the relevance of the information?	70 (97.22)	2 (2.78)
9. Did you assess the information you acquired critically?	59 (81.94)	13 (18.06)
10. Did you actively participate in interactive discussion?	60 (83.33)	12 (16.67)
11. Were you able to communicate ideas effectively with your group members?	61 (84.72)	11 (15.28)
12. Did your tutor consistently show enthusiasm with the PBL?	70 (97.22)	2 (2.78)
13. Did your group members and tutor offer feedback often to improve your learning?	69 (95.83)	3 (4.17)
14. Do you now have a better understanding of PBL?	72 (100)	0 (0.00)
15. Do you think WeChat is an effective app for online PBL in a dental clerkship?	72 (100)	0 (0.00)
16. Do you agree that online PBL based on WeChat can improve students' ability in the clerkship?	71 (98.61)	1 (1.39)

Analysis

The data from PBL cases and WeChat messages were collected for quantitative analysis in MS Excel 2013. The content of the periodic evaluations within the WeChat-PBL groups was analyzed qualitatively. The questionnaires for long-term evaluation were collected and analyzed, gathering participants' opinions and comparing them, in order to understand whether the students benefited from this WeChat-PBL mode.

Results

Problem-Based Learning Cases

In this study, there were 45 cases presented in the WeChat-PBL within the 3-month clerkship, and each group finished four or five PBL cases. These 45 cases were all real clinical cases, including diseases from different departments (Oral & Maxillofacial Surgery, n=12; Endodontics, n=10; Periodontics Department, n=7; Prosthodontics, n=5; Orthodontics, n=5; Pediatric Dentistry, n=4; Dental Implantology, n=2). The multidisciplinary team approach was used in the discussion of 30 cases (66.67%). It took 15-22 days (mean 20.12 days) to finish one PBL case.

WeChat Messages

All students had positive reactions to the communication within the groups. A total of 2063 text messages, 2325 voice messages, 364 images, 92 videos, 278 documents, and 129 webpage links were sent in the 45 PBL cases. In addition, there were 54 video conferences. Of the 2063 text messages, 1823 came from students and 240 from tutors. Of the 2325 voice messages, 2068 were from students and 257 from tutors. All of the images, videos, documents, and webpage links were sent by students.

Periodic Evaluation

The periodic evaluation showed that students and tutors were quite satisfied with the process of WeChat-PBL and appreciated the group members' performance. The students held positive attitudes toward the WeChat-PBL teaching, noting that WeChat-PBL encouraged the activation of prior knowledge and provided opportunities for elaboration on that knowledge. The tutors reported that the students had obvious increases in both general knowledge and the skills for analyzing problems.

Long-Term Evaluation

The response rate to the questionnaire was 100% (72/72). The results of the questionnaire indicated that the WeChat-PBL achieved positive effects (Table 1). All the students commented positively about communicating on WeChat in the online PBL. Most students enjoyed discussing the real PBL cases from the clerkship (70/72, 97.2%), felt free to propose questions (70/72, 97.2%), and paid attention to the relevance of the information (70/72, 97.2%). As well, most students reported that they were able to provide new information (65/72, 90.28%) and the group members and tutor often offered feedback to improve their learning (69/72, 95.8%).

In addition, most of the students evaluated themselves highly on other aspects in the PBL, including contributing to the group discussion (60/72, 83.3%), completing the assigned tasks on time (58/72, 80.5%), applying prior knowledge to solve problems (62/72, 86.1%), assessing the acquired data critically (59/72, 81.9%), participating in the interactive discussion actively (60/72, 83.3%), and communicating ideas effectively with group members (61/72, 84.7%).

The majority of the students (70/72, 97.2%) felt the tutors showed persistent enthusiasm in the WeChat-PBL, and the tutors gave a positive evaluation of all the students' behavior in the

XSL•FO
RenderX

PBL. All students reported having a better understanding of PBL and indicated that WeChat was an ideal app for online PBL in the dental practical clerkship. Similarly, almost all students agreed that the WeChat-PBL could improve students' skills in the clerkship.

Discussion

Principal Results

In this study, we designed a new PBL mode based on WeChat. This mode of online PBL successfully eliminated the physical and temporal limitations of traditional PBL in dental clerkships. It ensured the time needed for and the quality of PBL, broadened the manner that students gain knowledge, and promoted efficiency in solving problems in the dental practical clerkship. As a modern pedagogical philosophy, the importance of PBL is increasingly being recognized in student learning and innovation in medical education [17]. Many educators have tried to improve traditional PBL by modifying the instruction, hence, other PBL modes such as tutorless PBL, 3C3R Modified PBL, and Hybrid-PBL have emerged in PBL teaching [18-20]. However, compared with traditional PBL, the WeChat-PBL has several advantages that take PBL to a higher level.

The most important advantage of our WeChat-based PBL is its timesaving and convenience. Traditional PBL is time-consuming [21], since a regular schedule is needed to gather the students together. However, WeChat-PBL allows asynchronous discussions and supports discussions "anytime, anywhere." Furthermore, group members can upload images, videos, radiological examinations, or laboratory results of the case via mobile phone in a timely manner, so that other members can instantly access the case data.

Second, the new PBL mode offers an excellent platform for sharing educational resources and the newest information. Group members can use their mobile phones, searching valuable information on the Internet. Many universities offer a range of online e-Health courses and other online curricula [22,23]. Group members can send webpage links to these and other resources to the group chat directly, thus, improve the efficiency and quality of the discussions in the PBL.

Third, this new mode improves students' motivation to learn and strengthens their practical ability. The questionnaire results indicated that the students became more and more curious and active in WeChat-PBL, aiming to grasp the practical use of theoretical knowledge instead of just knowing about the theory. Group members became better at listening and communicating in the new PBL mode, which helped them achieve the ultimate intent of the clerkship.

Social media has been gaining steady support for its presence in medical education. The findings of this study illustrate a method through which online PBL teaching can be facilitated by the use of social networking sites. Compared with the use of similar social media platforms such as WhatsApp [24] and Twitter [25] in PBL, WeChat has some special functions that are particularly well adapted for online PBL discussions including video conferencing and sharing of documents (eg, MS PowerPoint, MS Word, and Adobe PDF). Another noteworthy feature of WeChat-PBL in our study is that the PBL case was directly selected from students' clinical work, and the PBL discussions aligned with the clinical treatment of the selected case. These real cases helped students recognize the authenticity of the study materials and become more motivated to explore and solve the problems [26]. Moreover, the students were able verify the solutions in the practice, truly integrating theory with practice. In this study, each PBL group consisted of students belonging to more than one department, and the teachers were randomly assigned across departments. These features contribute to the students' dental practice in different clinical departments in their clerkship rotation, integrating the knowledge of different disciplines in dentistry. Further, the tutor approved the case selected by the students instead of preparing a case, allowing the tutor more time and energy to focus on the guidance.

Limitations

In order to use this new PBL mode based on WeChat, there are several basic requirements: a suitable mobile phone, the WeChat app, and Internet or Wi-Fi access. For students in China, owning a mobile phone is not a financial burden. In this study, all students had their own phones. Furthermore, WeChat is a free app that anyone can download and install on a supported mobile phone. These days, all colleges and universities in China provide a Web-based information platform for their campuses. Free wireless networks are ubiquitous in student dormitories and in hospitals. However, there exist several shortcomings in this new PBL mode.

First, the dental students have more freedom using WeChat-PBL than in traditional PBL. This helps motivate students to spend more time in the PBL. When there is increased control in traditional PBL, students' motivation to assume responsibility for learning decreases [27]. However, too much freedom can lead to laziness, procrastination, and so on. Without the restrictions of the classroom setting, some undisciplined students may send off-topic pictures or messages to amuse others in the PBL group chat. Quality pedagogy needs to balance freedom and restraint [28]. Therefore, the PBL group members should set and enforce guidelines and rules in the group chat. If necessary, the tutor could warn the undisciplined students privately. Fortunately, this phenomenon never occurred in this study. Students were motivated and stayed on topic.

Second, this online PBL mode lacks ways to manage discussion topics effectively. The personality types of students and their group dynamics are linked to their PBL performance [29]. During an online discussion, such as the WeChat-PBL, the topics can be dominated by a few active users, since there is not the restriction of traditional classroom setting. Students can discuss what they are interested in and can go off topic unintentionally, potentially affecting the enthusiasm and activity of the other students. Keeping students engaged in a virtual environment requires a sustained instructor presence. In order to overcome this shortcoming, the tutor should always be active in the asynchronous discussion group, encouraging students' discussion and prompting them to think more deeply. In addition, students admitted to the WeChat-PBL should receive training on this new mode as part of the PBL orientation

XSL•FO
RenderX

program. In our study, we found that all students were able to focus on the discussion topics after receiving training.

As far as the study method is concerned, there were some limitations to our design. The study lacked a control group, and the evaluations focused only on perception/satisfaction with the PBL, rather than assessing actual changes in knowledge or skills. In addition, the tool used to evaluate the experience was not validated, and thus the satisfaction reported may be due to the novelty effect. Therefore, longer studies are needed to assess intermediate and long-term effects. Finally, the study lacked a rigorous sampling process, and all eligible students were recruited into our study.

Conclusions

The results indicate the feasibility and acceptability of WeChat in PBL teaching for students in a dental practical clerkship. This new PBL mode not only has the advantages of timesaving and convenience but also offers a suitable platform for sharing educational resources and the newest information. The WeChat-PBL improves students' learning motivation and strengthens their practical ability. It also contributes to students' dental practice in different clinical departments in their clerkship rotation, integrating the knowledge of different disciplines in dentistry. Although there are certain limitations in WeChat-PBL, we were able to find the solutions during the study process. We conclude from this study that our WeChat-PBL is an effective method for PBL teaching in a dental practical clerkship.

Acknowledgments

This work was supported by the Teaching Research Project in Colleges and Universities of Hubei Province, China (No. 2010006), the Fundamental Research Fund for the Central Universities of China (2042017kf0206), and the Hubei Province Health and Family Planning Scientific Research Project (WJ2017M047).

Authors' Contributions

WZ and ZL contributed to the study conception and design, the analysis and interpretation of data, the drafting and revising of the paper. Z-RL contributed to the study design, the analysis and interpretation of data, and the drafting of the paper.

Conflicts of Interest

None declared.

References

1. Tu M, Yu C, Wu L, Li T, Kwan C. Dental and medical students' perspectives on early exposure to PBL in Taiwan. J Dent Educ 2012 Jun;76(6):746-751 [FREE Full text] [Medline: 22659703]
2. Rui Z, Rong-Zheng Y, Hong-Yu Q, Jing Z, Xue-Hong W, Chuan Z. Preliminary investigation into application of problem-based learning in the practical teaching of diagnostics. Adv Med Educ Pract 2015;6:223-229 [FREE Full text] [doi: 10.2147/AMEP.S78893] [Medline: 25848334]
3. Amin Z, Hoon Eng K, Gwee M, Dow Rhoon K, Chay Hoon T. Medical education in Southeast Asia: emerging issues, challenges and opportunities. Med Educ 2005 Aug;39(8):829-832. [doi: 10.1111/j.1365-2929.2005.02229.x] [Medline: 16048625]
4. Neufeld VR, Barrows HS. The "McMaster Philosophy": an approach to medical education. J Med Educ 1974 Nov;49(11):1040-1050. [Medline: 4444006]
5. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. J Med Internet Res 2012 Sep 27;14(5):e128 [FREE Full text] [doi: 10.2196/jmir.1994] [Medline: 23017375]
6. Rat C, Hild S, Rault Sérandour J, Gaultier A, Quereux G, Dreno B, et al. Use of Smartphones for Early Detection of Melanoma: Systematic Review. J Med Internet Res 2018 Apr 13;20(4):e135 [FREE Full text] [doi: 10.2196/jmir.9392] [Medline: 29653918]
7. Hong J, Hwang M, Tai K, Chen Y. Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety. Computers & Education 2014 Mar;72:313-322 [FREE Full text] [doi: 10.1016/j.compedu.2013.11.011]
8. Nadal R, Lellis-Santos C. Using smartphones to learn exercise physiology. Med Educ 2015 Nov;49(11):1162-1163. [doi: 10.1111/medu.12877] [Medline: 26494096]
9. Diug B, Kendal E, Ilic D. Evaluating the use of twitter as a tool to increase engagement in medical education. Educ Health (Abingdon) 2016;29(3):223-230 [FREE Full text] [doi: 10.4103/1357-6283.204216] [Medline: 28406107]
10. Pickering JD, Bickerdike SR. Medical student use of Facebook to support preparation for anatomy assessments. Anat Sci Educ 2017 Jun;10(3):205-214. [doi: 10.1002/ase.1663] [Medline: 27806192]
11. Sterling M, Leung P, Wright D, Bishop TF. The Use of Social Media in Graduate Medical Education. Academic Medicine 2017;92(7):1043-1056. [doi: 10.1097/ACM.0000000000001617]

XSL•FO
RenderX

12. Wang J, Gao F, Li J, Zhang J, Li S, Xu G, et al. The usability of WeChat as a mobile and interactive medium in student-centered medical teaching. Biochem Mol Biol Educ 2017 Sep;45(5):421-425. [doi: 10.1002/bmb.21065] [Medline: 28581156]
13. Zeng F, Deng G, Wang Z, Liu L. WeChat: a new clinical teaching tool for problem-based learning. Int J Med Educ 2016 Apr 25;7:119-121 [FREE Full text] [doi: 10.5116/ijme.5708.e5c4] [Medline: 27111920]
14. Bosse HM, Huwendiek S, Skelin S, Kirschfink M, Nikendei C. Interactive film scenes for tutor training in problem-based learning (PBL): dealing with difficult situations. BMC Med Educ 2010 Jul 06;10(1):52 [FREE Full text] [doi: 10.1186/1472-6920-10-52] [Medline: 20604927]
15. Shankar PR, Nandy A. Student feedback on problem-based learning processes. Australas Med J 2014;7(12):522-529 [FREE Full text] [doi: 10.4066/AMJ.2014.2208] [Medline: 25646128]
16. Win N, Nadarajah V, Win D. The implementation of problem-based learning in collaborative groups in a chiropractic program in Malaysia. J Educ Eval Health Prof 2015;12:17 [FREE Full text] [doi: 10.3352/jeehp.2015.12.17] [Medline: 25961676]
17. Jin J, Bridges SM. Educational technologies in problem-based learning in health sciences education: a systematic review. J Med Internet Res 2014 Dec 10;16(12):e251 [FREE Full text] [doi: 10.2196/jmir.3240] [Medline: 25498126]
18. Kaliyadan F, Amri M, Dhufiri M, Amin TT, Khan MA. Effectiveness of a modified tutorless problem-based learning method in dermatology - a pilot study. J Eur Acad Dermatol Venereol 2012 Jan;26(1):111-113. [doi: 10.1111/j.1468-3083.2011.04016.x] [Medline: 21366714]
19. Xue H, Qian J, Wang L, Yuan X, Chen Y, Wu W, et al. 3C3R modified PBL pediatric teaching of Chinese medical students. PLoS One 2013;8(5):e63412 [FREE Full text] [doi: 10.1371/journal.pone.0063412] [Medline: 23667610]
20. Yan Q, Ma L, Zhu L, Zhang W. Learning effectiveness and satisfaction of international medical students: Introducing a Hybrid-PBL curriculum in biochemistry. Biochem Mol Biol Educ 2017 Jul 08;45(4):336-342. [doi: 10.1002/bmb.21046] [Medline: 28696053]
21. Seneviratne R, Samarasekera D, Karunathilake I, Ponnamperuma G. Students' perception of problem-based learning in the medical curriculum of the Faculty of Medicine, University of Colombo. Ann Acad Med Singapore 2001 Jul;30(4):379-381. [Medline: 11503544]
22. Edirippulige S, Smith AC, Armfield NR, Bensink M, Wootton R. Student perceptions of a hands-on practicum to supplement an online eHealth course. J Med Internet Res 2012 Dec 18;14(6):e182 [FREE Full text] [doi: 10.2196/jmir.2029] [Medline: 23246840]
23. Zhan X, Zhang Z, Sun F, Liu Q, Peng W, Zhang H, et al. Effects of Improving Primary Health Care Workers' Knowledge About Public Health Services in Rural China: A Comparative Study of Blended Learning and Pure E-Learning. J Med Internet Res 2017 Dec 01;19(5):e116 [FREE Full text] [doi: 10.2196/jmir.6453] [Medline: 28461286]
24. Raiman L, Antbring R, Mahmood A. WhatsApp messenger as a tool to supplement medical education for medical students on clinical attachment. BMC Med Educ 2017 Jan 06;17(1):7 [FREE Full text] [doi: 10.1186/s12909-017-0855-x] [Medline: 28061777]
25. Lasker R, Vicneswararajah N. Using Twitter to teach problem-based learning. Med Educ 2015 May;49(5):531. [doi: 10.1111/medu.12717] [Medline: 25924150]
26. Chye SY, Liau AK, Liu WC. Student Teachers' Motivation and Perceptions of E-Portfolio in the Context of Problem-Based Learning. Asia-Pacific Edu Res 2012 Dec 18;22(4):367-375 [FREE Full text] [doi: 10.1007/s40299-012-0022-4]
27. McAllister A, Aanstoot J, Hammarström IL, Samuelsson C, Johannesson E, Sandström K, et al. Learning in the tutorial group: a balance between individual freedom and institutional control. Clin Linguist Phon 2014;28(1-2):47-59. [doi: 10.3109/02699206.2013.809148] [Medline: 23848371]
28. Schaffar B. Changing the Definition of Education. On Kant's Educational Paradox Between Freedom and Restraint. Stud Philos Educ 2013 Feb 5;33(1):5-21. [doi: 10.1007/s11217-013-9357-4]
29. Ihm J, An S, Seo D. Do Dental Students' Personality Types and Group Dynamics Affect Their Performance in Problem-Based Learning? J Dent Educ 2017 Jun;81(6):744-751 [FREE Full text] [doi: 10.21815/JDE.017.015] [Medline: 28572421]

Abbreviations

PBL: problem-based learning

XSL•FO
RenderX

Edited by G Eysenbach; submitted 06.09.18; peer-reviewed by N Karim, S Amantini, M El Tantawi; comments to author 05.11.18; revised version received 27.12.18; accepted 29.12.18; published 19.03.19.

Please cite as:
Zhang W, Li ZR, Li Z
WeChat as a Platform for Problem-Based Learning in a Dental Practical Clerkship: Feasibility Study
J Med Internet Res 2019;21(3):e12127
URL: http://www.jmir.org/2019/3/e12127/
doi:10.2196/12127
PMID:30888328

©Wei Zhang, Zheng-Rong Li, Zhi Li. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Paid and Free Facebook Methods to Recruit Australian Parents to an Online Survey: An Evaluation

Shannon K Bennetts[1,2], PhD; Stacey Hokke[1], PhD; Sharinne Crawford[1], PhD; Naomi J Hackworth[1,2,3], DPsych; Liana S Leach[4], PhD; Cattram Nguyen[2], PhD; Jan M Nicholson[1], PhD; Amanda R Cooklin[1], PhD

[1]Judith Lumley Centre, School of Nursing and Midwifery, La Trobe University, Bundoora, Australia

[2]Murdoch Children's Research Institute, Parkville, Australia

[3]Parenting Research Centre, Melbourne, Australia

[4]Australian National University, Canberra, Australia

Corresponding Author:
Shannon K Bennetts, PhD
Judith Lumley Centre
School of Nursing and Midwifery
La Trobe University
Level 3, George Singer Building
Bundoora, 3086
Australia
Phone: 61 03 9479 ext 8763
Email: s.bennetts@latrobe.edu.au

Abstract

Background: The prevalence of social media makes it a potential alternative to traditional offline methods of recruiting and engaging participants in health research. Despite burgeoning use and interest, few studies have rigorously evaluated its effectiveness and feasibility in terms of recruitment rates and costs, sample representativeness, and retention.

Objective: This study aimed to determine the feasibility of using Facebook to recruit employed Australian parents to an online survey about managing work and family demands, specifically to examine (1) recruitment rates and costs; (2) sample representativeness, compared with a population-based cohort of parents; and (3) retention, including demographic and health characteristics of parents who returned to complete a follow-up survey 6 weeks later.

Methods: Recruitment was conducted using 20 paid Facebook advertising campaigns, supplemented with free advertising approaches such as posts on relevant Facebook pages and requests for professional networks to circulate the survey link via Facebook. Recruitment rates and costs were evaluated using the Checklist for Reporting Results of Internet E-Surveys, including view rate, participation rate, completion rate, cost per consent, and cost per completer. Sample representativeness was evaluated by comparing demographic and outcome variables with a comparable sample from the Longitudinal Study of Australian Children including educational attainment, marital status, country of birth, neighborhood disadvantage, work-family conflict, and psychological distress. Retention was evaluated by comparing the number and demographic characteristics of participants at recruitment and at 6-week follow-up.

Results: Recruitment strategies together resulted in 6653 clicks on the survey link, from which 5378 parents consented to participate and 4665 (86.74%) completed the survey. Of those who completed the survey, 85.94% (4009/4665) agreed to be recontacted, with 57.79% (2317/4009) completing the follow-up survey (ie, 43.08% [2317/5378] of parents who consented to the initial survey). Paid Facebook advertising recruited nearly 75% of the sample at Aus $2.32 per completed survey (Aus $7969 spent, 3440 surveys completed). Compared with a population-based sample, participants at baseline were more likely to be university educated (P<.001), experience greater work-family conflict (P<.001) and psychological distress (P<.001), and were less likely to be born outside Australia (P<.001) or live in a disadvantaged neighborhood (P<.001).

Conclusions: Facebook provided a feasible, rapid method to recruit a large national sample of parents for health research. However, some sample biases were observed and should be considered when recruiting participants via Facebook. Retention of participants at 6- to 8-week follow-up was less than half the initial sample; this may reflect limited ongoing participant engagement for those recruited through social media, compared with face-to-face.

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e11206) doi:10.2196/11206

KEYWORDS

parents; research subject recruitment; retention; social media; cost effectiveness; sampling bias; fathers; mothers; survey

Introduction

Overview

Busyness and time pressure are linked to reduced research participation [1], particularly for parents balancing the demands of work and family simultaneously [2]. This presents barriers to participation in health research, which is often time-intensive or requires face-to-face engagement. Social media platforms such as Facebook have the potential to reach parents who, as a cohort, are already strongly engaged in social media for parenting information, advice, or peer support [3,4]. Employed parents typically report high levels of time pressure [5]; therefore, online methodologies may allow more frequent and flexible opportunities for parents to engage in research. Although the omnipresence of social media in modern society offers new opportunities for efficient and low-cost participant recruitment, rigorous evaluation of Facebook as a recruitment method is scant and many unanswered questions remain. How feasible is this method for health research? Can it replace traditional offline methods of participant recruitment? Are samples of parents recruited via Facebook comparable to those recruited via offline methods? What are the retention rates for a convenience sample of parents recruited through Facebook? We address these questions salient to many researchers in a variety of health- and family-related fields and report on the methodology and feasibility of using Facebook to recruit employed Australian parents to an online survey, with a brief follow-up.

Background

Facebook is a free social networking website whereby users create profiles, share content, and connect with other users. It remains the most popular social media platform globally, with over 2 billion active monthly users [6], including 15 million Australians, of whom 50% log in daily [4,7]. Facebook is the main social media platform used by parents [8], attracting more parents than nonparents on a daily basis [9]. In recent years, parents' use of social media has increased [10], providing new ways for parents to maintain social ties, connect with other parents, share experiences, seek social and emotional support, and obtain parenting information [9,11]. Mothers typically use Facebook more frequently than fathers, and usage tends to increase during the transition to parenthood [3,9].

Given such widespread use of social media (and the ubiquity of smartphones), it is unsurprising that researchers are beginning to harness Facebook as a mobile, flexible, and potentially low-cost research tool. Traditional offline methods such as mail-outs and telephone interviews are becoming less feasible and less effective, evidenced by rising postage costs, increased refusal rates for household surveys, and reduced use of landline phones [12-14]. Extant literature varies enormously with respect to Facebook recruitment rates and costs. For example, Leach et al [15] recruited 819 new mothers to a health and well-being survey using Facebook advertisements in just 4 days, at a cost of Aus $0.55 per participant. Conversely, Kapp et al [16] used Facebook advertisements to recruit women to a survey of mammogram use but failed to recruit any eligible respondents, despite offering monetary incentives. Such variability likely reflects multiple factors such as the salience of the research topic, Facebook use among the target population, how the advertisement is presented, as well as the provision of incentives. A recent systematic review highlighted mixed results from Facebook recruitment studies [17]. From 54 studies, an average of 736 participants (range: 0 to 11799) were recruited at a cost of US $1.36 to US $110 per completing participant (per study average=US $17.48, of the 21 studies that provided sufficient data). Despite this variability, Facebook has been particularly effective for the recruitment of vulnerable and traditionally hard-to-reach populations, including young adults [18], HIV-positive participants [19], lesbians, bisexuals, gay, transgender, intersex, queer populations [20], new mothers [15], and low-income populations [21].

Although the use of Facebook for recruitment has grown dramatically in recent years, sample representativeness remains underreported. Thornton et al [17] found that only 16 (14.5%) studies examined sample representativeness, of which 86% concluded that their Facebook sample was similar to samples recruited using traditional methods. Leach et al [15] found that, compared with a representative population sample of mothers, mothers recruited via Facebook were younger and more likely to be in a de facto relationship, highly educated, first-time mothers, and speak only English at home. Importantly, these mothers also had poorer self-reported physical and mental health, suggesting potential bias in who self-selects into research participation about specific topics. Thornton et al [17] conclude that although the majority of Facebook samples have similar demographic characteristics to samples recruited via other methods, they are often not representative when compared with national population data. Evidence regarding the retention of samples recruited via Facebook is similarly limited. A 1-month follow-up of young adult veteran drinkers reported nearly 80% retention, with highly educated participants less likely to drop out [22].

This Study

Rigorous evaluation of the feasibility of Facebook recruitment remains scant. In particular, there is a lack of explicit detail about Facebook advertising settings and strategies, content of advertisements, or modifications to Facebook settings and advertising campaigns throughout recruitment (exceptions include the studies by Leach et al and Arcia [15,23]). It is, therefore, difficult to draw accurate conclusions about the feasibility and effectiveness of Facebook as a research tool. To address these gaps, we conducted a methodological evaluation to assess the feasibility of using Facebook to recruit a national sample of employed parents to an online survey, with respect to (1) recruitment rates and costs, (2) representativeness (compared with a national population-based cohort of employed

XSL•FO
RenderX

Australian parents), and (3) retention (including retention rate and demographic characteristics of those who returned to complete a follow-up survey 6 to 8 weeks later).

Methods

The evaluation reported here is based on the protocol and data collected for an online survey, the Families at Work Survey. We first describe this survey and the recruitment protocol, followed by the evaluation method.

The Families at Work Survey

The Families at Work online survey aimed to identify the employment conditions and workplace supports accessed by employed parents of children aged 0 to 18 years to manage work and family demands and to identify which strategies were associated with better parent well-being. Participants were required to be (1) aged 18 years or older, (2) living in Australia, (3) in paid employment (including self-employment or employees currently on leave), and (4) the parent of at least one child aged 18 years or younger. Upon survey completion, parents were invited to be contacted for a 6- to 8-week follow-up survey. The 15-min baseline survey (T1) was conducted between August and November 2016 (including school-term and holiday periods in Australia). The 10-min follow-up survey (T2) readministered core demographics and primary outcome measures and was conducted from October 2016 to February 2017.

Participants were invited to enter a draw to win 1 of the 10 Aus $50 gift cards at the end of each survey; winners were randomly selected and sent a gift card via email. The survey was administered via Qualtrics (Qualtrics Provo) [24]. Ethical approval was granted by La Trobe University Human Ethics Committee (S16-122).

Recruitment Protocol

A study-specific Facebook page was created before recruitment that included a study description and contact details of the research team. The page featured the university logo to support the perceived legitimacy of the survey. Relevant content was regularly posted to the page, such as news articles about managing the demands of work and family life and updates on the number of survey respondents. Participants were recruited through either (1) paid Facebook advertising or (2) free Facebook advertising. Both methods included passive *snowball* sampling, as users liked, shared, or circulated the link to others. Survey advertisements comprised the following: (1) a title (eg, “When it comes to balancing work and family life, what works for you? What do you find tricky?”), (2) an image (eg, mother hugging child, father walking child to school), (3) main text (eg, “Researchers are looking for working parents of children aged 18 years or younger to complete a 15-minute survey. You can go into a draw to win a $50 gift card!”), and (4) the survey link.

Paid Facebook Advertising

A total of 20 paid advertising campaigns were run at an overall cost of Aus $7969, approximately 1 campaign per week, most of which ran for 7 days each. A Facebook *campaign* has an overarching objective (eg, increase clicks to a website), targets specific Facebook users (eg, gender, age, location), and has a budget [25]. Data collection ceased after the 20th campaign due to a sufficiently large and diverse sample being obtained. The study was also advertised for a small fee (Aus $200) on a popular online single-parent community, whereby an administrator promoted the survey to members via email and across numerous single-parent Facebook pages.

Each paid advertising campaign was designed using Facebook’s Ads Manager, for which we selected the intended audience, schedule, format, and budget. Advertisements were displayed to users whose profiles indicated that they lived in Australia, were aged 18 to 60 years, and who matched on specific demographics, interests, or behaviors (eg, *mother*, *father*, *work-life balance*). Campaigns either targeted *parents* (all) or specifically targeted an underrepresented subgroup of our sample (eg, *fathers*, *fathers with teens*). For campaigns that specifically targeted mothers or fathers, users who identified their gender as *female* or *male* were selected in the Facebook Ads Manager to ensure that the advertisement was only shown to female or male users, respectively (Note: we have used the terms *gender*, *male*, and *female* for consistency with the Facebook Ads Manager, although we acknowledge the complexity and distinctions between sex, gender, and parenting roles). Campaigns targeting parents more broadly did not specify gender. A range of high-resolution stock images were used in the advertisements. Campaigns were closely monitored throughout data collection to identify subgroups of parents who were under-represented in the sample. No adjustments were made to campaigns once they had commenced. Campaigns were not run simultaneously, to avoid competition for the same target population (however, some overlap occurred with mother- and father-specific campaigns, given that these were targeting different populations). Features of the 20 campaigns are presented in Table 1. All advertisements were placed on Mobile News Feeds and Desktop News Feeds, given that this placement is highly visible and generates the strongest engagement at the lowest cost [26,27]. Users who viewed the paid advertisement on their News Feed could choose to *like*, *tag*, or *share* it with their friends through a passive *snowballing* approach.

XSL•FO
RenderX

Table 1. Characteristics of the 20 paid Facebook campaigns.

Number	Target	Image description	Format[a]	Duration (days)	Prize draw advertised	Time scheduling
1	Parents (all)	Including family, tradespeople, and florist at work	C	7	No	Anytime
2	Blue-collar workers	Including chefs, hairdresser, and factory worker	C	7	No	Anytime
3	Fathers	Including father doing laundry and tradesman at work	C	7	No	Anytime
4	Parents (all)	Including mother and son, father and daughter, and family	C	7	No	Anytime
5	Parents (all)	Mother and daughter and father and son	S	4	Yes	Anytime
6[b]	Parents (all)	Mother and daughter, and father and son	S	7	Yes	12-2 pm; 7-10 pm
7	Fathers	Father and son playing soccer	S	7	Yes	12-2 pm; 6-10 pm
8	Parents (all)	Father carrying daughter on shoulders	S	7	Yes	12-2 pm; 6-10 pm
9	Fathers	Father helping children build a bird feeder	S	7	Yes	12-2 pm; 6-10 pm
10	Parents (all)	Father helping son ride his bike to school	S	7	Yes	12-2 pm; 6-10 pm
11	Fathers	Father and children with a dog beside river	S	7	Yes	Excluding 12-2 pm; 6-10 pm
12	Fathers	Father and children with a dog beside river	S	7	Yes	Excluding 12-2 pm; 6-10 pm
13	Parents (all)	Mother and children	S	7	Yes	12-2 pm; 6-10 pm
14	Parents (all)	Father and son after a bike ride	S	7	Yes	12-2 pm; 6-10 pm
15	Fathers	Father and son playing football	S	5	Yes	Excluding 12-2 pm; 6-10 pm
16	Parents (all)	Mother helping son to do homework	S	7	Yes	12-2 pm; 6-10 pm
17	Fathers	Father and daughter reading a book	S	7	Yes	Excluding 12-2 pm; 6-10 pm
18	Mothers of teens	Mother and teenage son	S	7	Yes	12-2 pm; 6-10 pm
19	Fathers of teens	Father and teenage son	S	7	Yes	12-2 pm; 6-10 pm
20	Regional and rural parents	Family on cattle farm	S	7	Yes	Excluding 12-2 pm; 6-10 pm

[a]Carousel (C) format comprised a set of 5 scrolling images; single format (S) comprised 1 image.

[b]Campaign 6 was the first campaign using the revised survey landing page, with scheduled advertisements run at lunchtime (12-2 pm) and evenings (6-10 pm).

Free Facebook Advertising

To supplement the paid advertising, we used *cross-promotion* to harness Facebook's popularity and reach at no cost. We identified Facebook pages aimed at parents specifically and those used by the general adult population, by searching within Facebook or Google using keywords (eg, *parenting*, *mothers group*, *union*). We contacted page administrators to seek their support in promoting our survey link to their members or followers (see Multimedia Appendix 1 for example wording). The survey link was also circulated using an active *snowballing* approach, whereby the research team asked their professional networks to circulate the survey link through their personal Facebook account or within Facebook groups in which they were members.

Survey Data Collection

Unique survey URLs were generated for each recruitment strategy and advertising campaign, which reflected the source of recruitment (eg, *Campaign 14*). This allowed us to monitor responses to each recruitment method. Facebook users who clicked on the survey link were directed to the survey landing page. For both T1 and T2 surveys, participants were asked to provide electronic consent by selecting 6 statements to demonstrate that they had read and understood the Participant Information Statement and agreed to participate in the survey. Participants who consented to be recontacted were emailed a unique survey link 6 to 8 weeks later. Nonresponders or participants who had partially completed the follow-up survey were sent up to 2 email reminders at weekly intervals.

XSL•FO
RenderX

Evaluation Method

Facebook Metrics

Facebook metrics were collected through the Facebook Ads Manager, including reach (ie, the number of users who saw the adverts in their News Feed at least once), link clicks (ie, the number of users who clicked on the advertisement), cost per click (ie, campaign cost divided by the number of link clicks), and relevance (a score out of 10 generated by Facebook, which estimates how well the target audience is responding to the advertisement). Higher relevance scores indicate positive user engagement (eg, link clicks), whereas lower scores indicate negative interactions (eg, hiding or reporting an advertisement) [28].

Recruitment Rates and Costs

Recruitment rates and costs were calculated following the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [29]. The CHERRIES framework is designed to guide the quality reporting of online surveys in the same way as the Consolidated Standards of Reporting Trials statement guides the reporting of randomized controlled trials. As such, we report several metrics that speak to the quality and completeness of the data and of the overall feasibility of Facebook recruitment (that is, was a useable sample and dataset obtained?). These include the following: *view rate*: the ratio of Facebook users who clicked on the advertisement and visited the survey landing page divided by users who saw the advertisement (click to reach ratio); *participation rate*: the ratio of those who consented to participate divided by the number of visitors to the survey landing page (consent to click ratio); and *completion rate*: the ratio of the number of people who completed the survey divided by those who consented to participate (completion to consent ratio). Cost per consent and cost per completer for each campaign were also derived. It was not possible to calculate the view rates or response rates for our free recruitment strategies.

Determining Sample Representativeness

Comparison data were drawn from the Longitudinal Study of Australian Children (LSAC), Kindergarten (K) cohort (child age: 4-5 years at recruitment in 2004). Full sample details, design, and field methods are published elsewhere [30]. Briefly, LSAC employed a 2-stage cluster sampling design using Australian postcodes and Australia's universal health insurance database (Medicare Australia) to recruit parents through a mailout. The LSAC sample is considered broadly representative of all Australian children and their parents. Data are collected biennially (since 2004) via a face-to-face interview with parents and a parent-report questionnaire [31,32]. Data from employed parents were compared with this study's sample on baseline demographic characteristics and primary outcome measures at 3 waves: wave 1 (child age: 4-5 years), wave 4 (child age: 10-11 years), and wave 6 (child age: 14-15 years).

The demographic characteristics used for comparison were *marital status* (married or de facto; single); *country of birth* (born outside Australia; born in Australia); *education* (with or without tertiary qualification); and *neighborhood disadvantage*, assessed using the Socio-Economic Index of Areas (SEIFA) Disadvantage score [33]. Postcodes provided by the participants were matched with the corresponding SEIFA score (Australian *mean* 1000). To assess for self-selection bias, we also compared participants on the 2 main survey measures: work-family conflict and psychological distress. *Work-family conflict* was measured using 4 items on a 5-point scale, from 1="strongly disagree" to 5="strongly agree", adapted from Marshall and Barnett [34] and used widely in Australian research (eg, [35,36]). Scores across the 4 items were averaged, with higher scores indicating greater work-family conflict (alpha=.67). *Psychological distress* was assessed using the K6 [37] on a 5-point scale, from 1="none of the time" to 5="all of the time". Responses were summed (range 6-30), with higher scores indicating greater psychological distress (alpha=.87).

Statistical Analyses

To determine Facebook recruitment rates and costs (Aim 1), survey data were exported from Qualtrics into StataSE14 (StataCorp) [38] and the number of consenting participants and completed surveys were summarized by the recruitment source (identified by the unique survey URLs). Facebook campaign costs were then summarized by the number of participants who provided consent (cost per consent) and the number of participants who completed the survey (ie, reached the end of the survey and clicked *submit*; cost per completer). To assess sample representativeness (Aim 2), baseline (T1) demographic characteristics were compared with LSAC waves 1, 4, and 6. Only employed parents from the LSAC sample were used, to provide a meaningful comparison. Independent sample *t*-tests were used to compare continuous variables (ie, work-family conflict, psychological distress, neighborhood disadvantage), and chi-square tests were used to compare categorical variables (ie, educational attainment, marital status, country of birth). To assess participant retention (Aim 3), the number of participants who completed the follow-up survey (ie, reached the end of the survey and clicked *submit*) were compared with those who consented to be recontacted and with those who consented to the initial study. Demographic characteristics of T1 and T2 participants were compared using independent samples *t*-tests and chi-square tests, as appropriate.

Results

Survey Recruitment (T1)

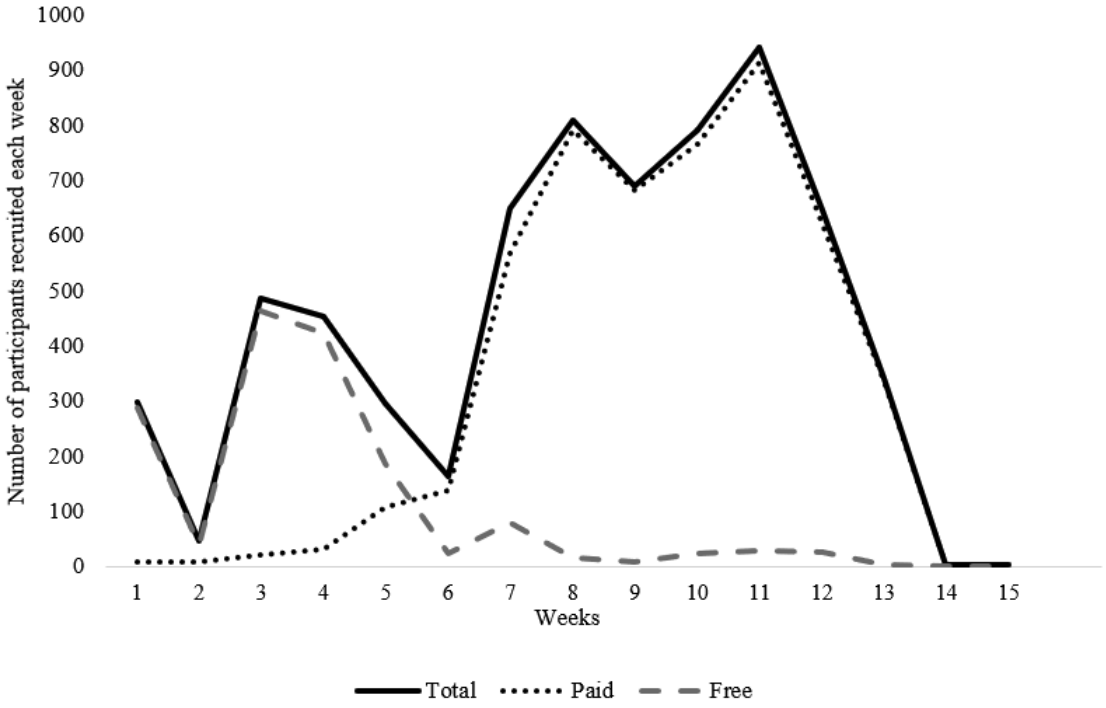
After a 15-week recruitment period, there were a total of 6653 clicks on the survey link, resulting in 5378 eligible participants consenting and commencing the T1 survey. Of these 5378, 4665 (86.74%) participants completed the survey (ie, pressed *submit* at the end of the survey) and a further 532 (9.89%) participants provided partial data (ie, exited before pressing *submit*); however, 181 (3.37%) participants provided consent but did not answer any questions. The proportion of consenting participants who provided complete, partial, or no data did not differ for the paid and free methods. Of the 4665 participants who provided complete data, 3440 (73.74%) were recruited through the 20 paid Facebook advertising campaigns, 79 (1.69%) through other paid online advertising, 782 (16.76%) through free Facebook advertising, and 364 (7.80%) through

XSL•FO
RenderX

circulation within the Facebook accounts of our professional networks.

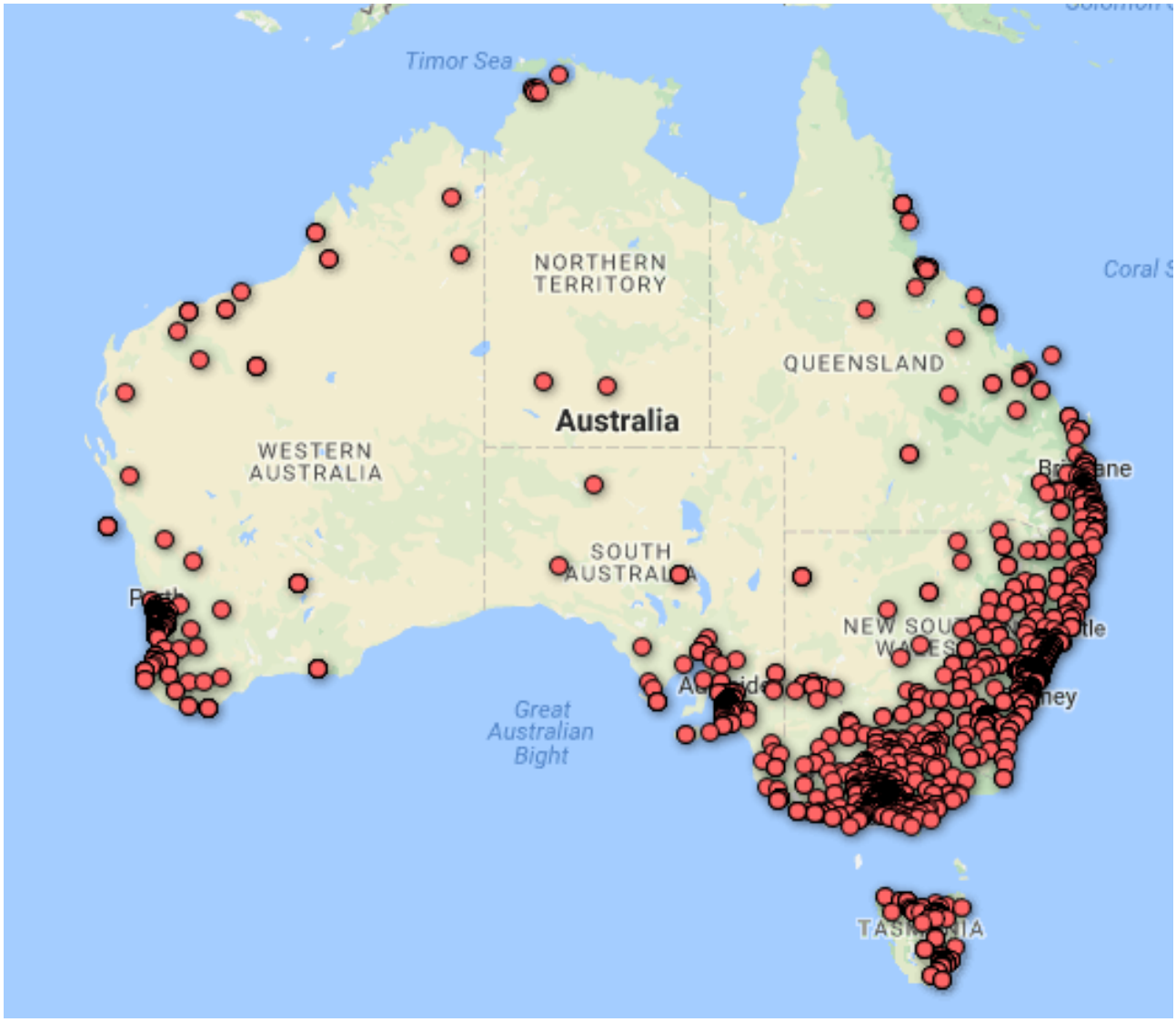
As shown in Figure 1, free advertising was the main source of recruitment during the first 5 weeks of data collection. After this time, our paid advertising campaigns had been fine-tuned to increase effectiveness and thus became the main recruitment method. This allowed a reduced focus on free methods, given that we had exhausted opportunities to post on pages with which we had established connections. In weeks 3 and 7, efforts to encourage circulation of the survey link within personal Facebook accounts of our professional networks were also effective. A pivotal point in recruitment can be observed around weeks 5 and 6, whereby changes to paid advertising strategies (described below) were implemented. Most participants completed the survey on a mobile device (72.80% [3396/4665]), followed by tablet (12.11% [565/4665]), laptop (8.00% [373/4665]), and desktop computer (7.01% [327/4665]). Respondents represented all Australian states and territories, with greater concentrations in or around the more populous cities of Melbourne, Sydney, Brisbane, Adelaide, Perth, and Hobart (see Figure 2).

Figure 1. Paid versus free recruitment rates during the T1 survey.



XSL•FO
RenderX

Figure 2. Geographical distribution of participants (for 5213 participants who provided postcodes).



Paid Facebook Advertising

The 20 paid Facebook advertising campaigns reached nearly half a million users, with a view rate (click to reach ratio) of 1.87%, a participation rate (consent to click ratio) of 47.80%, and a completion rate (completion to consent ratio) of 86.04%. The total cost of the paid Facebook advertisements was Aus $7969.25, with an average cost of Aus $1.99 per consenting participant and Aus $2.32 per completed survey. Table 2 summarizes the results of the 20 paid advertising campaigns, including the overall cost, reach, relevance, clicks, cost per click, cost per consenting participant, and cost per completed survey. As per our methodology, campaigns were regularly monitored, and subsequent campaigns were adjusted accordingly. The cost per consent started at Aus $31.79 and dropped to Aus $7.52 by Campaign 5. This decrease marked a shift in our recruitment strategies; we mentioned our incentive (ie, gift card prize draw) and changed from a multiple- to a single-image format. A further decrease in cost per consent occurred at Campaign 6 (Aus $2.54), after we reduced the text on the survey landing page and scheduled our advertisements to appear during specific time slots. Although cost per consent fluctuated across subsequent campaigns, it remained relatively low, ranging from Aus $0.68 (Campaign 14) to Aus $4.86 (Campaign 7) per consenting participant. Adjustments made to campaigns throughout data collection did not markedly change the cost per click but did improve the completion rates.

XSL•FO
RenderX

Table 2. Recruitment rates and costs for paid Facebook advertising campaigns.

Number	Cost (Aus $)	Reach[a]	Relevance score[b]	Link clicks (view rate), n (%)	Cost per click (Aus $)	Consented to participate (participation rate), n (%)	Cost per consent (Aus $)	Completed survey (completion rate), n (%)	Cost per completer (Aus $)
1	349.67	14,832	6	428 (2.89)	0.82	11 (2.57)	31.79	9 (81.82)	38.85
2	339.01	10,448	8	459 (4.39)	0.74	7 (1.53)	48.43	4 (57.14)	84.75
3	336.32	16,058	6	371 (2.31)	0.91	8 (2.16)	42.04	7 (87.50)	48.05
4	336.64	17,599	6	383 (2.18)	0.88	8 (2.09)	42.08	6 (75.00)	56.11
5	188.11	10,488	4	134 (1.28)	1.40	25 (18.66)	7.52	24 (96.00)	7.84
6	350.00	17,408	6	271 (1.56)	1.29	138 (50.92)	2.54	122 (88.41)	2.87
7	350.00	19,995	4	184 (0.92)	1.90	72 (38.59)	4.86	62 (86.11)	5.65
8	350.00	26,152	8	605 (2.31)	0.58	415 (59.11)	0.84	353 (85.06)	0.99
9	700.00	35,087	2	348 (0.91)	2.01	187 (51.72)	3.74	162 (86.63)	4.32
10	350.00	21,648	6	347 (1.60)	1.01	202 (58.50)	1.73	180 (89.11)	1.94
11	700.00	45,272	7	976 (2.16)	0.72	496 (50.82)	1.41	421 (84.88)	1.66
12	700.00	38,121	7	707 (1.85)	0.99	331 (46.68)	2.11	265 (80.06)	2.64
13	350.00	24,249	8	499 (2.06)	0.70	343 (68.94)	1.02	298 (86.88)	1.17
14	350.00	26,080	9	702 (2.69)	0.50	514 (73.65)	0.68	460 (89.49)	0.76
15	399.06	18,576	6	328 (1.77)	1.22	187 (57.62)	2.13	157 (83.96)	2.54
16	316.37	18,768	8	375 (2.00)	0.84	321 (62.93)	0.99	285 (88.79)	1.11
17	604.40	30,750	4	375 (1.22)	1.61	270 (48.27)	2.24	235 (87.04)	2.57
18	300.00	19,920	8	460 (2.31)	0.65	290 (64.57)	1.03	247 (85.17)	1.21
19	300.00	16,788	4	166 (1.00)	1.81	66 (42.17)	4.55	55 (83.33)	5.45
20	299.67	18,548	6	246 (1.33)	1.22	107 (43.90)	2.80	88 (82.24)	3.41
Total	7969.25	44,6787	—[c]	8364 (1.87)	—	3998 (47.80)	—	3440 (86.04)	—
Mean	—	—	6.2	—	1.50	—	1.9	—	2.32

[a]Reach is the number of users who saw the adverts in their News Feed at least once.

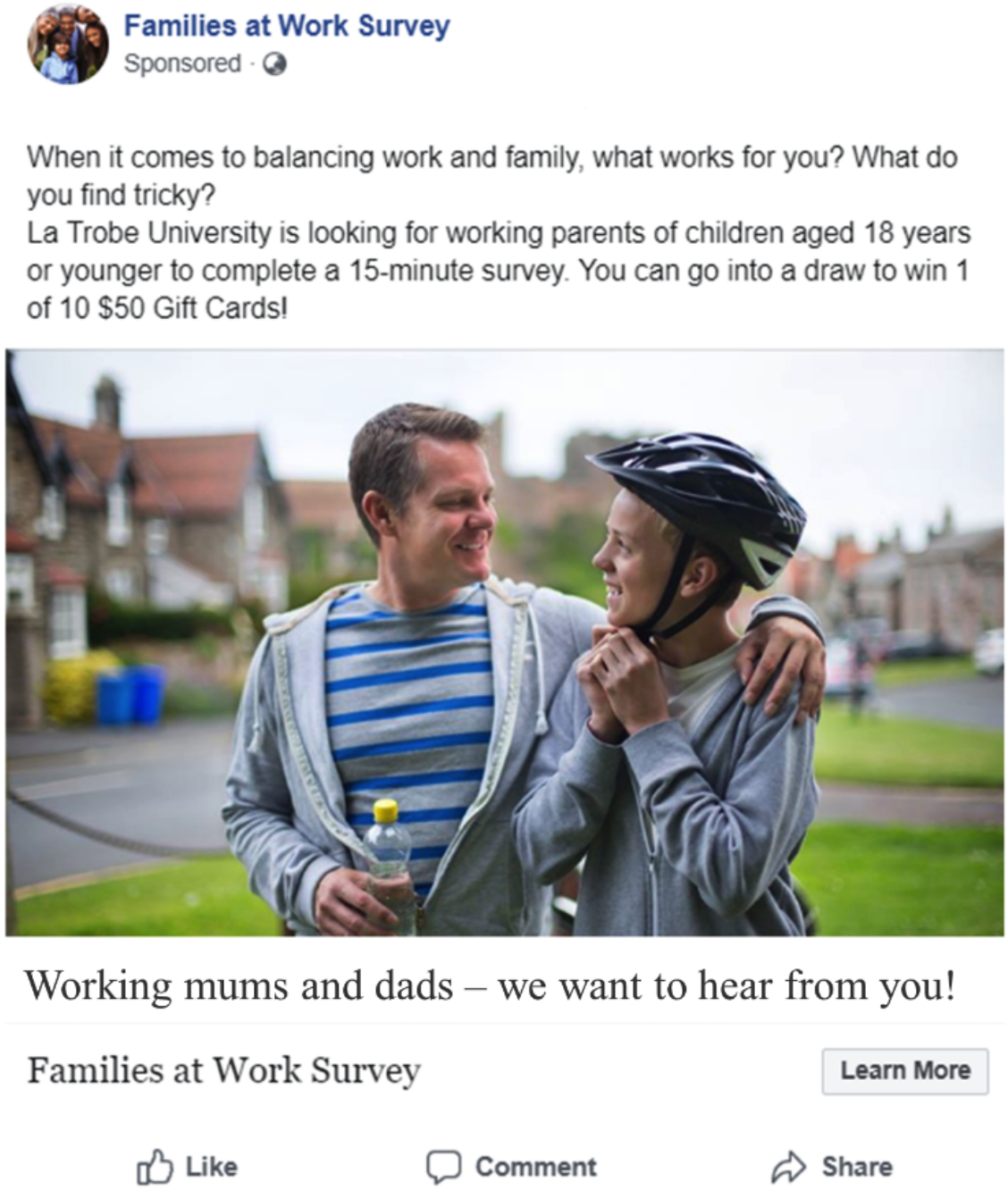
[b]Relevance score is out of 10 and estimates how well the target audience is responding to the advertisement; higher scores indicate positive engagement.

[c]Not applicable.

Although the changes implemented at Campaigns 5 and 6 improved the participation rates, we continued to monitor participant demographics and refine our advertising campaigns. For example, Campaign 6 targeted parents generally (ie, no gender specified), reaching 17408 users and attracting 271 link clicks. However, only 13.56% (2361/17,408) of views and 4.06% (11/271) of clicks were from male users. Overall, 8 of the 20 campaigns targeted male users (fathers), each of which recruited between 83.94% and 100% males. In comparison, the 11 campaigns in which no gender was specified, each recruited between 0% and 12.53% males. Thus, the father-specific campaigns effectively boosted participation from fathers; overall, the paid advertising campaigns recruited 1540 fathers (44.77% of the total sample recruited through the paid campaigns). Parents in regional Australia and parents of teenage children were also under-represented, prompting 3 new campaigns targeting regional parents, mothers of teens, and fathers of teens. The most successful paid campaign was Campaign 14 (Figure 3), which recruited 514 participants at a cost of Aus $0.68 per consented participant (ie, Aus $0.76 each for 460 completed surveys). This campaign coincided with *back to school* week, in which Australian children return to school after the holidays; the image was selected to maximize salience for parents during this period.

XSL•FO
RenderX

Figure 3. Most successful Facebook campaign (Campaign 14), which generated 514 survey consents at Aus $0.68 each (Aus $0.76 per completed survey) and a relevance score of 9 out of 10.



Free Facebook Advertising

We contacted the administrators of 85 Facebook pages, from which 22 responded, 13 agreed to endorse and cross-promote the study on their Facebook page, 6 invited us to post directly on their page as a visitor, and 3 declined as they did not promote surveys on their page. We did not receive a response to 73% of requests. Of the 6 Facebook pages that provided permission to post as a visitor, 1 post led to the completion of 6 surveys, whereas 5 posts generated none. In comparison, 8 Facebook administrators endorsed our survey on their page, resulting in the completion of 492 surveys. The success of each endorsed post generally reflected its number of *followers*. For example, a prominent parenting Facebook page with almost 200,000 followers led to the completion of 355 surveys, whereas a parenting page with 3000 followers generated 11 surveys. In all, 4 administrators shared our survey link with their members via email (generating 70 surveys) and 1 posted the survey link in their forum (generating 138 surveys). An additional 76 parents completed the survey via the link displayed on our study-specific Facebook page. Our requests for professional networks to circulate the survey link within their own Facebook account and groups were effective, recruiting 364 participants with complete data (7.8% of the full sample).

Retention (T2)

Of the 4665 participants who completed the T1 survey (ie, reached the end of the survey and clicked *submit*) 85.94%

XSL•FO
RenderX

(4009/4665) agreed to be contacted for the follow-up survey. Of these, 2463 (61.44%) participants consented to the T2 survey: 35.74% (1433/4009) responded to the initial email invitation, 15.71% (630/4009) responded to the first email reminder, and 9.98% (400/4009) to the second. One-third (32.98%, 1322/4009) of the participants did not respond to the email invitation or either reminder, a further 4.81% (193/4009) did not provide an email address in the T1 survey despite agreeing to be recontacted, and 31 (0.77%) provided an invalid email address. A small number of participants who consented at T2 provided no data (0.65%, 16/2463) or partial data (5.28%, 130/2463). Of the 4009 participants who agreed to be contacted for follow-up, 57.79% (2317) completed the T2 survey (ie, 43.08% [2317/5378] of parents who consented to the initial survey).

Sample characteristics of participants who completed the T1 and T2 surveys were compared. There were similar proportions of single parents (14% T1 vs 15% T2; *P*=.55) and parents born outside Australia (18% T1 vs 19% T2; *P*=.57) across both surveys. Parent age was also similar (mean 40.3 T1 vs mean 40.4 T2; *P*=.53), as was neighborhood disadvantage (SEIFA score mean 1019 T1 vs mean 1020 T2; *P*=.26). More males completed the T1 survey, compared with the T2 survey (32% T1 vs 27% T2; *P*<.001). Finally, T1 participants experienced greater work-family conflict (mean 3.4 T1 vs mean 3.3 T2; *P*<.001) and greater psychological distress (mean 11.5 T1 vs mean 11.2 T2; *P*=.005) than T2 participants.

Sample Representativeness

Demographic characteristics for T1 were compared with a representative sample of employed Australian parents participating in the LSAC. To compare across child age ranges, we repeated the comparisons with LSAC waves 1, 4, and 6. Similar findings were observed across each wave; therefore, only wave 1 (data collected in 2004) is reported here for comparative purposes (full results available on request). Compared with LSAC, the current sample consisted of more single parents (15% vs 14%; *P*<.001) and parents with tertiary education (65% vs 44%; *P*<.001) and fewer parents born outside Australia (18% vs 23%; *P*<.001). Parents in our sample also lived in less disadvantaged neighborhoods (SEIFA score mean 1018.7 vs mean 1011.1; *P*<.001) and reported greater psychological distress (mean 11.5 vs mean 10.1; *P*<.001) and greater work-family conflict (mean 3.4 vs mean 2.6; *P*<.001).

Discussion

Principal Findings

This paper reports on the feasibility of using Facebook to recruit a sample of working Australian parents to an online survey; this is one of few studies to systematically evaluate participant recruitment using social media for health research purposes in Australia (see also [15,27]). We report on the key parameters of recruitment rates and costs, sample representativeness, and retention, following the CHERRIES framework [29] for reporting online surveys. Our findings have implications for other researchers seeking to recruit participants through social media and also contribute to the emerging evidence about the nature of samples recruited using these methods.

Overall, the combination of paid and free Facebook advertising proved to be an effective strategy for recruiting a large sample of employed parents (4665 parents completed the survey over a 15-week period). After initially poor engagement with paid advertisements and high per-participant costs, adjustments to our paid advertisements resulted in significantly improved recruitment and participation rates (eg, advertising a prize draw incentive, reducing the amount of text on the survey landing page, using a single-image advertisement). These improvements reduced the need to focus our efforts on free methods. It should be noted that a relatively large proportion of participants (13.3% overall) indicated their consent but provided either no data or partial data. It has been reported elsewhere that participants recruited via social media may be less "conscientious" than those recruited using more traditional methods [18]. As such, we suggest that iterative methods and careful monitoring are required during Facebook recruitment.

We provide further support for the use of targeted Facebook recruitment as a low-cost means of recruiting parents [15,27,39]. Facebook may be a particularly versatile tool in today's tight monetary environment, in which researchers are under increasing pressure to obtain competitive funding and minimize research costs [40]. Our average cost per survey was substantially lower than most of those conducted to date (paid campaigns only: Aus $2.32 per survey; overall: Aus $1.75 per survey), with the exception of Leach et al [15], who achieved Aus $0.55 for a survey targeting women postpartum. The success of the Leach et al survey may reflect the salience of the research topic (ie, the "Living with a Young Baby" survey) and the high rate of Facebook use among new mothers [3]. Our recruitment costs also compare favorably with more traditional recruitment methods. For example, the authors of this study previously recruited 2002 Australian parents to complete a survey via Computer Assisted Telephone Interview at approximately Aus $119 per participant (direct research costs only) [41].

Engagement with our paid Facebook advertising campaigns compares favorably with previous studies. Nearly 2% of Facebook users exposed to our advertisements clicked on the link, and of those who did, nearly half (47.80%) consented to participate. The average click-through rate (1.87%) was also higher than reported elsewhere [21,23,42]. However, as illustrated in Table 2, the number of Facebook users an advertisement reaches or the number of *likes*, *shares*, and *clicks* it generates does not necessarily translate to survey completion.

Our sample of working Australian parents was broadly representative, compared with a large population-based sample. Consistent with Leach et al [15], some self-selection bias was evident in that we recruited parents experiencing more of the constructs being examined (ie, work-family conflict, psychological distress). It is likely that parents who viewed our advertisement about "juggling work and family life" were more likely to respond if they perceived this to be a salient issue. Some demographic characteristics, although statistically significant due to the large sample size, were not meaningfully different (ie, proportions of single parents, neighborhood disadvantage). Our sample, however, did under-represent parents with lower educational attainment and parents born outside

XSL•FO
RenderX

Australia, which is consistent with findings reported elsewhere [15,43]. Although 97% of Australian families with children under 15 years have internet access [44], those on low incomes are less likely to be internet users than those on high incomes, and migrants from non-English–speaking backgrounds are less likely to be online than their Australian-born counterparts [45]. It is possible that our method of harnessing professional networks to circulate the survey link over-recruited tertiary-educated parents; however, this tends to occur for both online and offline research methodologies [46].

Less than half of those initially recruited returned to complete the follow-up survey 6 to 8 weeks later, which is lower than the retention rates reported elsewhere (eg, [22]). This may reflect limited ongoing participant engagement for those recruited through social media, compared with face-to-face. Indeed, Frandsen et al [18] concluded that participants recruited via social media tended to be less invested or conscientious than those recruited via more traditional means such as flyers or newspapers. It is possible that some Facebook users impulsively click on a link out of curiosity, compared with more active engagement that might be required with traditional methods (eg, emailing the research team).

Facebook allows for a flexible and dynamic approach to recruitment, whereby strategies can be continually monitored, adjusted, and evaluated for effectiveness. This requires researchers to be highly responsive and open to trying different advertising strategies [46]. Although social media recruitment can lead to less contained or trackable recruitment than offline methods, users' ability to *like*, *share*, or *tag* other users in response to an advertisement can support ongoing *snowballing* and broad reach. A particular strength of our evaluation method was the use of unique survey URLs for each advertising campaign [24]. This enabled us to identify the specific campaign a participant engaged with. Facebook also provides a multitude of metrics through its Ads Manager for users to track the effectiveness of advertising campaigns and engagement with posted content. We used a flexible approach to develop targeted campaigns to reach specific subgroups, such as parents of teenagers, blue-collar workers, and those living in rural or remote areas.

Fathers were more difficult to recruit than mothers, requiring specific, targeted campaigns with explicit calls to action. This may reflect fewer fathers engaging in Facebook or around parenting-related topics [3], but also reflects evidence that fathers are typically under-represented in research [47,48]. Interestingly, advertisements targeted at *parents* engaged few fathers, and it became apparent that fathers required *specific* calls to action if they were to take part. This may be due, in part, to fathers engaging with less parent-related Facebook content than mothers or that the term *parents* was viewed as a reference to mothers rather than fathers. It may also be a product of Facebook functionality, whereby the campaign is presented to users who are similar to those who have already engaged with it (ie, mothers). This necessitates the use of different recruitment strategies for mothers and fathers.

The gradual adjustment of advertising strategies across the duration of recruitment allowed us to identify the most effective means of engaging our target population. A single-image advertisement was more successful than the multiple-image *carousel* format, although Facebook reports that "carousel link ads drive 30-50% lower cost per conversion and 20-30% lower cost per click than single-image link ads" [49]. It is possible that a single relevant image was most salient to our population of working parents or that the *carousel* format is more effective for campaigns that aim to promote a range of goods or products.

Limitations

We acknowledge several limitations of our evaluation of Facebook for recruitment. First, we made several changes to our recruitment strategy for campaigns 5 and 6, which resulted in a substantial improvement in participant recruitment rates and per-participant costs. We acknowledge that this change in format coincided with the decision to specifically mention our prize draw in the advertisement. It is, therefore, difficult to disentangle the relative contributions of these strategies to recruitment rates. Second, we compared our sample with a large cohort study (LSAC), and although it is considered broadly representative of the Australian population, it is not a contemporary sample given that wave 1 data were collected in 2004. Third, evaluation of Facebook as a recruitment tool can be impacted by its advertising algorithms and metrics, which are often difficult to comprehend and can change without notice [50]. This can pose problems for researchers interpreting metrics or seeking to replicate previously published Facebook recruitment protocols. In January 2018, Facebook founder, Mark Zuckerberg, announced a new algorithm that would result in "less public content like posts from businesses, brands and media" [51]; such changes are likely to impact the way in which Facebook is used for research. The constant evolution of Facebook also makes between-study comparisons problematic, which is exacerbated by the lack of consistency with which existing studies report on key parameters such as cost (eg, studies may report total cost, cost per click, cost per participant, or cost per completer).

Practical Challenges

We encountered several practical challenges during recruitment. Both paid and free methods were surprisingly time-intensive, requiring regular monitoring of response rates (daily in the first few weeks, reducing to weekly by the end of the recruitment period), designing new campaigns, contacting page administrators, and developing content for the study Facebook page. Currently, Facebook allows page administrators to block posts but not comments. Our decision to set the profanity filter to *strong* and to use a large selection of *moderation words* was effective; only a small number of offensive or negative comments were posted and were automatically hidden, keeping administrative requirements to a minimum. During the early phase of recruitment, our private messages to Facebook page administrators about promoting our survey were labeled as *spam*. Our account was subsequently blocked from posting or messaging for 1 month; this required the creation of a new account, to enable active recruitment to continue. Facebook does not provide any direct support service; therefore, we needed to rely on Facebook forums or the expertise of peers. Another challenge was the low response rate from administrators of

XSL•FO
RenderX

Facebook pages regarding requests to support our research by posting the survey link. It is possible that some Facebook pages or groups receive many requests from researchers or that our message was simply disregarded as *spam*. A preexisting connection with the group or page generally led to a greater likelihood of a response; therefore, researchers are encouraged to draw on personal or professional networks and to contact individuals directly (eg, via email) where possible. We also recognize that this free advertising is not necessarily a sustainable approach in the long term, given the potential risks of survey fatigue or of oversampling specific subgroups of the population.

Further Research and Recommendations

Future evaluations of Facebook recruitment may seek to harness qualitative methodologies to understand the reasons participants choose to engage or not engage in research advertised on Facebook, including the features of advertising campaigns and survey interfaces that may be most appealing for specific target groups, and the factors that may improve participant retention. This would be particularly helpful for fathers, who were more difficult than mothers to recruit and retain. It is also recommended that recruitment strategies are amended sequentially, rather than simultaneously, to identify the effectiveness of specific strategies with greater precision. Although we found Facebook to be an effective tool for the recruitment of working Australian parents, further research is needed to determine its feasibility for nonparent or unemployed populations and to examine the factors affecting retention following Facebook recruitment.

We recommend that careful consideration is given to engaging participants at each step of the way from viewing the advertisement, providing consent, to survey completion. An advertisement must be relevant and interesting; the survey landing page must be clear and concise with friendly, plain language; and the survey itself must be straightforward and not unreasonably lengthy. As documented elsewhere [46], regular communication with an institutional human research ethics committee is essential, from the study design phase and throughout the recruitment phase. However, the online space is remarkably dynamic, with social media platforms and functionality often changing rapidly and unpredictably. Researchers must stay on the cutting edge of what platforms are popular and how they function. Seeking guidance from an information technology or social media specialist during project design and implementation may be beneficial.

Conclusions

Findings suggest that Facebook has the potential to be a low-cost means of recruiting a large sample of working Australian parents, which is an important consideration given the competitive funding environment in which researchers work. A significant barrier is the ever-changing nature and functionality of social media; researchers may benefit from the support of social media professionals. Although we focus on the recruitment of parents, our methodology is applicable to the recruitment of other populations, providing access to real-time feedback and allowing recruitment gaps to be addressed using targeted campaigns. Our experience suggests that immediate success is unlikely; rather, sufficient lead-in time is required to build interest and momentum and to monitor and adjust recruitment strategies accordingly. Fathers were unlikely to respond to calls for *parents* but required specific invitations to *dads* using gender-specific campaigns. There was evidence of self-selection bias, given that we recruited parents with greater work-family conflict and psychological distress than the general parent population. Participants were also more highly educated and less likely to have been born outside of Australia than the general parent population. Retention to follow-up was less likely for males or for participants experiencing high work-family conflict and psychological distress. Further evidence is needed to understand the mechanisms of engagement and retention for populations recruited through social media.

Acknowledgments

The “Families at Work” study was funded by La Trobe University, Transforming Human Societies Research Focus Area and conducted by SB, AC, SC, SH, NJH, JN (La Trobe University), LL (Australian National University), and CN (Murdoch Children’s Research Institute). Authors SB, AC, SC, SH, NJ, and JN were additionally supported by the Roberta Holmes Transition to Contemporary Parenthood Program at La Trobe University. We acknowledge the contribution of Professor Lyndall Strazdins (Australian National University) to the design and implementation of the Families at Work study. This article uses unit record data from Growing Up in Australia, the Longitudinal Study of Australian Children. The study is conducted in partnership between the Australian Government Department of Social Services (DSS); the Australian Institute of Family Studies (AIFS); and the Australian Bureau of Statistics (ABS). The Longitudinal Study of Australian Children study design and data collection were funded by Australian Government Department of Social Services.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example messages sent to Facebook page administrators.

[PDF File (Adobe PDF File), 27KB - jmir_v21i3e11206_app1.pdf]

XSL•FO
RenderX

References

1. Vercruyssen A, Roose H, Carton A, van de Putte B. The effect of busyness on survey participation: being too busy or feeling too busy to cooperate? Int J Soc Res Methodol 2014;17(4):357-371. [doi: 10.1080/13645579.2013.799255]
2. Cooklin AR, Westrupp EM, Strazdins L, Giallo R, Martin A, Nicholson JM. Fathers at work: work-family conflict, work-family enrichment and parenting in an Australian cohort. J Fam Issues 2016;37(11):1611-1635. [doi: 10.1177/0192513x14553054]
3. Bartholomew MK, Schoppe-Sullivan SJ, Glassman M, Kamp Dush CM, Sullivan JM. New parents' Facebook use at the transition to parenthood. Fam Relat 2012;61(3):455-469. [doi: 10.1111/j.1741-3729.2012.00708.x] [Medline: 23671354]
4. Social Media News. 2018. Social Media Statistics Australia - January 2018 URL: https://www.socialmedianews.com.au/social-media-statistics-australia-january-2018/ [accessed 2018-06-01] [WebCite Cache ID 6zqAjyI04]
5. Rose J. "Never enough hours in the day": employed mothers' perceptions of time pressure. Aust J Soc Issues 2017;52(2):116-130. [doi: 10.1002/ajs4.2]
6. Statista. 2018. Number of monthly active Facebook users worldwide as of 3rd quarter 2018 (in millions) URL: https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ [accessed 2018-06-01] [WebCite Cache ID 6zqB2R29C]
7. Sensis. 2017. Sensis Social Media Report 2017 URL: https://www.sensis.com.au/asset/PDFdirectory/Sensis-Social-Media-Report-2017.pdf [accessed 2018-06-01] [WebCite Cache ID 75VEglIxh]
8. Lupton D, Pedersen S, Thomas GM. Parenting and digital media: from the early web to contemporary digital society. Sociology Compass 2016;10(8):730-743. [doi: 10.1111/soc4.12398]
9. Duggan M, Lenhart A, Lampe C, Ellison N. Pew Research Center. 2015. Parents and Social Media URL: http://www.pewinternet.org/2015/07/16/parents-and-social-media/ [accessed 2018-06-01] [WebCite Cache ID 6zqBQGt1Q]
10. Madden M, Cortesi S, Gasser U, Lenhart A. Pew Research Center. 2012. Parents, Teens, and Online Privacy URL: http://www.pewinternet.org/2012/11/20/parents-teens-and-online-privacy/ [accessed 2018-06-01] [WebCite Cache ID 6zqBaejeX]
11. Doty JL, Dworkin J. Online social support for parents: a critical review. Marriage Fam Rev 2014 Mar 04;50(2):174-198. [doi: 10.1080/01494929.2013.834027]
12. Batterham PJ. Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness. Int J Methods Psychiatr Res 2014 Jun;23(2):184-191. [doi: 10.1002/mpr.1421] [Medline: 24615785]
13. Brick JM, Williams D. Explaining rising nonresponse rates in cross-sectional surveys. Ann Am Acad Pol Soc Sci 2013;645(1):36-59. [doi: 10.1177/0002716212456834]
14. Blumberg SJ, Luke JV. Reevaluating the need for concern regarding noncoverage bias in landline surveys. Am J Public Health 2009 Oct;99(10):1806-1810. [doi: 10.2105/AJPH.2008.152835] [Medline: 19696381]
15. Leach LS, Butterworth P, Poyser C, Batterham PJ, Farrer LM. Online recruitment: feasibility, cost, and representativeness in a study of postpartum women. J Med Internet Res 2017;19(3):e61 [FREE Full text] [doi: 10.2196/jmir.5745] [Medline: 28274906]
16. Kapp JM, Peters C, Oliver DP. Research recruitment using Facebook advertising: big potential, big challenges. J Cancer Educ 2013 Mar;28(1):134-137. [doi: 10.1007/s13187-012-0443-z] [Medline: 23292877]
17. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: systematic review. Internet Interv 2016;4(1):72-81. [doi: 10.1016/j.invent.2016.02.001]
18. Frandsen M, Thow M, Ferguson SG. The effectiveness of social media (Facebook) compared with more traditional advertising methods for recruiting eligible participants to health research studies: a randomized, controlled clinical trial. JMIR Res Protoc 2016;5(3):e161 [FREE Full text] [doi: 10.2196/resprot.5747] [Medline: 27511829]
19. Yuan P, Bare MG, Johnson MO, Saberi P. Using online social media for recruitment of human immunodeficiency virus-positive participants: a cross-sectional survey. J Med Internet Res 2014 May 1;16(5):e117 [FREE Full text] [doi: 10.2196/jmir.3229] [Medline: 24784982]
20. Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del Rio C, et al. The comparability of men who have sex with men recruited from venue-time-space sampling and facebook: A cohort study. JMIR Res Protoc 2014 Jul 17;3(3):e37 [FREE Full text] [doi: 10.2196/resprot.3342] [Medline: 25048694]
21. Lohse B. Facebook is an effective strategy to recruit low-income women to online nutrition education. J Nutr Educ Behav 2013;45(1):69-76. [doi: 10.1016/j.jneb.2012.06.006] [Medline: 23305805]
22. Pedersen ER, Naranjo D, Marshall GN. Recruitment and retention of young adult veteran drinkers using Facebook. PLoS One 2017 Mar 1;12(3):e0172972 [FREE Full text] [doi: 10.1371/journal.pone.0172972] [Medline: 28249027]
23. Arcia A. Facebook advertisements for inexpensive participant recruitment among women in early pregnancy. Health Educ Behav 2014 Jun 1;41(3):237-241. [doi: 10.1177/1090198113504414] [Medline: 24082026]
24. Qualtrics. Qualtrics. 2016. URL: https://www.qualtrics.com/au [accessed 2019-01-18] [WebCite Cache ID 75VGMkv3s]
25. Facebook. 2018. Structuring your Facebook campaigns URL: https://www.facebook.com/business/a/campaign-structure [accessed 2018-02-13] [WebCite Cache ID 75VG8jeaV]
26. Ramo DE, Rodriguez TM, Chavez K, Sommer MJ, Prochaska JJ. Facebook recruitment of young adult smokers for a cessation trial: methods,metrics, and lessons learned. Internet Interv 2014 Apr;1(2):58-64 [FREE Full text] [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]

XSL•FO
RenderX

27. Harris ML, Loxton D, Wigginton B, Lucke JC. Recruiting online: lessons from a longitudinal survey of contraception and pregnancy intentions of young Australian women. Am J Epidemiol 2015 May 15;181(10):737-746. [doi: 10.1093/aje/kwv006] [Medline: 25883155]
28. Facebook. 2018. About relevance score URL: https://en-gb.facebook.com/business/help/768598746552352 [accessed 2019-01-17] [WebCite Cache ID 75VGrieNn]
29. 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]
30. Soloff C, Lawrence D, Misson S, Johnstone R. Growing Up in Australia. Melbourne: Australian Institute of Family Studies; 2006. Wave 1 weighting and non-response: LSAC Technical paper No 3 URL: http://data.growingupinaustralia.gov.au/pubs/technical/tp3.pdf [accessed 2018-03-23] [WebCite Cache ID 6zqCLrAzW]
31. Department of Family and Community Services. The Longitudinal Study of Australian Children: 2004 Annual Report. Melbourne: Australian Institute of Family Studies; 2005. URL: https://growingupinaustralia.gov.au/sites/default/files/annualreport2004.pdf [accessed 2018-01-18] [WebCite Cache ID 75VHJLAgj]
32. Soloff C, Lawrence D, Johnstone R. Longitudinal Study of Australian Children Technical Paper No 1: Sample Design. Melbourne: Australian Institute of Family Studies; 2005. URL: https://growingupinaustralia.gov.au/sites/default/files/tp1.pdf [accessed 2018-03-02] [WebCite Cache ID 6zqBtRKfE]
33. Socio-Economic Indexes for Areas. Canberra: Australian Bureau of Statistics; 2013. URL: http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifa [accessed 2018-06-01] [WebCite Cache ID 6zqCeJVID]
34. Marshall NL, Barnett RC. Work-family strains and gains among two-earner couples. J Community Psychol 1993 Jan;21(1):64-78. [doi: 10.1002/1520-6629(199301)21:1<64::AID-JCOP2290210108>3.0.CO;2-P]
35. Cooklin AR, Giallo R, Strazdins L, Martin A, Leach LS, Nicholson JM. What matters for working fathers? Job characteristics, work-family conflict and enrichment, and fathers' postpartum mental health in an Australian cohort. Soc Sci Med 2015 Dec;146:214-222. [doi: 10.1016/j.socscimed.2015.09.028] [Medline: 26520473]
36. Dinh H, Strazdins L, Welsh J. Hour-glass ceilings: work-hour thresholds, gendered health inequities. Soc Sci Med 2017 Mar;176:42-51. [doi: 10.1016/j.socscimed.2017.01.024] [Medline: 28129546]
37. Kessler R, Andrews G, Colpe LJ, Hiripe 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;32(6):959-976. [doi: 10.1017/S0033291702006074] [Medline: 12214795]
38. StataCorp. Stata. College Station TX: StataCorp LP; 2015. Stata Software URL: https://www.stata.com/ [accessed 2019-01-21] [WebCite Cache ID 75aGJeq7e]
39. Gilligan C, Kypri K, Bourke J. Social networking versus Facebook advertising to recruit survey respondents: a quasi-experimental study. JMIR Res Protoc 2014 Sep 17;3(3):e48 [FREE Full text] [doi: 10.2196/resprot.3317] [Medline: 25230740]
40. Edwards MA, Roy S. Academic research in the 21st century: maintaining scientific integrity in a climate of perverse incentives and hypercompetition. Environ Eng Sci 2017;34(1):51-61. [doi: 10.1089/ees.2016.0223] [Medline: 28115824]
41. Bennetts SK, Cooklin AR, Crawford S, D'Esposito F, Hackworth NJ, Green J, et al. What influences parents' fear about children's independent mobility? Evidence from a state-wide survey of Australian parents. Am J Health Promot 2018 Mar;32(3):667-676. [doi: 10.1177/0890117117740442] [Medline: 29166780]
42. Dworkin J, Hessel H, Gliske K, Rudi JH. A comparison of three online recruitment strategies for engaging parents. Fam Relat 2016 Oct;65(4):550-561 [FREE Full text] [doi: 10.1111/fare.12206] [Medline: 28804184]
43. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012 Feb 1;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
44. Australian Bureau of Statistics. 8146.0 - Household Use of Information Technology, Australia, 2016-17. Canberra, Australia: Australian Bureau of Statistics; 2018. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/8146.0 [WebCite Cache ID 73h1bTRjs]
45. Thomas J, Wilson CK, Park S. The Conversation. 2018. Australia's digital divide is not going away URL: https://theconversation.com/australias-digital-divide-is-not-going-away-91834 [accessed 2019-01-17] [WebCite Cache ID 75VIoW6Sk]
46. Loxton D, Powers J, Anderson AE, Townsend N, Harris ML, Tuckerman R, et al. Online and offline recruitment of young women for a longitudinal health survey: findings from the Australian Longitudinal Study on Women's Health 1989-95 cohort. J Med Internet Res 2015 May 4;17(5):e109 [FREE Full text] [doi: 10.2196/jmir.4261] [Medline: 25940876]
47. Costigan CL, Cox MJ. Fathers' participation in family research: is there a self-selection bias? J Fam Psychol 2001 Dec;15(4):706-720. [doi: 10.1037/0893-3200.15.4.706] [Medline: 11770476]
48. Davison KK, Charles JN, Khandpur N, Nelson TJ. Fathers' perceived reasons for their under-representation in child health research and strategies to increase their involvement. Matern Child Health J 2017 Dec;21(2):267-274 [FREE Full text] [doi: 10.1007/s10995-016-2157-z] [Medline: 27473093]
49. Facebook. 2015. Improving Ad Performance with the Carousel Format URL: https://www.facebook.com/business/news/carousel-ads [accessed 2019-01-17] [WebCite Cache ID 75VJ2ImSp]

XSL•FO
RenderX

50. Kayrouz R, Dear BF, Karin E, Titov N. Facebook as an effective recruitment strategy for mental health research of hard to reach populations. Internet Interv 2016 May;4:1-10 [FREE Full text] [doi: 10.1016/j.invent.2016.01.001] [Medline: 30135786]
51. Mosseri A. Facebook Newsroom. 2018. Bringing People Closer Together URL: https://newsroom.fb.com/news/2018/01/news-feed-fyi-bringing-people-closer-together/ [accessed 2018-11-05] [WebCite Cache ID 73h1hFidI]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
LSAC: Longitudinal Study of Australian Children
SEIFA: Socio-Economic Index of Areas
T1: Time point 1 (ie, baseline survey)
T2: Time point 2 (ie, follow-up survey)

Edited by G Eysenbach; submitted 01.06.18; peer-reviewed by A Cyr, J Martin-Kerry; comments to author 02.10.18; revised version received 06.11.18; accepted 10.12.18; published 06.03.19.

Please cite as:
Bennetts SK, Hokke S, Crawford S, Hackworth NJ, Leach LS, Nguyen C, Nicholson JM, Cooklin AR
Using Paid and Free Facebook Methods to Recruit Australian Parents to an Online Survey: An Evaluation
J Med Internet Res 2019;21(3):e11206
URL: https://www.jmir.org/2019/3/e11206/
doi:10.2196/11206
PMID:30839282

©Shannon K Bennetts, Stacey Hokke, Sharinne Crawford, Naomi J Hackworth, Liana S Leach, Cattram Nguyen, Jan M Nicholson, Amanda R Cooklin. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Searching for Information on the Risks of Combined Hormonal Contraceptives on the Internet: A Qualitative Study Across Six European Countries

Paula Gomes Alves[1], PhD; Irene Petersen[2], PhD; Fiona Stevenson[1], PhD

[1]eHealth Unit, Research Department of Primary Care and Population Health, University College London, London, United Kingdom

[2]Department of Primary Care and Population Health, University College London, London, United Kingdom

Corresponding Author:
Paula Gomes Alves, PhD
eHealth Unit
Research Department of Primary Care and Population Health
University College London
Upper 3rd Floor, Royal Free Campus
Rowland Hill Street
London, NW3 2PF
United Kingdom
Phone: 44 02080168011 ext 68011
Email: p.alves@ucl.ac.uk

Abstract

Background: Searching for health information online is increasingly common and is an obvious source of information about oral combined hormonal contraceptives (CHCs) and their risks. However, little is known about how publicly available websites address the risks of CHCs, particularly venous thromboembolism (VTE).

Objective: The aim was to explore the information available to women about VTE and other risks of CHCs on websites available through commonly used search engines.

Methods: A qualitative study was conducted to explore whether and how websites about CHCs in Denmark, Germany, Netherlands, Slovakia, Spain, and the United Kingdom make reference to VTE and other CHC risks. A systematic search procedure was adopted across the six countries, based on relevant keywords. The search was carried out using the Google search engine by fluent/native speakers of each language. A content analysis approach was conducted to extract information from the selected websites.

Results: A total of 357 websites were reviewed. Nearly all (343/357, 96.1%) the websites mentioned VTE as a risk of CHCs, with approximately half referring to other side effects as well. One-fifth (92/357, 25.8%) of the websites provided suggestions about the best contraceptive method to use, and only a minority (23/357, 6.4%) recommended women discuss CHCs with their health professionals. Sites were generally run by the media (110/357 30.8%) or medical services from nongovernmental organizations (140/357, 39.2%). Only a minority of websites referred to organizations such as the European Medicines Agency (11/357, 3.1%).

Conclusions: Despite the large number of websites containing information about oral CHCs and their risks, particularly VTE, only a limited number referred to information from accredited health agency sources. We argue this is a missed opportunity for accredited health agencies to share high-quality information to assist women using CHCs to make informed decisions about contraception.

(J Med Internet Res 2019;21(3):e10810) doi:10.2196/10810

KEYWORDS

oral combined hormonal contraceptives; venous thromboembolism; risks; information sources; internet; health information; qualitative research; content analysis

XSL•FO
RenderX

Introduction

Women using combined hormonal contraceptives (CHCs), which contain both estrogen and progesterone, may be three times more likely at risk of venous thromboembolism (VTE) compared to female nonusers [1]. VTE is a condition that involves formation of a blood clot in a vein and can lead to various health complications (eg, breathing difficulties, chest pain), some of which are life-threatening [2].

According to the European Medicines Agency (EMA), the overall risk of VTE is small and "the benefits of CHC in preventing unwanted pregnancies continue to outweigh their risks" [3]. In their 2014 review, the EMA emphasized that women should be provided with updated information about CHC risks, particularly VTE, so that an informed decision about contraception can be made [4]. It is unclear how women are being provided with this information when prescribed CHCs. A European Union-wide survey conducted by the European Commission in 2014 reported that most Europeans (>75%) believe that the internet is a good resource for health-related information, with 6 in 10 people reporting they "go online when looking for health information" [5]. Therefore, it is likely that women are using the internet to search for information about CHCs. However, little is known about the availability and the type of information that women can find online about CHC-related VTE and other risks of this contraceptive method.

According to the most recent statistics (2017), approximately 90% of females in all EU countries aged 16 to 29 years have used the internet in the last 3 months [6]. The reasons for using the internet are many and may vary between individuals. Similarly, in the 2014 European Commission study, 8 in 10 people aged 15 to 39 years reported using the internet as a health information resource [5]. A recent study by Haluza et al [7] reported that 79% of people referred to the internet as a primary source of health information, outnumbering people using sources such as the doctor (74%), books (63%), or family (62%).

Understanding the information available to women about CHCs and their associated risks is a pressing issue, particularly because previous literature has suggested that women may have knowledge gaps and misconceptions about contraceptives. Previous research has shown that providing women with information about CHC not only improves their knowledge and awareness of risks but also allows them to make a more informed choice [8-10]. Studies have also reported that receiving information about contraception has encouraged women to opt for alternative methods than those they had initially considered [9,10].

It is important to explore the type of information about CHCs and risks available when searching online via common search engines. The goals of this study were two-fold: (1) to identify European websites focusing on oral CHCs and their associated risks, particularly VTE, available through commonly used search engines; and (2) to explore the information provided about CHCs and their risks provided on these websites through a content analysis approach. This study was part of a larger project [11] exploring how women and health professionals communicate about the risks of using CHCs, the process by which women decide to use a CHC, and the information sources that prescribers and women rely on to support their decisions.

Methods

Overview

In this qualitative study we identified websites that provided information about CHCs available through common search engines and analyzed the information extracted from the websites about the risks associated with the use of CHCs. Websites were systematically identified across six countries in Europe: Denmark, Germany, Netherlands, Slovakia, Spain, and the United Kingdom. The selection was made to ensure geographical coverage of southern, central, eastern, northern, and noncontinental European regions.

Development of the Search Protocol

A protocol was developed in English to ensure consistency across the online searches for websites about CHCs across the six countries. Three English lay advisors were consulted for which keywords they would use to search for CHCs and their risks. The three lay advisors were women aged 20 to 30 years. Each woman was asked, "Which words would you use to search for information online about the risks of blood clots when taking the pill?" This led to the identification of keywords that were meaningful and relevant to women (lay terms) and to, as far as possible, ensure that our search would retrieve results that were equivalent to those obtained by women when searching for information about oral CHCs. English-speaking women were used because the protocol was developed in English.

The keywords generated by the three laywomen advisors were then discussed by the research team, who added relevant synonyms to ensure a larger coverage of results (see Figure 1).

After defining the keywords, we moved on to produce a step-by-guide guide on how to perform the online searches. A preliminary version of this search protocol was then revised and piloted by another laywoman who was a native speaker of English; she commented on the clarity and the appropriateness of the language used. The final version of the protocol was delivered in March 2017 and disseminated to the four women recruited to conduct the searches.

Search and Selection of Websites

The search for websites about CHCs was carried out using the Google search engine in each country. We opted for Google because it is considered to be one of the most popular search engines. Google was also used in previous studies about women-related health issues [12].

XSL•FO
RenderX

Figure 1. Keywords and the search strategy used to search for combined hormonal contraceptives.

Search term 1	*Combination word*	*Search term 2*
1.1 The pill OR 1.2 Combined pill OR 1.3 Contraceptive pill OR 1.4 Contraceptive OR 1.5 Birth control pill	AND	2.1 Blood clot OR 2.2 Clot in vein OR 2.3 Thrombosis 2.4 DVT (deep vein thrombosis) OR 2.5 Risk OR 2.6 Side effect

Textbox 1. Inclusion and exclusion criteria for website selection about combined hormonal contraceptives (CHCs).

Inclusion criteria

- Websites targeting women and/or health professionals
- Provides information about oral CHCs
- Contains discussion of the risks of CHCs, including venous thromboembolism
- Updated since the latest guidance from the European Medicines Agency was published (2013)

Exclusion criteria

- Scientific electronic databases with links to publications about this topic
- No reference to oral CHCs and/or their risks
- Non-European websites
- Not updated since the latest guidance from the European Medicines Agency was published (2013)

Due to the large number of results available on Google, we limited our search to the first 10 pages of results. We limited the page results because previous studies have shown that when it comes to online reviews of products, consumers rely on a very limited number of reviews to make their decisions [13]. The limit of 10 pages was chosen by the research team in consultation with another laywoman advisor independent to the study, who reported that in real-life searches it is unlikely that people will search further than this. The websites retrieved from each country were selected according to the criteria detailed in Textbox 1.

Search Procedure and Data Extraction

We recruited four laywomen via the research team network to conduct the Google searches. We opted for laywomen, as opposed to researchers, to conduct the searches because we wanted to replicate how women search for CHCs in naturalistic settings. Also, previous research has suggested that contraceptive responsibilities tend to lay primarily with women [14], making it more likely that searches about CHCs are conducted by women rather than men.

We recruited women of child-bearing age who were native or fluent English speakers, and also native or fluent speakers of at least one of the other five languages of the countries included in the study (see Table 1). We opted for women who were fluent in English to ensure that the same protocol could be used consistently across the team.

Before starting the search, all women were instructed to translate the keywords into their own language. Because the women were considered laypeople, they were instructed to adapt the English keywords to what they would mean in their language. This ensured that all non-English keywords used represented the lay terms that women would potentially rely on to search in the other non-UK countries. The research team discussed these translations with women to ensure they were equivalent to the English terms and that they represented lay, relevant terms in each country.

The internet search was independently conducted by the women in March 2017. They were all based in London; therefore, the computers used for the searches were connected to a virtual private network (VPN) service for remote access to local/national internet servers. This ensured the retrieval of results were equivalent to those obtained by individuals based in each country.

Once connected to the VPN service, women were instructed to open their local version of Google (eg, for Spain, www.google.es) and enter the keywords in the “Search” field. The search involved combining all the variations of search term 1 with each of the six variations of search term 2 (eg, “the pill OR combined pill OR contraceptive OR contraceptive pill OR birth control pill AND blood clot”), resulting in a total of six searches (see Figure 1). To avoid the identification of websites using the languages of interest, but whose origin was not in the countries under study (eg, a website written in German targeting Austria or Switzerland), women were also instructed to use the “Advanced search” menu on Google and to narrow the results to “Country/Region” (Denmark, Germany, Netherlands, Slovakia, Spain, and the United Kingdom) and “Site/Domain” (.dk, .de, .nl, .sk, .es, and .uk).

XSL•FO
RenderX

Table 1. Sociodemographic and language background of the women (N=4) who conducted the internet searches about combined hormonal contraceptives (CHCs).

Searcher	Age range	Education[a]	Employment	Previous user of oral CHC	Languages
Woman 1	25-34	Postgraduate education	Part-time	Yes	English (fluent); Spanish (fluent)
Woman 2	18-24	A-levels/GCSE	Student	No	English (native); Danish (native); German (fluent)
Woman 3	35-44	A-levels/GCSE	Part-time	Yes	English (fluent); Slovakian (native)
Woman 4	35-44	Postgraduate education	Full-time	Yes	English (fluent); Dutch (native)

[a]GCSE: General Certificate of Secondary School, equivalent to high school in the United Kindgom

Data Analysis

After the selection of websites in the first 10 pages of results, a standardized Microsoft Excel spreadsheet was created for data extraction (Multimedia Appendix 1). This spreadsheet was used by women to extract information about each website found in their respective countries. Women were instructed to complete the data extraction form in English. The data extraction form included both close-ended questions, which focused on general website characteristics, and an open-ended question for women to describe, in free text, which information was contained on the website. We summarized the women's responses to the close-ended items to gain an overview of the structure of the websites through a basic descriptive analysis (eg, counts and percentages).

However, the focus of our study was the qualitative information provided by the women when describing the websites' contents. These qualitative data were analyzed using a content analysis approach [15]. We started with a preliminary reading of the women's descriptions of the websites, from which categories of information emerged. These categories represented the content, or the topics covered, by the website. After developing this categorical system, the women's descriptions were revisited, and each website was assigned to categories according to the information they covered.

The first author conducted both the definition of the categories and the categorization of the websites. For quality control purposes, the development of the categories was discussed with a senior researcher with expertise in qualitative research. Then, we randomly selected 10% of the websites and invited one of the women who assisted with the internet search to discuss the categorizations made by the first author. For this process, we asked the woman to consider the categories assigned to each website and consider whether she agreed with the categorization. After discussion, the categorizations were readjusted, and the first author proceeded with the categorization of the remaining websites.

Results

Overview of the Website Characteristics

In total, 357 websites with information about oral CHCs were selected across the six countries. The majority of these were provided by either medical services from nongovernmental organizations (NGOs) (140/357, 39.2%) or media (110/357, 30.8%), including newspapers and online magazines. See Table 2 for detailed information on the types of websites with information about CHC-related VTE and other risks organized according to the country of origin. Websites tended to mostly target women (237/357, 66.3%), with a few focusing on particular subpopulations such as young women or students (4/357, 1.1%). One-fifth of the websites (81/357, 22.6%) did not refer to a specific group and, therefore, were considered to be targeting the general population. Half of the websites selected (185/357, 51.8%) were updated between 2015 and 2017. However, one-fifth of the websites (89/357, 24.9%) did not make any reference to when their website had last been updated.

Websites in Denmark, the Netherlands, and Slovakia did not make reference to the EMA guidelines, with mentions in the other countries minimal, namely Germany (3/113, 2.7%), Spain (1/22, 4.5%), and the UK (7/80, 8.8%). Few websites were interactive, for example with the provision for users to post questions and discuss topics with other individuals. The exceptions were websites in Spain (12/22, 54.5%) and the UK (34/80, 42.5%).

Contents Covered by Websites Focusing on Combined Hormonal Contraceptives and Their Risks

Twenty themes emerged from the content analysis of the free-text website descriptions provided by the women across the six countries. A description of these themes and their respective theme codes used to categorize the information covered by each website are provided in Textbox 2.

Having developed this schema of 20 themes, we categorized each of the 357 websites and explored the type of information that was covered by those websites. The results of this process are presented in Table 3.

XSL•FO
RenderX

Table 2. Distribution of websites per country according to their type (N=357).

Type of website	Country, n (%)						
	All countries (N=357)	Denmark (n=82)	Germany (n=113)	Netherlands (n=38)	Slovakia (n=22)	Spain (n=22)	UK (n=80)
Beauty/fitness/lifestyle website	25 (7)	6 (7)	15 (13)	0 (0)	1 (5)	2 (9)	1 (1)
Charity	6 (2)	0 (0)	1 (1)	3 (8)	0 (0)	0 (0)	2 (3)
Health blog/network	26 (7)	4 (5)	5 (4)	2 (5)	3 (14)	4 (18)	8 (10)
Insurance services	1 (0)	0 (0)	0 (0)	1 (3)	0 (0)	0 (0)	0 (0)
Legal services	2 (1)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	1 (1)
Media	110 (31)	16 (20)	42 (37)	1 (3)	0 (0)	9 (41)	42 (53)
NGO medical services	140 (39)	54 (66)	46 (41)	7 (18)	18 (82)	2 (9)	13 (16)
National health service	5 (1)	0 (0)	0 (0)	1 (3)	0 (0)	1 (5)	3 (4)
Personal blog	4 (1)	2 (2)	0 (0)	0 (0)	0 (0)	1 (5)	1 (1)
Petition website	1 (0)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Religious blog	1 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)
Science blog	2 (1)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	1 (1)
Support group	33 (9)	0 (0)	1 (1)	23 (61)	0 (0)	3 (14)	6 (8)
Travel blog	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)

Textbox 2. Themes (categories) covered by websites on combined hormonal contraceptives, including their codes and definitions/explanations.

ADP: Advantages/benefits about the pill

AP: Alternative contraceptive methods to using the pill

BC: Blood clots as risk

BWER: Bad/negative women experiences or reports about using the pill

DC: Doctor communication (ie, the importance of talking with doctor about the pill)

DP: Disadvantages about the pill

EA: Emergency advice (eg, warning symptoms, life-saving actions)

EB: Emphasis on beauty/lifestyle motives for using the pill (eg, improving skin, hair)

FC: Family communication (ie, the importance of talking to family members, in particular mothers, about the pill)

GIP: General information about the pill (eg, history, how it works, how to take it)

IC: Information about contraceptive methods in general

MP: Myths about the pill

NBC: No blood clots mentioned as a risk

RD: Resources for doctors (eg, symptom checklists, procedures to adopt)

RV: Risk variations when using the pill (eg, groups with higher risk, risk factors)

SPD: Suggests/advises pill is dangerous to use

SPS: Suggests/advises pill is safe to use

VSE: Various side effects of the pill (besides blood clots)

WP: Which is the most dangerous/safe type of pill

WW: Women worries about the pill

XSL•FO
RenderX

Table 3. Themes covered by the selected websites across the six countries (ranked by frequency for all countries).

Themes[a]	Country, n (%)						
	All countries (N=357)	Denmark (n=82)	Germany (n=113)	Netherlands (n=38)	Slovakia (n=22)	Spain (n=22)	UK (n=80)
BC	343 (96)	78 (95)	113 (100)	37 (97)	19 (86)	22 (100)	80 (100)
VSE	164 (46)	37 (45)	112 (99)	20 (53)	6 (27)	13 (59)	77 (96)
GIP	151 (42)	53 (65)	33 (29)	20 (53)	12 (55)	8 (36)	27 (34)
RV	121 (34)	41 (50)	109 (96)	4 (11)	5 (23)	2 (9)	16 (20)
WP	92 (26)	41 (50)	40 (35)	2 (5)	1 (5)	0 (0)	8 (10)
ADP	80 (22)	26 (32)	24 (21)	16 (42)	0 (0)	2 (9)	12 (15)
EA	65 (18)	26 (32)	23 (20)	5 (13)	2 (9)	0 (0)	9 (11)
BWER	62 (17)	9 (11)	17 (15)	10 (26)	0 (0)	0 (0)	70 (88)
DP	54 (15)	11 (13)	23 (20)	12 (32)	1 (5)	1 (5)	6 (8)
AP	41 (11)	13 (16)	22 (19)	0 (0)	1 (5)	0 (0)	5 (6)
SPS	35 (10)	18 (22)	14 (12)	1 (3)	0 (0)	0 (0)	2 (3)
IC	26 (7)	11 (13)	2 (2)	1 (3)	0 (0)	9 (41)	3 (4)
DC	23 (6)	5 (6)	13 (12)	0 (0)	0 (0)	0 (0)	5 (6)
SPD	21 (6)	8 (10)	9 (8)	0 (0)	1 (5)	1 (5)	2 (3)
EB	18 (5)	0 (0)	17 (15)	0 (0)	0 (0)	0 (0)	1 (1)
NBC	10 (3)	3 (4)	4 (4)	1 (3)	3 (14)	0 (0)	0 (0)
MP	9 (3)	1 (1)	1 (1)	2 (0)	0 (0)	1 (5)	4 (5)
RD	3 (1)	0 (0)	3 (3)	0 (0)	0 (0)	0 (0)	0 (0)
WW	2 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (3)
FC	1 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

[a]ADP: Advantages/benefits about the pill; AP: alternative contraceptive methods to using the pill; BC: blood clots as risk; BWER: bad/negative women experiences or reports about using the pill; DC: doctor communication (ie, the importance of talking with doctor about the pill); DP: disadvantages about the pill; EA: emergency advice (eg, warning symptoms, life-saving actions); EB: emphasis on beauty/lifestyle motives for using the pill (eg, improving skin, hair); FC: family communication (ie, the importance of talking to family members, in particular mothers, about the pill); GIP: general information about the pill (eg, history, how it works, how to take it); IC: information about contraceptive methods in general; MP: myths about the pill; NBC: no blood clots mentioned as a risk; RD: resources for doctors (eg, symptom checklists, procedures to adopt); RV: risk variations when using the pill (eg, groups with higher risk, risk factors); SPD: suggests/advises pill is dangerous to use; SPS: suggests/advises pill is safe to use; VSE: various side effects of the pill (besides blood clots); WP: which is the most dangerous/safe type of pill; WW: women worries about the pill.

Nearly all selected websites mentioned VTE as a risk of CHCs (343/357, 96.1%). Just under half of the websites (164/357, 45.9%) mentioned various side effects of CHCs and general information about CHCs (151/357, 42.2%). One-third of websites mentioned risk variations in taking CHCs according to personal characteristics/clinical history (121/357, 33.9%). A quarter of the websites provided suggestions to which oral contraceptive was the best to use (92/357, 25.8%) and the advantages of CHCs (80/357, 22.4%). Only a minority of websites emphasized the importance of talking to health professionals about CHCs (23/357, 6.4%).

Discussion

To our knowledge, this is the first and largest study to date to qualitatively explore the information available to women searching online for information about CHCs and their risks.

This study aimed to explore online information sources which mention CHCs and their risks, which are currently and publicly available to women across six EU countries. A content analysis approach was used to explore what was covered by these websites. The websites were identified through a systematic, standardized Google search in six languages. We found many websites that provided information about oral CHCs and their risks. There were a limited number of personal blogs providing women's accounts and experiences with CHCs. Overall, the websites did not offer the possibility for users to ask further questions and share their experiences and rarely included references to health agencies such as the EMA. Most websites included information on VTE and other risks of oral CHCs, but only a minority advised users to seek further information from a health professional.

The fact that information about CHCs is largely provided by media websites and NGO medical services raises questions about the quality of the information provided. As in any other medium, the quality of the information provided across websites is likely to vary, and regulation of quality standards is hard to implement [16]. Rowe et al [17] stated that, in the UK and

XSL•FO
RenderX

Sweden, media reports tend to present health risks in an "alarmist rather than reassuring" way, and seldom provide details about statistics to "express degrees of risk." Almost a decade later, Walsh-Childers et al [18] reported an improvement over the years in the completeness and usefulness of news stories about health, but also a decline in the number of articles reporting health news.

When it comes to information provided by medical services, it is possible that the contents are not always easy to understand from a general population perspective. The concept of patient health literacy has been explored in many other contexts. For instance, in a study by Dickens et al [19], it was suggested that health professionals may overestimate patients' health literacy. This conclusion was because of 68% of patients whose health literacy was judged as adequate by health professionals, only 22% were considered to be health literate according to the Newest Vital Sign test, which aims to measure low health literacy [19]. In the field of contraception, a study conducted in 2007 reported that the reading level for contraceptive instructions was the sixth to twelfth grade for condoms, ninth to tenth grade for spermicides, and tenth to twelfth grade for emergency contraception, with the twelfth grade corresponding to education attained at 18 years of age. However, according to the latest Eurostat 2017 statistics [20], there are numerous individuals in our countries of interest with "less than primary, primary, and lower secondary education." Lower education corresponds to education attained at 13 years of age (Denmark 17%, Germany 13%, Netherlands 19%, Slovakia 7%, Spain 37%, and UK 19%). The scenario becomes even more complex when it comes to using health information which was searched online. According to authors such as Norman and Skinner [21], consumers of eHealth information need to have additional literacy skills to explore the data available on websites, which include not only being able to understand the information itself, but also the ability to seek, find, and appraise the quality of the information found. Hence, further research is needed to explore whether both the literacy and eHealth literacy of those searching for online information about CHC and its risks to support contraception choices.

This study also suggests that websites about contraceptive are generally aimed at women. Literature has also referred to contraception as a topic that is socially perceived to be within the "women's sphere" [14]. If information provided in websites is not gender neutral, this may be perpetuating the idea that contraception is a decision placed on women as opposed to partnerships. Future research should explore the perceptions of males about CHCs, their level of involvement in contraception choices, and which information sources do men seek when it comes to contraception.

Our study demonstrated that most websites were not interactive (ie, did not allow its users to make comments, share experiences, or ask questions about the content provided). Previous literature has demonstrated that the internet, including social media platforms, is becoming an important platform for communication between patients and health professionals. According to Rolls et al [22], health professionals are increasingly using social media to share their knowledge. As a result, several guidelines have emerged about how professionals should use online communication tools to share health information [23,24]. The internet is a powerful source of health information, including topics such as contraception; therefore, professionals may be missing an opportunity to communicate with women about oral CHCs using the internet. We acknowledge that having websites monitored by health regulators or health professionals with active communication features available to their users is unlikely to be feasible. However, we propose that popular health-related websites may develop algorithms which allow them to understand (eg, the information most frequently accessed by website users) and use that as starting point to improve the way in which such information is shared with the target audience.

Despite giving information about the risks of oral CHCs, there are several websites which also advise about the best oral CHC, including references to different brands. This topic has been widely addressed, with many authors referring to "Dr Google" as an expression that reflects the role that the internet has assumed in searching for health information. In the study by Lee et al [25], it was reported that nearly half of individuals with chronic health conditions use Web-based health information to help manage their condition. Individuals seek health information online due to the convenience, coverage, and anonymity [26]. For this reason, the health community must acknowledge the role the internet is playing in helping people making decisions about their health care, and increase involvement with the internet as a prime medium for communication of the latest evidence.

Citation of either European (eg, EMA) or local (eg, National Health Service) health agencies was limited. Variation in reference to health agencies presents the possibility of further involvement in promoting initiatives to raise awareness about the extent of their work among the general population. A possible example for modeling is the project "Decides Espana" [27], which launched in Spain in 2014. This project focused on improving sexual and reproductive rights, and one of its goals was to "raise the awareness of health institutions on the notion of equity in health." Although this project targeted cultural minorities, it provides an example of initiatives that could be promoted at a local and European level to increase the population knowledge about health organizations that individuals can contact or access to search for reliable and up-to-date information.

This study is not without limitations. First, the search was conducted in different languages, derived from a protocol prepared in English; therefore, it is uncertain whether all key terms used were completely equivalent across countries. We consider the impact of this to be minimal because the women conducting the searches were bilingual or fluent in English. Second, this study only discusses information about CHCs included on the websites and was unable to assess the quality of this information. Third, the six countries included did represent some diversity across Europe, with four of the six being part of Northern Europe, one from Eastern Europe, and one from Southern Europe. However, further research could target other Southern and Eastern European countries. Nevertheless, we consider this study to be a pioneer in its approach of systematically searching for and reviewing health information which is readily and freely available to all internet

XSL•FO
RenderX

users. We hope this might encourage other researchers to focus on evaluating the quality of health-related information provided online. This is particularly important when it comes to using oral CHCs. CHCs have been marketed as a "lifestyle drug" (ie, "medications to improve a person's quality of life"), which may also have a "cosmetic, life-enhancing, recreational, or discretionary" value [28], yet carries a small but important potential risk to health through VTE, which may not be evident among the normalization of use of CHCs.

In conclusion, this study has shown that there are numerous online information sources about oral CHCs. It is crucial that regulatory health agencies and health professionals engage in this space to inform potential and current users of CHCs both inside and outside the consultation room. In a digital era, and when there is little control over the contents that internet users post online (eg, personal blogs), such active involvement by health authorities would allow for CHC users to be able to make fully informed decisions about their contraceptive needs.

Acknowledgments

This study was funded by the European Medicines Agency (EMA/2015/24/PH). The study team is a part of an international consortium which conducts studies on effectiveness and pharmacoepidemiology studies on medicines led by Department of Clinical Epidemiology, Aarhus University. The views expressed are those of the authors and not necessarily those of the European Medicines Agency. The authors would like to particularly express their gratitude to Dr Vera Ehrenstein (Aarhus University) for her collaboration, as well as to the women who assisted with the internet search.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Data Extraction Form (Table format).

[PDF File (Adobe PDF File), 22KB - jmir_v21i3e10810_app1.pdf]

References

1. Vinogradova Y, Coupland C, Hippisley-Cox J. Use of combined oral contraceptives and risk of venous thromboembolism: nested case-control studies using the QResearch and CPRD databases. BMJ 2015 May 26;350:h2135 [FREE Full text] [Medline: 26013557]
2. Moheimani F, Jackson D. Venous thromboembolism: classification, risk factors, diagnosis, and management. ISRN Hematol 2011:10. [doi: 10.5402/2011/124610] [Medline: PMC3196154]
3. European Medicines Agency. 2013 Oct 11. PRAC confirms that benefits of all combined hormonal contraceptives (CHCs) continue to outweigh risks [press release] URL: https://tinyurl.com/y5us96r8 [accessed 2019-01-28] [WebCite Cache ID 75lGavxTt]
4. European Medicines Agency. 2014 Jan 31. Combined hormonal contraceptives URL: https://www.ema.europa.eu/en/medicines/human/referrals/combined-hormonal-contraceptives [accessed 2019-01-28] [WebCite Cache ID 75lGiLRVV]
5. European Commission. 2014 Nov 28. Europeans becoming enthusiastic users of online health information URL: https://tinyurl.com/y3zddt5w [accessed 2019-01-28] [WebCite Cache ID 75lH2zAPy]
6. Eurostats. Eurostats Statistics Explained. 2017. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Main_Page [accessed 2019-01-28] [WebCite Cache ID 75lHGFomZ]
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. de Reilhac P, Plu-Bureau G, Serfaty D, Letombe B, Gondry J, Christin-Maitre S. The CORALIE study: improving patient education to help new users better understand their oral contraceptive. Eur J Contracept Reprod Health Care 2016 Oct;21(5):388-394. [doi: 10.1080/13625187.2016.1217323] [Medline: 27530618]
9. Costa AR, Palma F, Sá JL, Vicente L, Bombas T, Nogueira AM, et al. Impact of a women's counselling programme on combined hormonal contraception in Portugal--the IMAGINE Study. Eur J Contracept Reprod Health Care 2011 Dec;16(6):409-417. [doi: 10.3109/13625187.2011.608441] [Medline: 21942578]
10. Merckx M, Donders GG, Grandjean P, Van de Sande T, Weyers S. Does structured counselling influence combined hormonal contraceptive choice? Eur J Contracept Reprod Health Care 2011 Dec;16(6):418-429 [FREE Full text] [doi: 10.3109/13625187.2011.625882] [Medline: 22066890]
11. European Networks of Centres for Pharmacoepidemiology Pharmacovigilance. 2018. Study of regulatory communication and risk awareness following the Article 31 referral of Combined Hormonal Contraceptives in relation to thromboembolism URL: https://tinyurl.com/y3xftmln [accessed 2019-01-28] [WebCite Cache ID 75lZSHcdf]

XSL•FO
RenderX

12. Baazeem M, Abenhaim H. Google and women's health-related issues: what does the search engine data reveal? Online J Public Health Inform 2014;6(2):e187. [doi: 10.5210/ojphi.v6i2.5470]
13. Kwon B, Kim S, Duket T, Catalán A, Yi J. Do people really experience information overload while reading online reviews? Int J Human Computer Inter 2015;31:959-973. [doi: 10.1080/10447318.2015.1072785]
14. Fennell J. Men bring condoms, women take pills: men's and women's roles in contraceptive decision making. Gend Soc 2011;25:496-521. [doi: 10.1177/0891243211416113]
15. Hsieh H, Shannon S. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687]
16. Purcell G, Wilson P, Delamothe T. The quality of health information on the internet. BMJ 2002;324(7337):17-18. [Medline: PMC1122494]
17. Rowe G, Frewer L, Sjoberg L. Newspaper reporting of hazards in the UK and Sweden. Public Underst Sci Jan 2000;9(1):59-78. [doi: 10.1088/0963-6625/9/1/304]
18. Walsh-Childers K, Braddock J, Rabaza C, Schwitzer G. One step forward, one step back: changes in news coverage of medical interventions. Health Commun 2018 Feb;33(2):174-187. [doi: 10.1080/10410236.2016.1250706] [Medline: 27983868]
19. Dickens C, Lambert BL, Cromwell T, Piano MR. Nurse overestimation of patients' health literacy. J Health Commun 2013;18 Suppl 1:62-69 [FREE Full text] [doi: 10.1080/10810730.2013.825670] [Medline: 24093346]
20. Eurostats. 2016. Educational attainment statistics URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/Educational_attainment_statistics [accessed 2019-01-28] [WebCite Cache ID 75lI8kqsk]
21. Norman C, Skinner H. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
22. Rolls K, Hansen M, Jackson D, Elliott D. How health care professionals use social media to create virtual communities: an integrative review. J Med Internet Res 2016 Jun 16;18(6):e166 [FREE Full text] [doi: 10.2196/jmir.5312] [Medline: 27328967]
23. Chauhan B, George R, Coffin J. Social media and you: what every physician needs to know. J Med Pract Manage 2012;28(3):206-209. [Medline: 23373164]
24. 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]
25. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer use of "Dr Google": a survey on health information-seeking behaviors and navigational needs. J Med Internet Res 2015 Dec 29;17(12):e288 [FREE Full text] [doi: 10.2196/jmir.4345] [Medline: 26715363]
26. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
27. Derechos Culturales para Impulsar el Desarrollo y la Salud. Decides. 2017. Decides Espana [press release] URL: http://www.decides.pro/wp-content/uploads/2014/12/FICHA-PA--S-DECIDES_Espa--a-EN.pdf [accessed 2019-01-28] [WebCite Cache ID 75lSDX2CR]
28. Watkins ES. How the pill became a lifestyle drug: the pharmaceutical industry and birth control in the United States since 1960. Am J Public Health 2012 Aug;102(8):1462-1472. [doi: 10.2105/AJPH.2012.300706] [Medline: 22698049]

Abbreviations

CHC: combined hormonal contraceptives
EMA: European Medicines Agency
NGO: nongovernmental organizations
VTE: venous thromboembolism

Edited by G Eysenbach; submitted 19.04.18; peer-reviewed by J Fleming, H Hah; comments to author 29.09.18; revised version received 23.11.18; accepted 10.12.18; published 18.03.19.

Please cite as:
Gomes Alves P, Petersen I, Stevenson F
Searching for Information on the Risks of Combined Hormonal Contraceptives on the Internet: A Qualitative Study Across Six European Countries
J Med Internet Res 2019;21(3):e10810
URL: http://www.jmir.org/2019/3/e10810/
doi:10.2196/10810
PMID:30882358

XSL•FO
RenderX

©Paula Gomes Alves, Irene Petersen, Fiona Stevenson. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Challenges of Implementing a Web-Based Telemonitoring Strategy for Teenagers With Inflammatory Bowel Disease: Empirical Case Study

Alie Dijkstra[1], MSc; Anke Heida[2], MD, PhD; Patrick Ferry van Rheenen[1], MD, PhD

[1]Department of Pediatric Gastroenterology, Hepatology and Nutrition, University of Groningen, University Medical Centre Groningen, Groningen, Netherlands

[2]Department of General Practice, University of Groningen, University Medical Centre Groningen, Groningen, Netherlands

Corresponding Author:
Patrick Ferry van Rheenen, MD, PhD
Department of Pediatric Gastroenterology, Hepatology and Nutrition
University of Groningen
University Medical Centre Groningen
Internal Code CA31
Hanzeplein 1
Groningen, 9713 GZ
Netherlands
Phone: 31 503614151
Fax: 31 503611671
Email: p.f.van.rheenen@umcg.nl

Abstract

Background: We designed a telemonitoring strategy for teenagers with inflammatory bowel disease to prevent an anticipated disease flare and avert unplanned office visits and day care procedures. The strategy was evaluated in a randomized controlled trial that involved 11 Dutch pediatric gastroenterology centers, each using repeated symptom scores and stool calprotectin measurements. In the telemonitoring arm of the trial, teenagers (n=84) as well as their health providers were alerted to out-of-range results, and suggestions for change in therapy were offered. We demonstrated that the technology was a safe and cost saving alternative to health checks by the specialist at fixed intervals.

Objective: The aim of this study was to evaluate whether we could move our telemonitoring strategy from a demonstration project to one that is sustained within existing sites.

Methods: In this empirical case study, we used the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework to explore the challenges to implementing our strategy. The framework distinguishes 7 domains: (1) the illness, (2) the technology, (3) the value proposition, (4) the adopter system, (5) the organization, (6) the societal system, and (7) the time dimension. We summarized the challenges across all 7 domains and classified them as simple (+++), complicated (++), or complex (+). Technologies in which multiple domains are complicated have proven difficult to implement, whereas those with multiple complex domains may not even become mainstreamed.

Results: The technology that we used and the linked program (*IBD-live*) allowed us to select and target the teenagers who were most likely to benefit from a face-to-face encounter with their specialist (+++). The value proposition of the technology was clear, with a distinct benefit for patients and an affordable service model, but health providers had plausible personal reasons to resist (double data entry, ++). The organization was not yet ready for the innovation, as it requires a shift to new ways of working (+). We had no concerns about reimbursement, as Dutch health insurers agreed that screen-to-screen consultations will be reimbursed at a rate equivalent to face-to-face consultations (+++). Finally, the technology was considered easy to adapt and evolve over time to meet the needs of its users (+++).

Conclusions: The challenges to be addressed are merely *complicated* (++) rather than *complex* (+), which means that our program may be difficult but not impossible to sustain within existing sites. After integrating the technology and its use with local workflows first, we believe that our telemonitoring strategy will be ready for sustained adoption. In contrast with what we did ourselves, we recommend others to use the NASSS framework prospectively and in real time to predict and explore the challenges to implementing new technologies.

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e11761) doi:10.2196/11761

KEYWORDS

eHealth; inflammatory bowel disease; health care improvement; implementation science; quality of care

Introduction

Crohn's disease and ulcerative colitis are the 2 main forms of inflammatory bowel disease (IBD). Both are chronic lifelong conditions that cause intestinal inflammation. The inflammation waxes and wanes over time in a seemingly unpredictable fashion. The prevalence of IBD exceeds 0.3% in North America, Oceania, and many countries in Europe [1]. A third of patients experience at least one flare per year, which is associated with pain, disruption of normal activities, and reduction of quality of life [2].

Follow-up of patients with IBD traditionally comprises scheduled visits regardless of how well the patients feel. For every scheduled visit that is instrumental in preventing unnecessary pain or disability, many more will be uninformative. Selecting and targeting the patients who are most likely to benefit from a face-to-face encounter with their specialist are logical solutions.

We have recently shown that repeated quantitative measurements of fecal calprotectin help specialists and patients recognize disease flares at an early stage [3]. Monitoring calprotectin values longitudinally enables timely adjustment of therapy plans. In modern-day IBD care, the treatment target is to keep calprotectin values below 250 µg/g, which is the equivalent of disease remission [4]. When calprotectin values exceed this critical threshold, the patient has a high risk of disease flare and treatment intensification is justified. Conversely, when calprotectin values drop again below the critical threshold, induction therapy can be safely tapered down.

Although it seems that without monitoring, patients with IBD would not be in a position to receive optimal care, there are several potential difficulties and unresolved matters that relate to the use of calprotectin. It may disrupt doctor-patient communication, unduly focusing attention on inflammatory markers (medicalization and creating worried well), increase health care costs, and bears the risk of overdiagnosing flares as calprotectin may also rise in bacterial gastrointestinal infections.

Between June 2013 and January 2016, we conducted a randomized controlled trial (RCT) comparing a Web-based telemonitoring strategy for teenagers with IBD (*IBD-live*, designed for early detection of relapse) and conventional follow-up [5]. We concluded that patients in the telemonitoring group followed a similar disease course as those in conventional follow-up, although they had significantly less face-to-face encounters with their specialist. Moreover, the number of times blood was taken, trips to the hospital, and hours of absence from school or social activities were lower in the telemonitoring group. The majority of patients assigned to the telemonitoring arm took an active interest in their readings, engaged enthusiastically with the feedback, and found this technical innovation reassuring. In a patient satisfaction survey that was performed after the close-out visit (after 52 weeks of participation), the majority of respondents agreed that telemonitoring was time saving, increased their understanding of the disease, and did not disturb them. Overall, 81% of respondents wished to use the telemonitoring services in future.

We sought to evaluate whether we could move our telemonitoring strategy from a demonstration project to one that is sustained within existing sites.

Methods

Context

The Web-based telemonitoring strategy was tested in 11 pediatric gastroenterology practices, which were recruited on the basis of their interest in IBD. The principal investigators at the various sites were pediatric gastroenterologist and members of the *Kids with Crohn's and Colitis* working group for Collaborative Research in the Netherlands. Together they treat about two-thirds of the total pediatric IBD population in the Netherlands. National treatment guidelines provided uniformity in treatment among practices [6]. Table 1 summarizes practice characteristics for the 11 sites.

Outline of the Telemonitoring Strategy

Patients in the telemonitoring group had access to the IBD-live website, a cloud-based portal with 3 modules, as follows: (1) the Flarometer, an automatic cumulation of patient-rated symptom scores and fecal calprotectin measurements, (2) a platform for direct communication with the local IBD team with guaranteed feedback within 2 working days, and (3) a module with study questionnaires (on quality of life, absenteeism, and health care utilization). Patients could freely explore the website, but the majority only accessed it after receiving an email notification that it was time for the next Flarometer measurement (Figure 1). In case the Flarometer was not used despite 3 email notifications with 7-day intervals, the IBD-nurse or specialist phoned the patients to encourage them to fill in the Flarometer and send in a stool sample. The email notification had a link that took the user in a few steps in sequential order to the Flarometer (tunnel design) [7]. The Flarometer strategy is presented in Figure 2. Patients as well as their specialists were alerted to out-of-range results, and suggestions for change in therapy were offered.

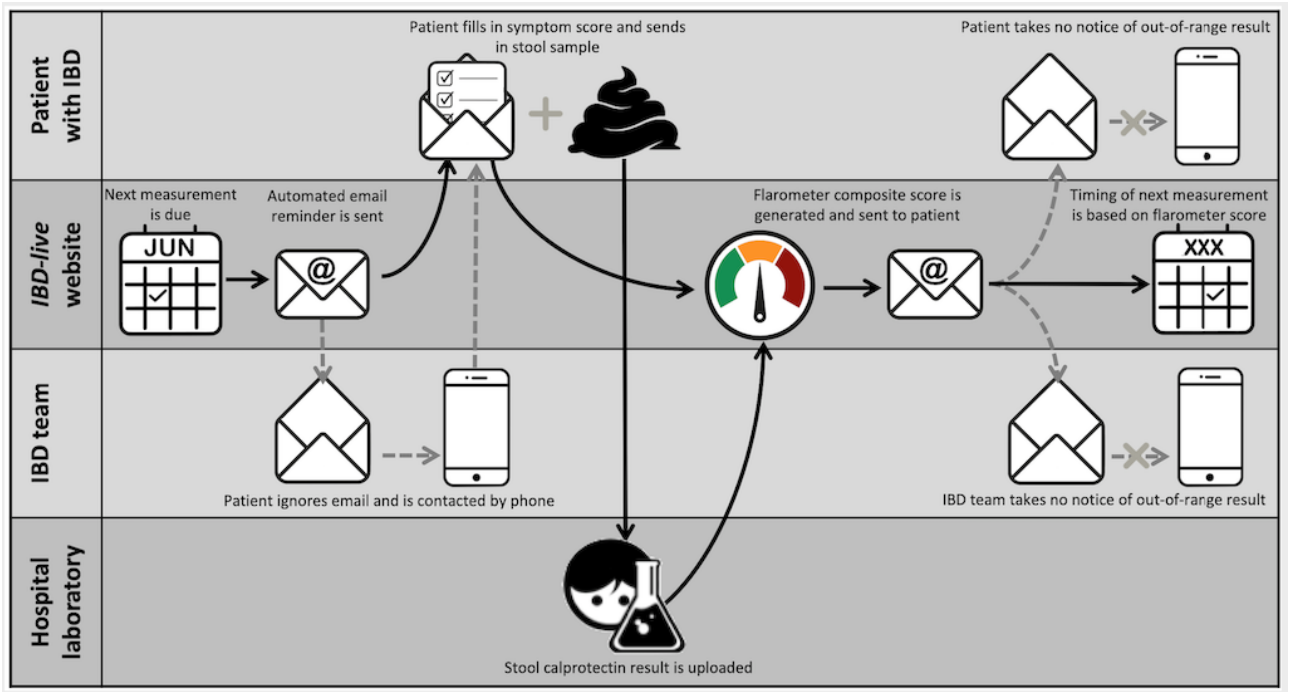
XSL•FO
RenderX

Table 1. Practice characteristics (N=11).

Characteristics	Number of practices, n
Practice type	
Tertiary care hospital	6
Large regional general hospital	5
Practice size	
10 to 100 pediatric IBD[a] patients	7
More than 100 pediatric IBD patients	4
Adolescents transitioned to adult-oriented care	
More than 10 per year	7
≥10 per year	4
Pediatric IBD nurse present	
Yes	8
No	3
Pediatric endoscopy unit available	
Yes	11
No	0

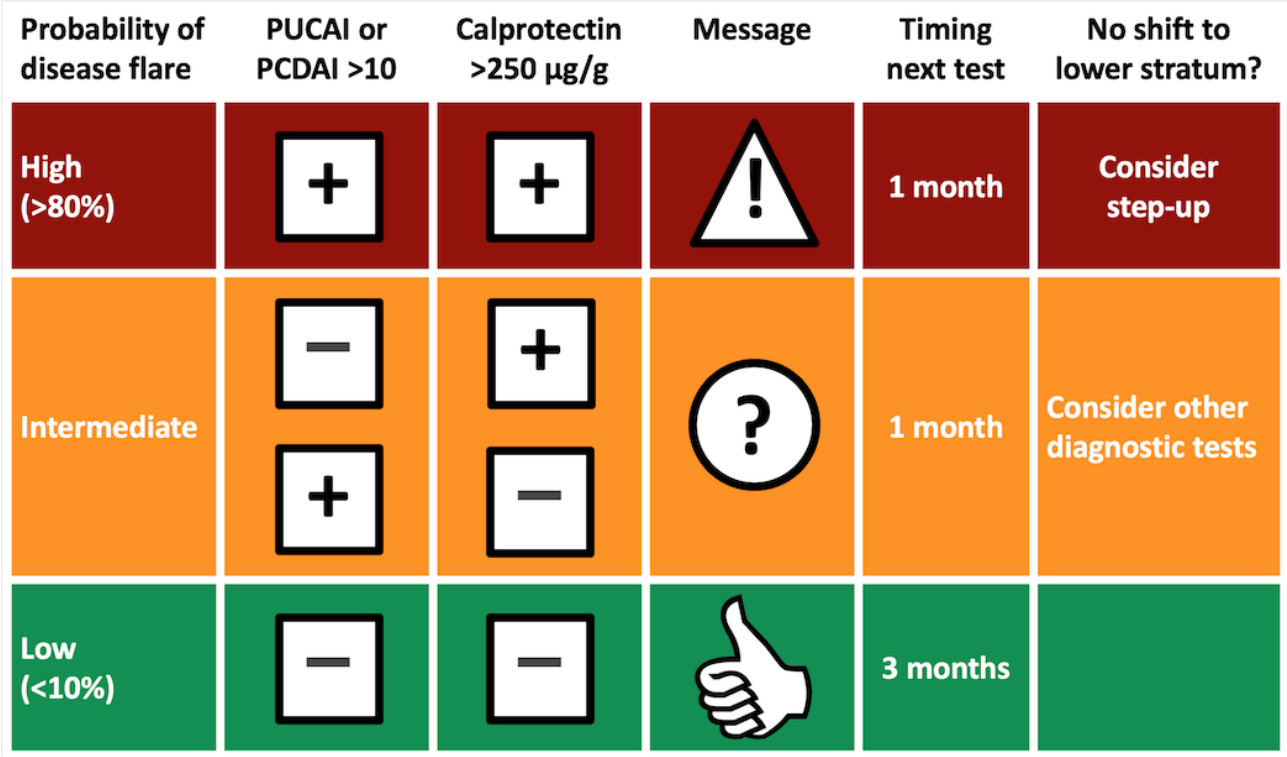
[a]IBD: Inflammatory bowel disease.

Figure 1. Swim lane process diagram showing the sequence of events from sending an automated reminder by email to the advice on treatment and timing of re-measurement. The parallel lines divide the diagram into lanes, with one lane for each person or subprocess. The horizontal direction represents the sequence of events in the overall process. Arrows between the lanes represent how information or material is passed among the subprocesses. Black arrows represent the ideal flow; gray dotted arrows represent the weakest links in the chain of events. IBD: inflammatory bowel disease.



XSL•FO
RenderX

Figure 2. Flarometer strategy. Algorithm with advice on treatment and the timing of re-measurement (figure printed with permission from Heida et al [5]). PUCAI: pediatric ulcerative colitis activity index; PCDAI: pediatric Crohn's disease activity index.



Evaluation

During the RCT, we gathered information on its potential for implementation in a real-world setting. Data sources included study parameters as specified in the study protocol [8] and ethics paperwork, the log of technical failures, and 1 additional semistructured interview for all patient participants. In addition, for the purpose of this case study, we collected information from 5 research staff including the principal investigator (PFvR) and 2 research coordinators (AD and AH), 13 clinical staff involved in the trial, 4 clinical staff not involved in the trial, 2 hospital decision makers, and 1 Web designer. This information was extracted from 15 semistructured and 4 narrative interviews.

We used the recently published nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework to systematically explore the challenges and decide whether we should abandon our telehealth technology or aim to achieve sustained adoption within the existing sites [9].

The NASSS framework comprises measures categorized in 7 domains: (1) the illness, (2) the technology, (3) the value proposition, (4) the adopter system (professional staff, patients, and lay carers), (5) the organization, (6) the wider (institutional and societal) system, and (7) interaction and mutual adaptation among all these domains over time. In Table 2, we summarize all the original measures and how we translated them to assess local applicability. All measures were classified as simple (straightforward, predictable, and few components—as in making a sandwich), complicated (remains predictable, but with multiple interacting components or issues—as in building a rocket), or complex (dynamic, unpredictable, and not easily disaggregated into constituent components—as in raising a child). The choice for 1 of 3 classes was further substantiated by tabulated examples in the NASSS index paper [9].

Technologies in which multiple domains are complicated have proven difficult, slow, and expensive to implement. Those in which multiple domains are complex may not even become mainstreamed.

Overall, 2 assessors (AD, PFvR) independently rated the measures with the use of a predefined matrix. Disagreements were resolved by discussion and, if necessary, a third independent assessor (AH).

Ethical Considerations

The patient-related data in this paper as well as the log of technical failures were collected as part of the initial RCT. The Medical Ethical Committee of the University Medical Centre Groningen approved the study protocol (METC 2013/010). Secondary approval was obtained from all participating centers. Use of additional data collected from nonpatient participants (including semistructured interviews with specialists and IBD nurses and qualitative interviews with managers and Web designer) in this empirical case study did not require approval according to the Dutch Medical Research Involving Human Subject Act.

XSL•FO
RenderX

Table 2. Overview of the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework used to evaluate the success of the *IBD-live* telemonitoring strategy. The left column shows the original NASSS questions, the right column shows how we adapted some to assess local applicability.

Domain and questions	Adapted questions
Domain 1: The illness	
1A. What is the nature of the condition or illness?	How well does the Flarometer strategy predict out-of-range or all-is-well events?
1B. What are the relevant sociocultural factors and comorbidities?	What are the sociocultural factors associated with good adherence to the technology? (Original RCT[a])[b]
Domain 2: The technology	
2A. What are the key features of the technology?	No changes
2B. What kind of knowledge does the technology bring into play?	No changes
2C. What knowledge and/or support is required to use the technology?	No changes
2D. What is the technology supply model?	No changes
Domain 3: The value proposition	
3A. What is the developer's business case for the technology (supply-side value)?	No changes
3B. What is the desirability, efficacy, safety, and cost effectiveness of the technology (demand-side value)?	What is the change from baseline in quality-of-life and cost-effectiveness? (Original RCT)[b]
Domain 4: The adopter system	
4A. What changes in staff roles, practices, and identities are implied?	No changes
4B. What is expected of the patient (and/or immediate caregiver)—is this achievable by, and acceptable to, them?	No changes
4C. What is assumed about the extended network of lay caregivers?	Not applicable
Domain 5: The organization	
5A. What is the organization's capacity to innovate?	No changes
5B. How ready is the organization for this technology-supported change?	No changes
5C. How easy will the adoption and funding decision be?	No changes
5D. What changes will be needed in team interactions and routines?	No changes
5E. What work is involved in implementation and who will do it?	No changes
Domain 6: The wider context	
6A. What is the political, economic, regulatory, professional (eg, medicolegal), and sociocultural context for program rollout?	No changes
Domain 7: The time dimension	
7A. How much scope is there for adapting and coevolving the technology and the service over time?	No changes
7B. How resilient is the organization to handling critical events and adapting to unforeseen eventualities?	No changes

[a]RCT: randomized controlled trial.

[b]We already published the results of this measure in the original RCT [5].

Results

A summary of the challenges of the IBD-live telemonitoring strategy per domain is given in Table 3.

In the following sections, we give a detailed overview of the considerations we have made during the rating process.

Domain 1: The Illness

Patients were followed for 52 weeks. During this observation period, 448 Flarometer scores were generated, of which 176 (39.3%) were in the green stratum, 181 (40.4%) were in the orange stratum, and 91 (20.3%) were in the red stratum. Overall, 30 of 176 green *all-is-well* results were generated during a close-out visit and were excluded from analysis. The median change in fecal calprotectin concentration in the follow-up stool

XSL•FO
RenderX

of the 146 remaining cases was 0 (interquartile range, IQR 0-70) ug/g (Multimedia Appendix 1). A total of 11 of 91 red *out-of-range* alerts were generated during a close-out visit and were excluded from analysis for not leading to a response from patient or health provider anymore. Among the remaining 80 red alerts were the patients who were most likely to benefit from an encounter with their specialist because of an imminent flare. In 39 of 80 cases, the local IBD team followed the automated advice to intensify treatment. The median change in fecal calprotectin concentration in the follow-up stool in this group was −255 (IQR −975 to 20) ug/g. In 21 of 80 cases, there was a well-founded reason to deviate from the treatment advice and follow a watchful waiting strategy. These reasons included the presence of colon pathogens and high symptom scores caused by disorders that were not related to IBD such as irritable bowel syndrome. In these cases, the median change in fecal calprotectin concentration in the follow-up stool was −216 (IQR −1965 to 0) ug/g. In 20 of 80 red alerts, there was no contact between patient and health provider, despite 3 email notifications and an attempt to contact the patient by telephone after 2 weeks of passivity. This strategy violation led to a median change in fecal calprotectin concentration of 167 (IQR 0-1047) ug/g. Multimedia Appendix 1 shows that the direction of change in calprotectin concentration in the latter group was different from the others and toward further deterioration.

The foregoing made us decide that ad hoc face-to-face follow-up consultations were fairly predictable and consistent in the group with out-of-range results. We had few unpredictable eventualities in the all-is-well results. Both disease relapses and sustained remissions were predictable, and the treatment advice generated by the Flarometer was reliable. This measure was therefore classified as *simple*.

Domain 2: The Technology

The telemonitoring strategy relied on calprotectin measurements in stool samples that were sent to the hospital laboratory by ordinary mail. Immediately after arrival, the stool sample was analyzed with a point-of-care test that offers quantitative results within minutes. This rapid test has good agreement with the established enzyme-linked immunosorbent assay (ELISA) and can therefore serve as a reliable alternative to the time-consuming ELISA in the follow-up of patients with IBD [10]. We plotted the calprotectin results per patient over time to give both patient and doctor an overview of the inflammatory state of the gut during the experiment. Patients found the longitudinal tracking of calprotectin results helpful and trusted the results. However, doctors were equivocal about whether a rise in calprotectin above 250 ug/g was sufficient to justify treatment intensification. Some wished to perform additional diagnostic tests to rule out gastrointestinal infections before progressing to a treatment decision. Multimedia Appendix 2 shows that concurrent gastrointestinal infections were found in 14% of stools with out-of-range results, and this proportion was not statistically different from stools with calprotectin values in the target range (below 250 ug/g).

Use of the telemonitoring strategy required no previous knowledge from the patients except for recognition of what conditions count as urgent, such as a drug reaction. All measures within this domain were classified as *simple*, with the exception of 2D: the sequence of events from sending an automated reminder by email to the formulation of a treatment advice (Figure 1), relying on a bespoke solution from a small-sized enterprise, which carries a risk of supplier withdrawal. However, its relatively substitutability is relatively straightforward, and the complexity of this measure was therefore classified as *complicated*.

Domain 3: The Value Proposition

Our telemonitoring strategy depended on a predictive algorithm on the basis of calprotectin changes over time to prevent an anticipated disease flare and avert unplanned outpatient visits and day care procedures (including colonoscopies). Although the technology created new possibilities for *personalized* medicine, such as selecting and targeting the patients who are most likely to benefit from a face-to-face encounter with their specialist, we did not observe a reduction in workload. Patients assigned to the telemonitoring group had fewer face-to-face encounters with the IBD team compared with those in conventional follow-up, but the number of telecontacts (emails and telephone calls) increased by 40%. The telemonitoring strategy was cost effective, as it demonstrated comparable effectiveness with conventional follow-up at lower costs. In the intention-to-treat analysis, telemonitoring led to a mean annual cost saving of €89 per patient, and this amount increased to €360 in those who were adherent to the technology. The time-saving aspect of the technology and the better sense of disease control were highly valued by patients and their caretakers. Not all patients were suitable to be followed in a telemonitoring program. It is important to realize that differences exist in people's desire for information. Some patients seek as much information as possible about the threat of a disease flare (*monitors*), whereas others try to avoid potentially threatening information (*blunters*) [11]. The Flarometer technology will give the former category of patients and parents a greater sense of control, as opposed to the latter category who may feel more vulnerable because of the constant confrontation with their chronic disease. This aspect may partially explain attrition in the telemonitoring arm.

We are not sure that use of the telemonitoring strategy reduces the demand on health services, but it certainly allows selecting and targeting the patients who are most likely to benefit from a face-to-face encounter with their specialist. Furthermore, the business case is still underdeveloped. We therefore classified measure 3A as *complicated*. However, the technology itself is desirable for patients, and it is safe and cost effective, particularly in those who are adherent to the telemonitoring strategy, and measure 3B was therefore classified as *simple*.

XSL•FO
RenderX

Table 3. Summary of the challenges of the *IBD-live* telemonitoring strategy based on the nonadoption, abandonment, scale-up, spread, and sustainability framework.

Domain and question	Rating[a]	Explanation
Domain 1: The illness		
1A. How well does the Flarometer strategy predict out-of-range or all-is-well events?	+++	Ad hoc face-to-face follow-up consultations were fairly predictable and consistent in the group with out-of-range results; few unpredictable eventualities in the all-is-well results.
1B. What are the sociocultural factors associated with good adherence to the technology?	+++	Majority of teenage patients were considered *suitable* for the technology, except for those with concurrent diseases such as arthritis and sclerosing cholangitis; Patients with delayed emotional maturity, prominent functional gastrointestinal disease, or with little knowledge of the Dutch language were considered *less suitable* for the technology.
Domain 2: The technology		
2A. Perceived usability	+++	Access to Web-based portal was easy via a link in the email notification. The tunnel design (Figure 1) guided the patients in a few steps in sequential order through the Flarometer. Successful stool-based monitoring system for teenagers heavily relies on active and coordinating role of the parent.
2B. Appropriateness of the automated treatment advice	+++	Treatment advice was accepted and trusted by patients. Concurrent gastrointestinal infection was found to be the cause of the out-of-range result in a minority of cases.
2C. Knowledge and/or support required to use the technology	+++	Use of the technology requires no previous knowledge from the patient, except for recognizing what conditions count as urgent.
2D. Technology supply model	++	Technology relies on bespoke solution from a small-sized enterprise with risk of supplier withdrawal.
Domain 3: The value proposition		
3A. Developer's business case for the technology (supply-side value)	++	Not sure that use of the technology reduces the demand on health services, but it certainly allows selecting and targeting the patients who are most likely to benefit from a face-to-face encounter with their specialist.
3B. Efficacy, safety, and cost-effectiveness of the technology (demand-side value)	+++	Technology is desirable for patients, it is safe and cost effective, particularly in those who are adherent to the telemonitoring strategy.
Domain 4: The adopter system		
4A. Implications for staff roles, practices, and identities in case of adoption	+++	Difference between innovators and early adopters (who embraced the technology) on 1 side and a minority of other health providers (who were reluctant to adopt the technology). Technology was not seen as a threat to job security.
4B. What is expected of the patient (and/or immediate caregiver)—and is this achievable by, and acceptable to, them?	+++	Patients were already familiar with the stool collection procedure. Logging on to the system was easy. The study presupposed that parents or carers were actively supporting their child during the study observation period.
Domain 5: The organization		
5A. Organization's capacity to innovate	+	Technology follows the natural work flow but conflicts with established hospital electronic databases and therefore requires double data entry. This will put a strain on the already overstretched health service.
5B. Readiness for this technology-supported change	++	No linked routine for booking face-to-face appointments in case of a red alert.
5C. Easiness of adoption and funding	++	Anticipated reduction in costs were not realized as case management was not always successful in avoiding follow-up consultations and day care admissions. Neutral cost-benefit balance.
5D. Changes required in team interactions and routines	++	Variation in clinician engagement was based on the vision of local teams of whether remote biomarker monitoring enhances rather than threatens the existing service. Expansion of *rapid access* possibilities is required.
5E. Work and persons involved in implementation	+	Significant work needed to build shared vision, engage staff, enact new practices, and monitor impact.
Domain 6: The wider context		

XSL•FO
RenderX

Domain and question	Rating[a]	Explanation
6A. Political, economic, regulatory, professional, and sociocultural context for program rollout	+++	Effect of January 2018, the Dutch Health care Authority has agreed that screen-to-screen consultations will be reimbursed at a rate equivalent to face-to-face consultations, provided that a substantive report is added to the patient's medical record.
Domain 7: The time dimension		
7A. Scope for adapting and coevolving the technology over time	++	The technology can easily be adapted over time.
7B. Handling critical events and adaptation to unforeseen eventualities	+++	The research head quarter and the Web designer were able to detect critical events quickly and respond to these through coordinated action.

[a]Rating: Simple +++; Complicated ++; Complex +.

Domain 4: The Adopter System

The structured questionnaire that was spread among health providers after termination of the RCT revealed the difference between innovators and early adopters (who embraced the technology) on 1 side and a minority of other health providers who were reluctant to adopt the technology on the other. The hurdle was not learning to use a new technology but rather the prospect of double data entry when running a telemonitoring service in parallel with conventional follow-up. This perception of overload will put a strain on already overstretched health services. The telemonitoring strategy was not seen as a threat to job security, although a considerable proportion of health providers felt that it changed the professional distance to their patients (Figure 3).

For patients, the easiest way to log on to our system was via a link in the email notification. Some patients forgot their username or password, and the time involved in recovering their username or resetting their account password demoralized them. The tunnel design guided the patients in a few steps in sequential order through the Flarometer, and entering data was merely clicking in a data entry form. The patients were already familiar with the stool collection process with its inherent obstacles [12].

As neither existing staff nor patients and caregivers had to learn new skills, we classified all measures within this domain as *simple*.

Domain 5: The Organization

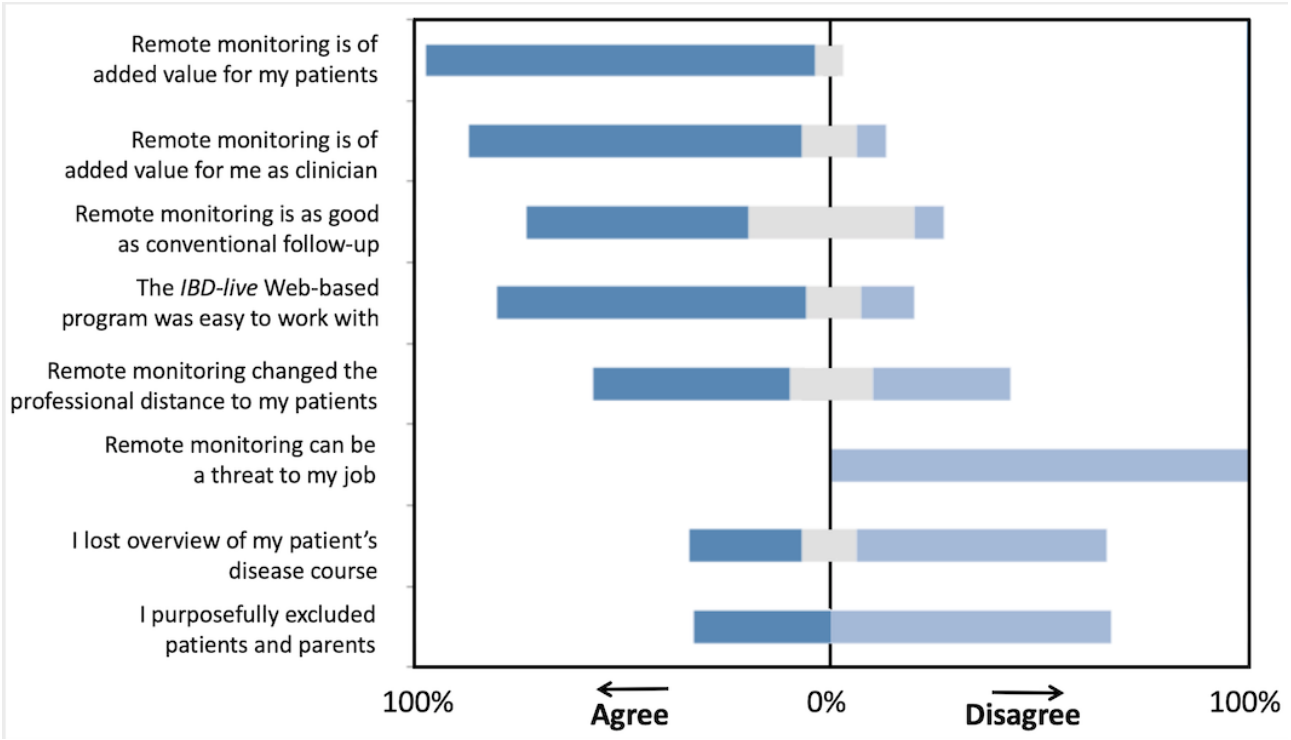
Decision makers may be disappointed to read that cost-savings by the use of telemonitoring were only modest, but physician-nurse task shifting, on the other hand, appeals greatly to health decision makers. As the technology allows to select and target patients who benefit from a face-to-face encounter with their physician at short notice for timely adjustment of therapy plans, nurse specialists in IBD treatment teams may dedicate their time to other problems, including repeat prescriptions, quality-of-life issues, and absenteeism from school or work, not to mention that time will be needed to motivate patients who ignored automated reminders. Successfully accomplishing a task shift probably involves the expansion of the number of rapid access clinic consultation slots. In addition, currently, there is no built-in routine for booking rapid consultations in case of out-of-range results. Although the telemonitoring strategy follows the natural work flow, it conflicts with established hospital electronic databases and therefore requires double data entry. This will put a strain on already overstretched health services. To reduce double documentation, clinicians may want to import data from electronic health records into the cloud-based telemonitoring system or vice versa. This is currently not possible. The organization was perceived as the most problematic of all domains. We classified measures 5A and 5E as *complex*. The last question in domain 5 concerns the work involved in implementation. We regret to say that such a study is extensive and often hidden, and we typically underestimated it at the planning stage. Measures 5B to 5D were classified as *complicated*.

Domain 6: The Wider Context

The most significant system-level challenge at the time that we performed our study was the absence of an established national tariff for reimbursing electronic health consultations. In the meantime, the Dutch Health care Authority agreed that with effect from January 2018, screen-to-screen consultations are reimbursed at a rate equivalent to face-to-face consultations, provided a substantive report is added to the patient's medical record [13]. Reimbursement is only possible when the screen-to-screen consultation is really comparable in both content and duration with a regular face-to-face consultation at the clinic. Second, the screen-to-screen contact should be booked in advance. As financial and regulatory requirements are now in place nationally, we classified this measure as *simple*.

XSL•FO
RenderX

Figure 3. Health provider's opinions about home telemonitoring (n=15). The proportion of respondents who agreed to the statements (left of the neutral line) versus those who disagreed (right of the neutral line). IBD: inflammatory bowel disease.



Domain 7: The Time Dimension

We used email notifications to communicate with the target population, but teenagers hardly use this *ancient form* of digital communication anymore. Despite being an over 40-year-old technology, email has many advantages to newer ways of communication such as WhatsApp, and it is still considered a more trustful way of communication. Email is persistent, easily searchable, one is able to send it from virtually every internet-able device in the world, it uses small data volume, one can attach whatever is wanted, and it is expandable (adding encryption or signing for instance). Having said all this, better, safer, and more appealing tools for communication in health care are a good thing. It will be interesting to see what swoops in to fill the gaps for patient-related data exchange in health care.

As the potential for adapting and coevolving the technology and service over time is uncertain, we classified this measure as *complicated*. With the technology that we used during the RCT, the research head quarter and the Web designer were able to detect critical events quickly and responded to these through coordinated action, therefore question 7B was classified as *simple*.

Discussion

Principal Findings

We recently reported on a telemonitoring strategy for teenagers with IBD (*IBD-live*) that was designed for early detection of relapse and follow-up of appropriateness of treatment. We concluded that the technology was a safe and cost-saving alternative to conventional follow-up. For patients and parents in the intervention arm, the time-saving aspect and the better sense of disease control were highly valued. In this empirical case study, we used the newly developed NASSS framework to explore the challenges to implementing the *IBD-live* strategy and decide whether we should abandon this technology or sustain it within existing sites [9]. The framework allowed us to explore complexity in multiple interacting domains. Aspects of the different domains were classified as simple (+++: straightforward, predictable, and with few components—as in making a sandwich), complicated (++: predictable, but with multiple interacting components or issues—as in building a rocket), or complex (+: dynamic, unpredictable, and not easily disaggregated into constituent components—as in raising a child). Our technology-supported telemonitoring strategy was designed to recognize and manage disease flares at a distance, and following the treatment advice of the Flarometer led to similar patient outcomes as the doctor's advice. This observation led to the conclusion that the condition that we aim to monitor (domain 1) is well characterized, well understood, and predictable (in other words, simple). The technology that we used (domain 2) was user-friendly for both patients and doctors, but a minority of health providers did not trust the treatment advice generated by the Flarometer. As colon pathogens were only occasionally detected in the stool samples with out-of-range results, we feel that their apprehension may not be well founded. The Flarometer algorithm relied on a bespoke solution from a small-sized enterprise, which always carries a risk of supplier withdrawal. However, its substitutability is relatively straightforward, and this aspect of the technology was therefore classified as *complicated*. For patients, parents, and health providers, the value of telemonitoring (domain 3) is in the possibility for *personalized* medicine, such as selecting and targeting the patients who are most likely to benefit from a face-to-face encounter with their specialist. Access to care was

XSL•FO
RenderX

improved in those patients who completed the Flarometer from their home with a result in the low-risk stratum and were reassured without missing school or work, whereas those with an out-of-range result were triaged by an IBD nurse for a physical office visit. However, in a busy, appointment-based practice, scheduling an urgent patient can be a challenge. Organizations that wish to use the telemonitoring strategy to the fullest may want to think of a flexible system to cater for urgent patients, for example, by expanding the number of rapid access clinic consultation slots. The supply-side value was therefore not merely classified as *simple* but rather as *complicated*.

The greatest challenge to scale-up, spread, and sustain our telemonitoring strategy is, without a doubt, the organization's capacity (domain 5) to build the new technology into this workflow, for example, by automating the steps for booking office visits in case of suspected disease flares.

Particular Strengths and Limitations of the Project

The main strength of our telemonitoring project was its randomized and pragmatic design. A heterogeneous group of patients, parents, health providers, and clinical practices were included to maximize the applicability of the results to everyday practice. IBD-live was developed through close collaboration among developers, clinicians, and representatives of the Dutch Crohn's and Colitis patient organization (CCUVN). The director of the CCUVN was consulted at the inception stage of the project and was also on the committee of reviewers who decided on the ZonMw funding.

However, this empirical case study shows that a sturdy evidence base for the efficacy of the intervention does not guarantee a spontaneous movement of the intervention into routine clinical usage. Historically, this research-to-practice gap has not been the concern of academic clinical researchers [14]. We used the NASSS framework retrospectively and were only able to explain why we failed to implement telemonitoring in our daily clinic till now. If we would have used the framework at an early stage of our telemonitoring program development, it would have saved us precious time. In the future, we will definitely use this framework prospectively and in real time to support new programs, and we also encourage others to do so. This case study hinges on the rating of challenges summarized in Table 3. Although 2 assessors (AD and PFvR) independently used a rating matrix, the generalizability of our findings is problematic. There is a great deal of subjectivity in how the challenges were rated, which may have affected the conclusions of this paper.

Comparison With Other Studies

In contrast with our telemonitoring strategy that is not yet mainstreamed, we know of 1 Dutch IBD telemonitoring initiative that has been successfully implemented—myIBDcoach [15,16]. This telemedicine system monitors and registers disease activity in patients with all subtypes of IBD as well as patient-reported outcome measures and quality metrics. We have tried to comprehend why this initiative, which focuses on adult patients with IBD, has a broad support and is expanding to both academic and nonacademic IBD centers. The most important difference with our own project relates to the institution of a foundation to facilitate implementation and to improve cooperation among the various stakeholders. The director of the Dutch IBD patients' organization, 2 gastroenterologists, an accountant, and an assistant professor of health analytics systems constitute the board. Representatives of the IBD section of the Dutch Association for Gastroenterology, the Dutch Association for Gastroenterologists, and the Association of IBD Nurse Specialists form a separate committee which decides on the design and content of myIBDcoach.

Several other telemonitoring initiatives in IBD care, with direct connection to the IBD specialist team, have been developed in Denmark, United States, and the United Kingdom [17-21]. However, in spite of the promising results in the field of responsive and cost-effective care, these initiatives are all in the stage of sustained adoption and none are ready for distant spread.

Interpretation

This case study illustrates that our system with remote biomarker monitoring enhances rather than threatens the existing hospital service with periodic office visits. Home telemonitoring is attractive for teenagers and their families, and health professionals may be interested in using it to keep teenagers who are well out of hospital and ease the pressure on overstretched outpatient services. The telemonitoring system is ideally run with IBD nurses to provide full coverage during working hours. Patients with disease in remission may have up to 4 Flarometer measures in 1 year and 1 routine physical office visit. Patients with an anticipated disease flare will be booked in between for a rapid access clinic consultation. Patients who require periodic day care infusions with antitumor necrosis factor (TNF) monoclonal antibodies may not profit from the telemonitoring service as they will already have unavoidable frequent contact with health care providers. However, those on injectable anti-TNFs, may also highly value our telemonitoring strategy. Home telemonitoring can move IBD care into a new era in which teenagers take ownership of their chronic disease and participate in the therapeutic decision-making process on the basis of longitudinal tracking of Flarometer results. It is important to realize that a wide range of services should be available, as patients and their families have different needs related to both disease activity and personal circumstances. Individual needs also change over time, which require a flexible contact to the hospital.

Abandon or Ready for Sustained Adoption?

Application of the NASSS framework has shown what issues will be mission-critical to move our telemonitoring strategy from a demonstration project to one that is fully mainstreamed and part of business as usual locally. The challenges to be addressed are merely *complicated* (++, ie, predictable and controllable) rather than *complex* (+), which means that our program may be difficult but not impossible to sustain within existing sites.

On the basis of the NASSS framework, we conclude the following:

XSL•FO
RenderX

- The technology and linked program were designed around a well predictable condition (disease flare) in patients with IBD.
- The prototype of the technology was sufficiently tested in an RCT.
- The value proposition of the technology was clear, with a distinct benefit for patients and an affordable service model.
- Health providers have plausible personal reasons to resist (double data entry).
- The organization may not be ready for the innovation, as it requires a shift to new ways of working.
- The concerns about reimbursement are not valid reasons to stall the mainstreaming of the technology.
- The technology is able to adapt and evolve over time in a way that continues to meet the needs of its users.

After integrating the technology and its use with local workflows, we believe that our telemonitoring strategy will be ready for distant spread and long-term sustainability.

Acknowledgments

The authors would like to thank all the colleagues who played an active role in this project—Johanna Escher, Daniëlle Hendriks, Thalia Hummel, Angelika Kindermann, Freddy Kokke, Bart Koot, Tim de Meij, Obbe Norbruis, Sarah Teklenburg-Roord, Herbert van Wering, and Margreet Wessels (pediatric gastroenterologists), and Martha van Gaalen (IBD nurse).

This work was supported by ZonMw Health Care Efficiency Research (grant number 837001001), Innovation Fund Dutch Insurance Companies (grant number B12-204–2509), and NutsOhra Fund (grant number 1301-002). Reagents for the Quantum Blue calprotectin point-of-care tests were an unrestricted donation by Bu¨hlmann Laboratories AG. An unrestricted start-up grant for the development of the Web-based program IBD-live was awarded by Ferring Pharmaceuticals BV. Neither company had a role in the design of this study, nor in the execution, analyses, interpretation of the data or decision to submit results.

Authors' Contributions

Guarantor of the article was PFvR. AD and PFvR conceived the study. AD conducted the primary analysis. AD and PFvR drafted the first version of the article. AH critically revised the article for important intellectual content. All authors have approved the final version of the article, including the authorship list.

Conflicts of Interest

PFvR and AH received funding for joint research projects from BÜHLMANN Laboratories and CisBio Bioassays. AD had no support from any organization for the submitted study, no financial relationships with any organization that might have an interest in the submitted study in the previous 2 years, and no other relationships or activities that could appear to have influenced the submitted study.

Multimedia Appendix 1

Median (inter quartile range, IQR) change in fecal calprotectin. The time interval after a green all-is-well result and a red out-of-range alert was 4 and 2 months respectively. A positive value represents an increase in fecal calprotectin.

[PNG File, 75KB - jmir_v21i3e11761_app1.png]

Multimedia Appendix 2

Presence of colon pathogens in stool samples as per flarometer color code.

[PNG File, 86KB - jmir_v21i3e11761_app2.png]

References

1. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2018 Dec 23;390(10114):2769-2778. [doi: 10.1016/S0140-6736(17)32448-0] [Medline: 29050646]
2. Burisch J, Munkholm P. The epidemiology of inflammatory bowel disease. Scand J Gastroenterol 2015 Aug;50(8):942-951. [doi: 10.3109/00365521.2015.1014407] [Medline: 25687629]
3. Heida A, Park KT, van Rheenen PF. Clinical utility of fecal calprotectin monitoring in asymptomatic patients with inflammatory bowel disease: a systematic review and practical guide. Inflamm Bowel Dis 2017 Dec;23(6):894-902 [FREE Full text] [doi: 10.1097/MIB.0000000000001082] [Medline: 28511198]
4. Maaser C, Sturm A, Vavricka SR, Kucharzik T, Fiorino G, Annese V, European Crohn's Colitis Organisation [ECCO] the European Society of Gastrointestinal Abdominal Radiology [ESGAR]. ECCO-ESGAR guideline for diagnostic assessment

XSL•FO
RenderX

in IBD part 1: initial diagnosis, monitoring of known IBD, detection of complications. J Crohns Colitis 2019 Feb 01;13(2):144-164. [doi: 10.1093/ecco-jcc/jjy113] [Medline: 30137275]
5. Heida A, Dijkstra A, Muller KA, Rossen JW, Kindermann A, Kokke F, et al. The efficacy of home telemonitoring versus conventional follow-up: a randomised controlled trial among teenagers with inflammatory bowel disease. J Crohns Colitis 2018 Mar 28;12(4):432-441. [doi: 10.1093/ecco-jcc/jjx169] [Medline: 29228230]
6. Escher J, Hagemeijer J, de Ridder L, Rings E. Dutch Society for Pediatrics (NVK). 2018 Jul. Diagnosis and treatment of inflammatory bowel disease in children URL: https://www.nvk.nl/Portals/0/richtlijnen/inflammatoire%20darmziekten/inflammatoiredarmziekten.pdf [accessed 2018-07-24] [WebCite Cache ID 719IRHjqz]
7. Pugatch J, Grenen E, Surla S, Schwarz M, Cole-Lewis H. Information architecture of web-based interventions to improve health outcomes: systematic review. J Med Internet Res 2018 Mar 21;20(3):e97 [FREE Full text] [doi: 10.2196/jmir.7867] [Medline: 29563076]
8. Heida A, Dijkstra A, Groen H, Muller KA, Verkade H, van Rheenen P. Comparing the efficacy of a web-assisted calprotectin-based treatment algorithm (IBD-live) with usual practices in teenagers with inflammatory bowel disease: study protocol for a randomized controlled trial. Trials 2015 Jun 16;16:271 [FREE Full text] [doi: 10.1186/s13063-015-0787-x] [Medline: 26073770]
9. 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]
10. Coorevits L, Baert FJ, Vanpoucke HJ. Faecal calprotectin: comparative study of the Quantum Blue rapid test and an established ELISA method. Clin Chem Lab Med 2013 Apr;51(4):825-831. [doi: 10.1515/cclm-2012-0386] [Medline: 23001318]
11. Michie S, McCaffry K, Heneghan C. Monitoring as a learning and motivational tool. In: Glasziou PP, Irwig L, Aronson JK, editors. Evidence Based Medical Monitoring from Principles to Practice. Oxford: Blackwell Publishing; 2008:123.
12. Heida A, Dijkstra A, Dantuma SK, van Rheenen PF. A cross-sectional study on the perceptions and practices of teenagers with inflammatory bowel disease about repeated stool sampling. J Adolesc Health 2016 Oct;59(4):479-481. [doi: 10.1016/j.jadohealth.2016.06.009] [Medline: 27506279]
13. Dutch Healtcare Authority. Directions in reimbursing eHealth URL: https://puc.overheid.nl/nza/doc/PUC_3654_22/1/ [accessed 2018-07-01] [WebCite Cache ID 717yIJ1GS]
14. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol 2015 Sep 16;3:32 [FREE Full text] [doi: 10.1186/s40359-015-0089-9] [Medline: 26376626]
15. de Jong MJ, van der Meulen Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet 2017 Sep 02;390(10098):959-968. [doi: 10.1016/S0140-6736(17)31327-2] [Medline: 28716313]
16. de Jong M, van der Meulen-de Jong A, Romberg-Camps M, Degens J, Becx M, Markus T, et al. Development and feasibility study of a telemedicine tool for all patients with IBD: MyIBDcoach. Inflamm Bowel Dis 2017 Dec;23(4):485-493. [doi: 10.1097/MIB.0000000000001034] [Medline: 28267047]
17. Walsh A, Travis S. What's app? Electronic health technology in inflammatory bowel disease. Intest Res 2018 Jul;16(3):366-373 [FREE Full text] [doi: 10.5217/ir.2018.16.3.366] [Medline: 30090035]
18. Pedersen N, Elkjaer M, Duricova D, Burisch J, Dobrzanski C, Andersen NN, et al. eHealth: individualisation of infliximab treatment and disease course via a self-managed web-based solution in Crohn's disease. Aliment Pharmacol Ther 2012 Nov;36(9):840-849 [FREE Full text] [doi: 10.1111/apt.12043] [Medline: 22971016]
19. Pedersen N, Thielsen P, Martinsen L, Bennedsen M, Haaber A, Langholz E, et al. eHealth: individualization of mesalazine treatment through a self-managed web-based solution in mild-to-moderate ulcerative colitis. Inflamm Bowel Dis 2014 Dec;20(12):2276-2285. [doi: 10.1097/MIB.0000000000000199] [Medline: 25248002]
20. Carlsen K, Houen G, Jakobsen C, Kallemose T, Paerregaard A, Riis LB, et al. Individualized infliximab treatment guided by patient-managed eHealth in children and adolescents with inflammatory bowel disease. Inflamm Bowel Dis 2017 Dec;23(9):1473-1482. [doi: 10.1097/MIB.0000000000001170] [Medline: 28617758]
21. Cross RK, Langenberg P, Regueiro M, Schwartz DA, Tracy JK, Collins JF, et al. A Randomized Controlled Trial of TELEmedicine for Patients with Inflammatory Bowel Disease (TELE-IBD). Am J Gastroenterol 2019 Mar 11;114(3):472-482. [doi: 10.1038/s41395-018-0272-8] [Medline: 30410041]

Abbreviations

CCUVN: Dutch Crohn's and Colitis patient organization
ELISA: enzyme-linked immunosorbent assay
IBD: inflammatory bowel disease
IQR: interquartile range
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability
RCT: randomized controlled trial

XSL•FO
RenderX

TNF: tumor necrosis factor

Edited by G Eysenbach; submitted 01.08.18; peer-reviewed by K Carlsen, P Ware; comments to author 09.10.18; revised version received 11.01.19; accepted 27.01.19; published 29.03.19.

Please cite as:
Dijkstra A, Heida A, van Rheenen PF
Exploring the Challenges of Implementing a Web-Based Telemonitoring Strategy for Teenagers With Inflammatory Bowel Disease: Empirical Case Study
J Med Internet Res 2019;21(3):e11761
URL: http://www.jmir.org/2019/3/e11761/
doi:10.2196/11761
PMID:30924785

©Alie Dijkstra, Anke Heida, Patrick Ferry van Rheenen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring How Evidence is Used in Care Through an Organizational Ethnography of Two Teaching Hospitals

Bryn Lander[1,2], PhD; Ellen Balka[1,2], PhD

[1]Centre for Clinical Epidemiology, Vancouver Coastal Health Research Institute, Vancouver, BC, Canada

[2]School of Communication, Simon Fraser University, Burnaby, BC, Canada

Corresponding Author:
Ellen Balka, PhD
Centre for Clinical Epidemiology
Vancouver Coastal Health Research Institute
702-828 West 10th Avenue
Vancouver, BC, V5Z 1M9
Canada
Phone: 1 604 725 2756
Fax: 1 604 875 5179
Email: ellenb@sfu.ca

Abstract

Background: Numerous published articles show that clinicians do not follow clinical practice guidelines (CPGs). However, a few studies explore what clinicians consider evidence and how they use different forms of evidence in their care decisions. Many of these existing studies occurred before the advent of smartphones and advanced Web-based information retrieval technologies. It is important to understand how these new technologies influence the ways clinicians use evidence in their clinical practice. Mindlines are a concept that explores how clinicians draw on different sources of information (including context, experience, medical training, and evidence) to develop collectively reinforced, internalized tacit guidelines.

Objective: The aim of this paper was to explore how evidence is integrated into mindline development and the everyday use of mindlines and evidence in care.

Methods: We draw on ethnographic data collected by shadowing internal medicine teams at 2 teaching hospitals. Fieldnotes were tagged by evidence category, teaching and care, and role of the person referencing evidence. Counts of these tags were integrated with fieldnote vignettes and memos. The findings were verified with an advisory council and through member checks.

Results: CPGs represent just one of several sources of evidence used when making care decisions. Some forms of evidence were predominately invoked from mindlines, whereas other forms were read to supplement mindlines. The majority of scientific evidence was accessed on the Web, often through smartphones. How evidence was used varied by role. As team members gained experience, they increasingly incorporated evidence into their mindlines. Evidence was often blended together to arrive at shared understandings and approaches to patient care that included ways to filter evidence.

Conclusions: This paper outlines one way through which the ethos of evidence-based medicine has been incorporated into the daily work of care. Here, multiple Web-based forms of evidence were mixed with other information. This is different from the way that is often articulated by health administrators and policy makers whereby clinical practice guideline adherence is equated with practicing evidence-based medicine.

(J Med Internet Res 2019;21(3):e10769) doi:10.2196/10769

KEYWORDS

clinical practice guidelines; evidence-based medicine; mindlines; ethnography

Introduction

Clinical Practice Guidelines (CPGs) have become ubiquitous, they provide concrete practice recommendations for care providers and are often viewed as an integral component of evidence-based medicine [1-3]. Figure 1 shows an idealized evidence-based medicine hierarchy whereby research is appraised, compared, consolidated, and rewritten into CPGs. Despite the significant resources now devoted to producing, adapting, and implementing CPGs, as highlighted by Eby [4]

XSL•FO
RenderX

and others [5,6], CPG adoption in practice has been problematic. Several reviews of existing studies concerned with the underutilization of CPGs focus on barriers to CPG adoption [7-10]. These studies assume that increased CPG uptake leads to more evidence-based care [11]. A few, if any, of these studies considered how information is accessed in the day-to-day work of clinicians.

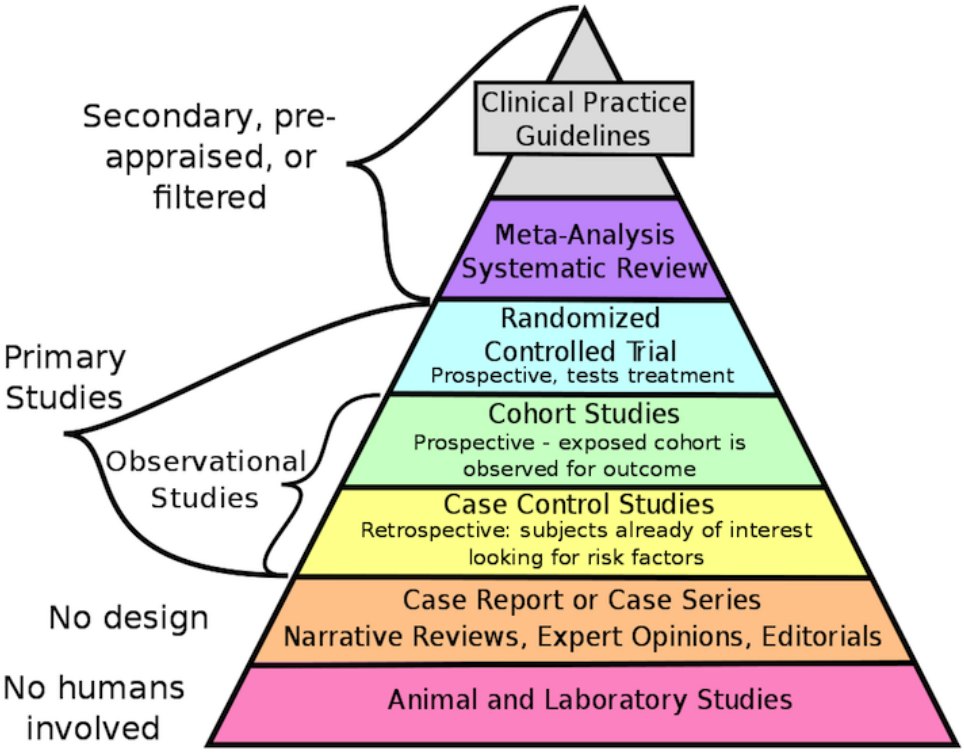
Gabbay and le May's ethnographies [12,13] explored how clinicians use research in their day-to-day work. They argue that general practitioners "normally found [it] neither necessary nor helpful to refer to guidelines or other sources of evidence directly during their day-to-day practice" [12]. Instead, clinicians draw on mindlines, which are defined as "collectively reinforced, internalized tacit guidelines" [13]. Mindlines are socially constructed through interactions with health care professionals and patients, personal medical training, experience, and context. They draw on knowledge in a fluid, multidirectional, and context-specific way. This enables clinicians to prioritize relevant information, reduce possible options for action, and decrease clinical uncertainty [11].

Through interviews, Timmermans and Angell [14] found that residents consult an array of evidence including MD Consult, "cheat" books, textbooks, CPGs, review articles, and primary research articles.

Gabbay and le May's mindlines research was nominated as 1 of 20 influential BMJ papers over the last 20 years [15]. Wieringa and Greenhalgh [11] found in a systematic review of "mindline(s)" conducted 10 years after the publication that a few studies empirically explored mindlines. They argued key components of mindlines—such as knowledge coconstruction and shared sense-making—need further research. Mindlines are formed in teaching hospitals as trainees learn how to think like clinicians [12].

The studies of Gabbay and le May [12,13] and Timmermans and Angell [14] occurred before smartphones and advanced Web-based systems exponentially increased accessibility of information. More recent studies assessed the impact of smartphones and information retrieval technologies on primary care clinicians [16,17] and internal medicine residents [18], arguing that these technologies increase the degree and variety of evidence accessed by clinicians. Others have assessed how virtual networks of clinicians inform the development of mindlines [19]. None of these studies observed how clinicians used scientific evidence in their daily work. This paper explores the ways that scientific evidence was used in the context of everyday practice and how clinicians and trainees draw on a multiplicity of knowledge in their decision-making processes in the age of smartphones and Web-based information systems. Its objective is to understand how evidence is integrated into mindline development and the everyday use of mindlines and evidence. It does so by exploring the ways clinicians teach trainees to use evidence in care, which sorts of evidence clinicians invoke from mindlines and which are read, and how evidence use changes as trainees develop their mindlines.

Figure 1. The evidentiary pyramid in evidence-based medicine. Source: "Research design and evidence" by CFCF. Used with permission by Wikimedia Commons (CC BY-SA 4.0, https://creativecommons.org/licenses/by-sa/4.0).



XSL•FO
RenderX

Methods

Design

This paper is based on an integrated knowledge translation project [20]. Hospital practitioners and administrators formed a project advisory committee, which advised during research design, data collection and analysis, and acted on project findings. The project focused on a problem identified by a health administrator—doctors did not follow CPGs. We used an ethnographic approach to study how evidence, including CPGs, was used in Internal Medicine at 2 teaching hospitals (identified as "Hospital A" and "Hospital B"). Multiple methods were employed, including ethnographic shadowing of care workers, in situ interviewing, and patient chart audits (reported elsewhere). This study received ethics approval from the Research Ethics Boards at the University of British Columbia (CREB#H15-0118) and Simon Fraser University (REB#2015s0388). Both Hospitals A and B granted organizational approval for this project. Informed consent was obtained from all participants directly shadowed or interviewed.

Data Collection

Hospital care in Internal Medicine represents a complex, dynamic, and fast-paced environment. Attendings, residents, and medical students rotate between specialties and hospitals. Rotations range from 1 to 6 weeks, leading to high turnover as teams of workers are constantly being formed and reformed. Internal Medicine at both hospitals is called the Clinical Teaching Unit (CTU) as it is actively involved in training residents and medical students.

One of this paper's authors, BL, shadowed care providers. Shadowing involved following and observing participants during their normal work and asking them, during lull periods, to interpret actions through informal interviews [21,22]. BL shadowed 6 CTU teams between October 2015 and January 2016 in 4 4-week blocks that alternated hospitals. Patients are predominately admitted to the CTU from the Emergency Department (ED). During the first 4-week block at each hospital, BL shadowed the CTU team—comprised an attending and 1 to 3 senior residents—involved in ED patient assessments and admissions as well as senior residents overseeing ED night admissions. During the next 4-week block at each hospital, BL shadowed 2 CTU teams—comprised an attending, a senior resident, 2 to 3 junior residents, and 2 to 4 medical students—involved in patient care on the wards during the day and ED patient admissions during their on-call night shifts. We refer to both residents and medical students as "trainees" unless a subgroup is identified.

All CTU members approached for this study agreed to participate. In total, 35 CTU team members consented to be shadowed or interviewed for the project. Many more CTU team members who were incidentally encountered during shadowing were recorded in fieldnotes. Shadowing occurred in 4-hour blocks to observe work processes while maintaining fieldnote quality. Shadowing was supplemented by a small number of observations of formal teaching during an annual CTU junior resident orientation. Table 1 outlines data collected at Hospitals A and B by data type, hours, location, and number of shifts. In total, BL collected 168 hours of data.

Table 1. Internal medicine team data collection in Hospitals A and B by type, hours, and number of shifts. Includes attending clinicians, residents, and medical students.

Type of data collected	Hospital A, hours (shifts)	Hospital B, hours (shifts)	Total, hours (shifts)
Shadowing			
Emergency department	35 hours, 5 minutes (8)	58 hours, 38 minutes (13)	93 hours, 43 minutes (21)
Internal medicine ward	28 hours (7)	38 hours, 20 minutes (13)	66 hours, 20 minutes (20)
Formal teaching observation	—[a]	4 hours, 15 minutes (3)	4 hours, 15 minutes (3)
Total	63 hours, 5 minutes (15)	101 hours, 13 minutes (29)	168 hours, 18 minutes (44)

[a]—: not applicable.

Data Analysis

Fieldnotes were transcribed into more complete computer-typed notes at the earliest possible opportunity and imported into NVivo (qualitative data analysis software by QSR International). Analysis of data began during data collection. A key early insight was that CPGs were 1 of many evidence sources CTU members referenced during work. This became a focus of subsequent observations and fieldnotes. By reading fieldnotes, we created evidence categories: CPGs, peer-reviewed articles, UpToDate and Lexi-Interact, Google, Pocket Medicine, phone apps, and experience. References to evidence were also categorized by events (teaching or care) and role of person referencing evidence. See Multimedia Appendix 1 for a description of these categories.

We read fieldnotes tagging passages by evidence type, event, and role. Each passage could be tagged multiple times if different sources of evidence were referenced by CTU team members in discussion. Fieldnotes were reread and tags were checked. Using the NVivo query function, we created 2 matrices that counted tags of (1) evidence by event and (2) within care, evidence by role. These matrices were exported into R (the R foundation) [23] for further analysis. We also developed memos that used thick description and vignettes to describe categories.

Preliminary findings were outlined and presented to the project's advisory committee at regular intervals for feedback. Emergent findings were also discussed with informants as a form of "member check" [24]. Results from quantitative category tags were triangulated memo descriptions and feedback to obtain a more holistic understanding of evidence use. All names used

XSL•FO
RenderX

here are pseudonyms. Fieldnote excerpts refer to a fieldnote number followed by a line number (ie, 0116:23).

Results

Evidence Use in Teaching

We found that scientific evidence was used differently in teaching and care. Figure 2 compares teaching and care events and presents counted evidence tags by category as a percent of total evidence tags in each event group. In teaching, the largest percent of counted evidence tags are peer-reviewed articles and experience. Our observations indicated that teaching sessions often began with the introduction of a patient case followed by a patient care plan informed by evidence:

> *Teal [senior resident] is teaching to a room of trainees. She begins by presenting a case where a pregnant woman has a fever, feels dizzy, falls over, and gets admitted to hospital. Over the next half hour, Teal[with input from the residents in the audience] writes out a strategy to initially manage, diagnose, and treat the patient for sepsis. One of the treatment strategies for sepsis is to give patients IV fluids. Teal tells the room that this patient got 2 liters of normal saline solution and asks the audience if that's a good amount to give or if they think the patient should get more or less. She continues, saying that "I want to challenge" the amount of fluids given to a septic patient. "We were all taught the Rivers Protocol." Teal adds "The new approach is to give a certain amount of fluids and then reassess. We were all taught to keep on giving fluids but that is now changing because fluids given above a certain point may be associated with increased mortality." Teal compares these recommendations to the UK sepsis guidelines for pregnant women and mentions that she couldn't find any equivalent Canadian guidelines. She reminds the room of the mantra "healthy mommy is healthy baby" and that in these types of situations the focus should be on caring for the mother not the baby.* [1028:30-174]

Teal used this session to model evidence use that predominately draws from peer-reviewed articles and CPGs. Even though the trainees in the audience graduated from medical school within 3 years, Teal told them their knowledge is out of date as scientific evidence has evolved. In another sepsis teaching session, a Fellow (completing subspecialty training after his internal medicine residency) acting as a guest lecturer discussed the peer-reviewed articles that support the new sepsis approach and rhetorically asked if new evidence means the Rivers dogma is broken. A resident replied by asking the following question:

> *What protocol do we follow if we have a septic patient?* [0721:187-466]

The Fellow left this question unanswered.

In formal teaching, experience was often used to fill holes and cover grey areas in scientific evidence. These were cases where guidelines were not clear:

> *There are no clear guidelines, it's a clinical decision.* [Nicola, attending, 0107:148]

A resident stated the following:

> *Treatments have been used for 100 years and they work so we use them* [Ray, attending, 0719:151]

Another resident stated the following about the treatment:

> *Is empiric, based on what people do not based on a clinical trial.* [Adam, senior resident, 0219:65]

Invoking and Reading Evidence in Patient Care

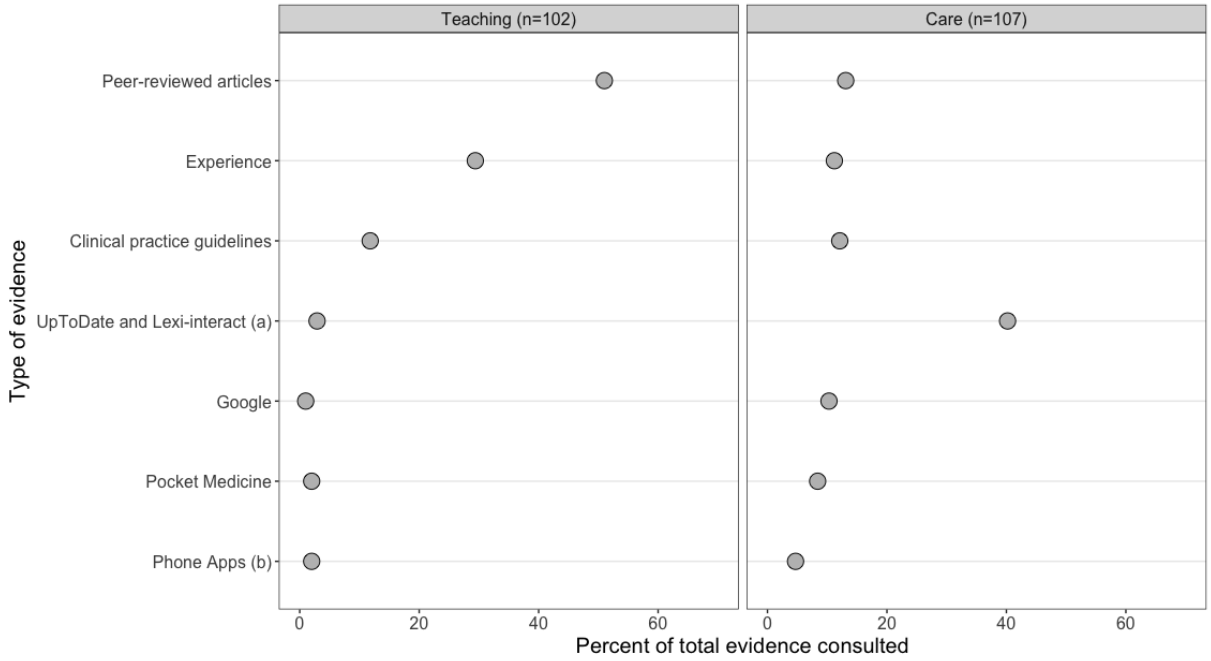
In care, some types of evidence such as experience and peer-reviewed articles were predominately invoked from mindlines, whereas other sources such as UpToDate, Lexi-Interact, Pocket Medicine, and phone apps were read. CPGs were sometimes read and sometimes invoked. The following example illustrates how evidence was invoked in care:

> *Sara [attending] is reviewing a patient admission with Phil [medical student] in the ED. The rest of Sara's CTU team listens. The patient in question is an elderly man who is being admitted due to delirium and a fall. Phil begins to summarise the patient's history and provisional diagnosis, listing delirium as the patient's primary issue. He then begins to outline possible causes for the delirium using the DIMS mnemonic [drugs, infections, metabolic and structural, a mnemonic commonly used to treat delirious patients with no explicit grounding in research or guidelines]. Sara tells the team that "mine is DIMS UFO where U is urine, F is fecal, and O is more involved. We will go over it this afternoon during teaching but the U is based on a paper by [Sara gives the author's name] who is the Godmother of delirium."* [1215:89]

In this example, Sara talked with trainees in a process to coconstruct an appropriate approach to caring for their patients. She made explicit her tacit thought processes and how what sorts of information were incorporated into her mindline.

XSL•FO
RenderX

Figure 2. Observations of different types of evidence consulted in teaching and care; (a) includes UpToDate and Lexi-interact consulted through a phone app, (b) excludes UpToDate and Lexi-interact consulted through a phone app.



Other evidence was read. Attendings would read from Lexi-Interact or UpToDate during conversations with trainees to double-check medication dosing, their inferences, and supplement gaps in their mindlines. Trainees explicitly searched and read evidence when developing care plans:

> *Fred [senior resident] stands outside of a patient's [P2] room. Kate [medical student] comes out of P2's room...Kate tells Fred that P2 grew e. coli in his nephrostomy tube [like a catheter but inserted directly into the kidney] and she thinks that they should prescribe antibiotics. Kate adds that she isn't sure which antibiotics to prescribe...Fred goes to UpToDate and types in the search terms "nephrostomy UTI" and then scans the first UpToDate document that comes up. He goes to the health authority's formulary site and looks for antibiotics for catheter associated infection and skims the article and reads a section out loud to Kate saying 'do not treat positive cultures without symptoms.' Fred asks Kate if she has heard of CAUTI [catheter associated urinary tract infection] and searches for, then shows Kate, a page on UpToDate about the condition. Fred adds that he's not sure if the same approach would apply to a nephrostomy infection. Fred then Googles nephrostomy and infection and finds an article by the Infectious Diseases Society of America about nephrostomy and infection. Kate smiles at Fred and tells him that she likes that organisation. They skim the article together but it has no recommendations about what they should do if they get a positive culture. Kate takes her IPad out and begins to search for information as well. She finds another entry in UpToDate that says that cultures in asymptomatic patients shouldn't be treated but that the tube should be changed as soon as possible. Fred looks at P2's lab results. P2 has no white blood count and he is afebrile. Kate adds that she worries about P2 because he is immunocompromised...Fred decides that they won't treat for the positive culture but that they might try and get P2's tube changed.* [0105:431]

In this case, Fred and Kate sifted through various sources of evidence and personal knowledge together to make sense of a clinical problem and develop a care plan. Fred drew from his own knowledge of how to treat catheter associated urinary tract infections, but he was unsure if nephrostomy associated infections were similar enough for his knowledge to be applicable. Both Fred and Kate read to supplement what they know, using computers and iPads to access information on the Web. The information selected for reading was not peer-reviewed articles but was summaries of existing evidence with concrete recommendations. Multiple Web-based sources were skimmed for relevant information. Through discussions, Fred and Kate combined the knowledge extracted from summary documents, their nascent mindlines, patient laboratory results, and their knowledge of the patient to form a provisional care plan for P2.

CPGs were invoked or read. For example, CPGs were invoked when 2 patients were admitted to the CTU for carbon monoxide poisoning. Still in the ED, the patients were scheduled for 3 hyperbaric dives (where they would be placed in a chamber with high pressured oxygen to reduce the amount of carbon monoxide in their blood). Matt (attending) told BL that he believed the carbon monoxide level in both patients was low and he was not sure if the dives were necessary. Matt greeted

XSL•FO
RenderX

the ED doctor overseeing the hyperbaric chamber and asked him if the hyperbaric dives were necessary. He explained that the patients' carbon monoxide levels were already low and asked the following question:

> *How much does it help at this point?* [Matt, attending]

The ED doctor answered that current best practice said to do the 3 dives and added the following:

> *There is absolutely no evidence but it is standard of care so we'll do it. We'll do 3. They'll get the benefits that they'll get...the literature is unclear, no consistent evidence of benefit.* [1023:372]

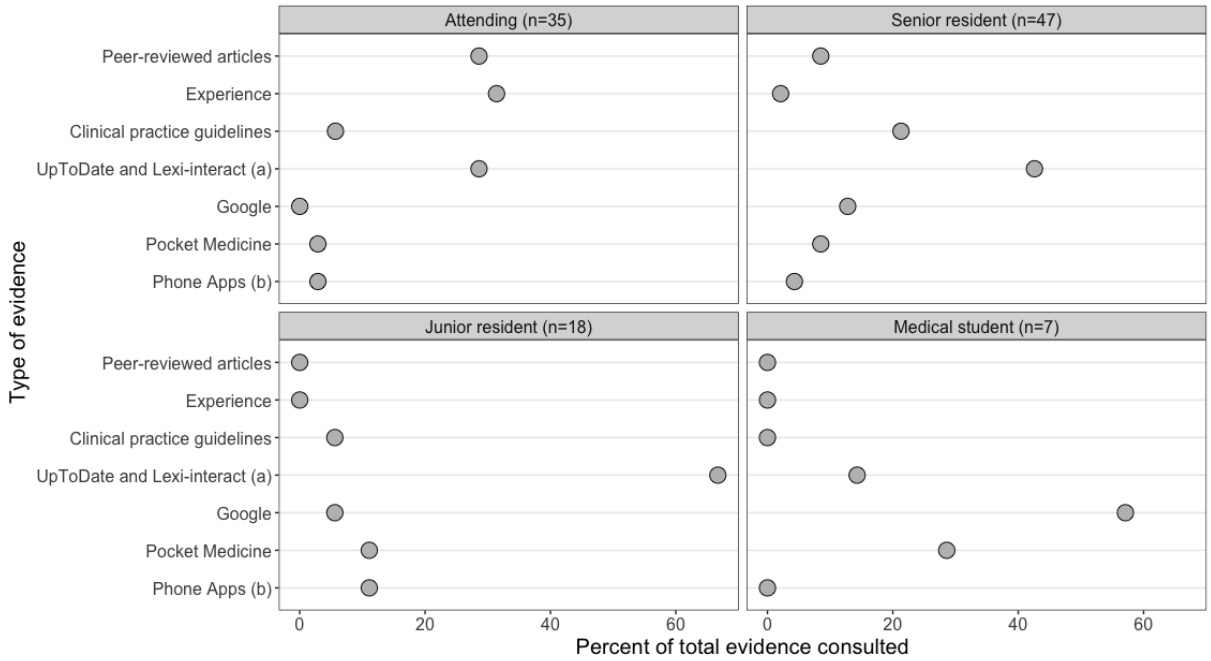
Here, both Matt and an ED doctor made sense of an appropriate treatment plan for 2 patients. Matt questioned the agreed-upon approach, citing specific laboratory results for these 2 patients. The ED doctor invoked "best practice" to justify his own mindline for treating carbon monoxide poisoning. In this case, "best practice" was not viewed as evidence based. Rather, the ED doctor believed that following the best practice is a prudent approach in the absence of clear evidence. In other situations, trainees read CPGs to help inform care decisions. Local guidelines are one of several Web-based evidence sources used by Fred and Kate to formulate a care plan for P2.

Developing Mindlines in Patient Care

In care, evidence use varied by role. Figure 3 compares care roles and presents counted evidence tags by category as a percent of total evidence tags for each role. Figure 3 shows that individuals in all roles read UpToDate. Medical students and junior residents predominately read summary sources of evidence such as UpToDate, Lexi-interact, Pocket Medicine, and phone apps. Conversely, attendings predominately invoked their own experience or knowledge of peer-reviewed articles from their mindlines. Senior residents consulted the widest variety of evidence sources not only drawing on summary sources (particularly UpToDate) but also invoking peer-reviewed articles and experience. Senior CTU team members may be more capable of using their mindlines to filter scientific uncertainty than more junior trainees who are developing their mindlines. Consistent with the mindlines concept, it was widely recognized that clinical and scientific uncertainty often remain despite referring to a broad array of evidence and information and that clinicians may ultimately make choices on the basis of their interpretation of this uncertainty in the context of patient care needs.

We observed that trainees would often gather multiple kinds of information to develop a narrative of patient diagnosis and care that was presented to an attending. This formal presentation was used by attendings to review trainees' reasoning processes and how knowledge of existing evidence was being incorporated into their developing mindlines. Attendings guided trainees on how mindlines were formed by quizzing them on how to think logically and identify and filter evidence and information when developing patient care plans.

Figure 3. Observations of different types of evidence consulted in care by role; (a) includes UpToDate and Lexi-interact consulted through a phone app, (b) excludes UpToDate and Lexi-interact consulted through a phone app.



XSL•FO
RenderX

Discussion

Principal Findings

This study explored the integration of evidence into mindline development and how mindlines and evidence are used. In this study, multiple forms of Web-based evidence were accessed—often through smartphones—and mixed with other forms of information. Evidence use varied between teaching and care and, within care, by role. Teaching modeled optimal evidence use and predominately drew on peer-reviewed journals and experience. In care, attendings' evidence use most closely approximated the evidence use modeled in teaching. Some forms of evidence—such as experience, peer-reviewed journals, and some CPGs—were predominately invoked through mindlines. Other forms of evidence—such as UpToDate, Lexi-Interact, Pocket Medicine, and phone apps—were read to supplement mindlines. As CTU team members gained experience, they increasingly incorporated evidence into their mindlines. An alternative explanation to the observed variation of evidence use by role is a cohort effect whereby trainees are more comfortable accessing Web-based information systems and using smartphones. Although a cohort effect may explain some of the observed variance, it does not explain differences in how senior and junior residents (separated by a year) consult evidence.

In 1961, Becker et al [25] reported instances where experience was invoked to overrule scientifically verified knowledge—a phenomenon we did not observe. Although Gabbay and le May [12] state that they rarely observed care providers reading summary evidence to inform their daily practice, we observed this practice across all groups. These 3 studies may represent different stages in the institutionalization of the evidence-based medicine paradigm into clinical practice representing pre [25], early [12], and later (seen here) stages. Our observations may also be a result of the increasing codification of evidence in commercially produced apps and other tools and their ease of access on smartphones and computers. Gabbay and le May [12] noted that by the end of their observations, clinicians were consulting evidence more frequently. Many of the clinicians observed by Gabbay and le May [12] trained before 1990, when evidence-based medicine was introduced as part of medical school training, whereas our study occurred in teaching hospitals where evidence-based medicine is emphasized. Gabbay and le May [12] noted that UpToDate was referred to as a "bible" during their observations in a teaching hospital's internal medicine specialty.

The sepsis example illustrates that the implications of evolving scientific knowledge on daily clinical practice are not always explicit [14]. CTU team members were also cognizant that existing evidence is occasionally of low quality, can make competing claims, and may not cover a clinical problem.

Our research suggests that evidence-based medicine has been incorporated into the work of the CTU teams studied here, just not in the ways envisioned by health administrators and policy makers who frequently equate low CPG uptake and a lack of standardization with not practicing evidence-based medicine. Instead of drawing on CPGs, CTU teams invoke and read a variety of evidence sources, blending them together to arrive at shared understandings of patient care [19,26,27]. As clinicians increasingly rely on their smartphones to access information, health administrators and policy makers will decreasingly be able to control the type and quality of information they access. The ways that managers assess and support clinician use evidence need to be expanded. It is important that these players understand current use patterns to support the application of quality evidence into clinical practice decisions.

Limitations and Future Research

Many insights gained from shadowing internal medicine teams were based on a small number of teams at 2 hospitals, which limits generalizability. There is a chance that participants increased their use of evidence as they knew they were being shadowed (ie, Hawthorne effect). To decrease this impact, BL maintained a neutral attitude throughout her observations, emphasized that her primary interest was to understand how work was conducted, and shadowed an internal medicine team for at least 1 week.

Ethnographic research enables a rich understanding of the complex world around us. Writing fieldnotes necessarily distills from this complexity. Observing evidence use was a primary objective in these fieldnotes, but it is probable that not all instances where evidence was used were captured. By observing in depth, ethnographic work cannot capture the same breadth of data as other methodologies. The numbers for some observations—such as counts of evidence use tags by junior residents and medical students in care—are small. Specific numbers or detailed analyses of variance are not provided in the text as counts cannot be considered exact. Rather, they represent trends that were combined with memos and general observations and then checked with informants and experts to derive insight. Additional quantitative studies that track information use on phones and computers are essential to validate findings.

UpToDate, Lexi-Interact, and Pocket Medicine are all published by the same private company. Combined, they represent over half of the evidence we observed being consulted. An avenue of future research is to compare the recommendations provided by this company to other sources of evidence to understand if a bias is being created in how medicine is practiced. It is unclear if the work practices observed in these 2 teaching hospitals have diffused into community-based hospitals and general practices. Further research is needed to understand whether CPGs and other tools such as order sets may be more effective vehicles to incorporate evidence into practice in other settings.

Conclusions

This paper has responded to calls for more empirical work exploring the complex ways that evidence and mindlines are incorporated into teaching and care [4,11]. It shows how mindlines are enacted [11-13] as clinicians blend multiple forms of evidence together to arrive at shared understandings and approaches to patient care and understand how this process varies as trainees develop their mindlines.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the members of our Advisory Committee whose valuable insights helped to steer this project. Finally, this study would not have been possible without the numerous health care providers who let BL follow them around during their daily work. BL was supported by the Michael Smith Foundation for Health Research Postdoctoral Trainee award in Knowledge Translation Science.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Fieldnote categorization.

[PDF File (Adobe PDF File), 91KB - jmir_v21i3e10769_app1.pdf]

References

1. Mykhalovskiy E, Weir L. The problem of evidence-based medicine: directions for social science. Soc Sci Med 2004 Sep;59(5):1059-1069. [doi: 10.1016/j.socscimed.2003.12.002] [Medline: 15186905]
2. Pope C. Resisting evidence: the study of evidence-based medicine as a contemporary social movement. Health (London) 2003 Jul 21;7(3):267-282. [doi: 10.1177/1363459303007003002]
3. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay V. The emergence of clinical practice guidelines. Milbank Q 2007 Dec;85(4):691-727 [FREE Full text] [doi: 10.1111/j.1468-0009.2007.00505.x] [Medline: 18070334]
4. Eby D. Physicians don't follow guidelines: what are we to do? Can Med Assoc J 2016 Feb 16;188(3):217 [FREE Full text] [doi: 10.1503/cmaj.1150080] [Medline: 26883888]
5. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003 Oct 11;362(9391):1225-1230. [doi: 10.1016/S0140-6736(03)14546-1] [Medline: 14568747]
6. Timmermans S, Berg M. The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care. Philadelphia: Temple University Press; 2010.
7. Kastner M, Bhattacharyya O, Hayden L, Makarski J, Estey E, Durocher L, et al. Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review. J Clin Epidemiol 2015 May;68(5):498-509 [FREE Full text] [doi: 10.1016/j.jclinepi.2014.12.013] [Medline: 25684154]
8. Grimshaw J, Thomas R, MacLennan G, Fraser C, Ramsay C, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004 Feb;8(6):iii-iv, 1 [FREE Full text] [Medline: 14960256]
9. Grimshaw J, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001 Aug;39(8 Suppl 2):II2-I45. [Medline: 11583120]
10. Cabana M, Rand C, Powe N, Wu A, Wilson M, Abboud P, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. J Am Med Assoc 1999 Oct 20;282(15):1458-1465. [doi: 10.1001/jama.282.15.1458] [Medline: 10535437]
11. Wieringa S, Greenhalgh T. 10 years of mindlines: a systematic review and commentary. Implement Sci 2015 Apr 09;10(1):45 [FREE Full text] [doi: 10.1186/s13012-015-0229-x] [Medline: 25890280]
12. Gabbay J. Practice-based Evidence For Healthcare: Clinical Mindlines. Oxford, United Kingdom: Routledge; 2011.
13. Gabbay J, le May A. Evidence based guidelines or collectively constructed "mindlines?" Ethnographic study of knowledge management in primary care. Br Med J 2004 Oct 30;329(7473):1013 [FREE Full text] [doi: 10.1136/bmj.329.7473.1013] [Medline: 15514347]
14. Timmermans S, Angell A. Evidence-based medicine, clinical uncertainty, and learning to doctor. J Health Soc Behav 2001 Dec;42(4):342-359. [doi: 10.2307/3090183] [Medline: 11831136]
15. Greenhalgh T, Annandale E, Ashcroft R, Barlow J, Black N, Bleakley A, et al. An open letter to The BMJ editors on qualitative research. Br Med J 2016 Feb 10;352:i563. [doi: 10.1136/bmj.i563] [Medline: 26865572]
16. McKibbon K, Fridsma D. Effectiveness of clinician-selected electronic information resources for answering primary care physicians' information needs. J Am Med Inform Assoc 2006;13(6):653-659 [FREE Full text] [doi: 10.1197/jamia.M2087] [Medline: 16929042]
17. Pluye P, Grad R. How information retrieval technology may impact on physician practice: an organizational case study in family medicine. J Eval Clin Pract 2004 Aug;10(3):413-430. [doi: 10.1111/j.1365-2753.2004.00498.x] [Medline: 15304142]
18. León SA, Fontelo P, Green L, Ackerman M, Liu F. Evidence-based medicine among internal medicine residents in a community hospital program using smart phones. BMC Med Inform Decis Mak 2007 Feb 21;7:5 [FREE Full text] [doi: 10.1186/1472-6947-7-5] [Medline: 17313680]

XSL•FO
RenderX

19. Wieringa S, Engebretsen E, Heggen K, Greenhalgh T. How knowledge is constructed and exchanged in virtual communities of physicians: qualitative study of mindlines online. J Med Internet Res 2018 Feb 02;20(2):e34 [FREE Full text] [doi: 10.2196/jmir.8325] [Medline: 29396385]
20. Kothari A, Wathen CN. A critical second look at integrated knowledge translation. Health Policy 2013 Feb;109(2):187-191. [doi: 10.1016/j.healthpol.2012.11.004] [Medline: 23228520]
21. McDonald S, Simpson B. Shadowing research in organizations: the methodological debates. Qual Research Organ Manag 2014 Mar 04;9(1):3-20. [doi: 10.1108/QROM-02-2014-1204]
22. Quinlan E. The 'actualities' of knowledge work: an institutional ethnography of multi-disciplinary primary health care teams. Sociol Health Illn 2009 Jul;31(5):625-641 [FREE Full text] [doi: 10.1111/j.1467-9566.2009.01167.x] [Medline: 19392938]
23. R Core Team. Vienna, Austria: R foundation for statistical computing; 2016. A language and environment for statistical computing URL: https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing [accessed 2019-02-21] [WebCite Cache ID 76LlgRvOO]
24. Creswell J, editor. Qualitative Inquiry And Research Design: Choosing Among Five Approaches (Third Edition). USA: Sage Publications, Inc; 2013.
25. Becker H, Geer B, Hughes E, Strauss A. Boys in white: Student Culture in Medical School. New Brunswick: Transaction Publishers; 2005.
26. Bal R. Evidence-based policy as reflexive practice. What can we learn from evidence-based medicine? J Health Serv Res Policy 2017 Dec;22(2):113-119. [doi: 10.1177/1355819616670680] [Medline: 28429971]
27. Oertle M, Bal R. Understanding non-adherence in chronic heart failure: a mixed-method case study. Qual Saf Health Care 2010 Dec;19(6):e37. [doi: 10.1136/qshc.2009.033563] [Medline: 21127097]

Abbreviations

CPG: clinical practice guideline
CTU: clinical teaching unit
ED: emergency department

Edited by G Eysenbach; submitted 12.04.18; peer-reviewed by R Bal, T Arcury, A Cyr; comments to author 28.08.18; revised version received 23.10.18; accepted 10.12.18; published 28.03.19.

Please cite as:
Lander B, Balka E
Exploring How Evidence is Used in Care Through an Organizational Ethnography of Two Teaching Hospitals
J Med Internet Res 2019;21(3):e10769
URL: http://www.jmir.org/2019/3/e10769/
doi:10.2196/10769
PMID:30920371

©Bryn Lander, Ellen Balka. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Online Health Information Seeking and eHealth Literacy Among Patients Attending a Primary Care Clinic in Hong Kong: A Cross-Sectional Survey

David Ka-Ki Wong[1], MBChB, MSc, FHKAM; Man-Kuen Cheung[1], MBBS, FHKAM

University Health Service, The University of Hong Kong, Pokfulam, China (Hong Kong)

Corresponding Author:
David Ka-Ki Wong, MBChB, MSc, FHKAM
University Health Service
The University of Hong Kong
2/F Meng Wah Complex
The University of Hong Kong
Pokfulam,
China (Hong Kong)
Phone: 852 39172502
Fax: 852 25481430
Email: dkkw@hku.hk

Abstract

Background: Previous studies have suggested that patients' online health information seeking affects their medical consultations and patient-doctor relationships. An up-to-date picture of patients' online health information-seeking behaviors can inform and prepare frontline health care professionals to collaborate, facilitate, or empower their patients to access and manage health information found online.

Objective: This study explores the prevalence, patterns, and predictors of online health information-seeking behaviors among primary care patients in Hong Kong, and the relations between online health information seeking and electronic health (eHealth) literacy.

Methods: Patients attending a university primary care clinic in Hong Kong were asked to complete a questionnaire survey on their demographic backgrounds; health status; frequency and pattern of online health information seeking; contents, sources, and reasons for online health information seeking; and their eHealth literacy. eHealth literacy was measured by the validated eHealth Literacy Scale (eHEALS). Regression analyses explored various demographic and behavioral predictors to online health information seeking, and predictors to eHealth literacy.

Results: In all, 97.32% (1162/1194) respondents used the internet, of which 87.44% (1016/1162) had used the internet to find health information. Most respondents (65.97%, 665/1008) searched once monthly or more. Few (26.88%, 271/1008) asked their doctor about health information found online, but most doctors (56.1%, 152/271) showed little or no interest at all. The most sought topic was symptom (81.59%, 829/1016), the top reason was noticing new symptoms or change in health (70.08%, 712/1016), the most popular source was online encyclopedia (69.98%, 711/1016), and the top reason for choosing a source was convenience (55.41%, 563/1016). Poisson regression analysis identified high eHEALS score, fair or poor self-rated health, having a chronic medical condition, and using the internet several times a day as significant predictors of online health information seeking. Multiple regression analysis identified lower age, better self-rated health, more frequent internet use, more frequent online health information seeking, and more types of health information sought as significant predictors to higher eHealth literacy.

Conclusions: Online health information seeking is prevalent among primary care patients in Hong Kong, but only a minority shared the information with doctors. Websites were chosen more for convenience than for accuracy or authoritativeness. Doctors should recognize patients' online health information-seeking behavior, and facilitate and empower them to search for high-quality online health information.

(J Med Internet Res 2019;21(3):e10831) doi:10.2196/10831

KEYWORDS

online health information seeking; eHealth literacy; primary care; Hong Kong

XSL•FO
RenderX

Introduction

The rapid development of the internet has significantly changed the way people obtain information. Now a vast amount of information can be accessed instantly. In the health care domain, the internet has changed the way people find and receive health information, from passive information received from doctors' advice and the mass media to active information sought through Web searches (asking "Dr Google" [1]). Sources of information have also evolved from static and authoritative sources such as books and printed journals to the more dynamic and user-contributed contents such as blogs, online forums, and social networking sites [2-4].

Online health information seeking by patients has been shown to affect medical consultations and patient-doctor relationships [5-8]. Patients increasingly turn to the internet to prepare for doctors' consultations, to discuss information they found online with their doctors, or to complement, validate, and challenge the information offered by doctors [7,9-11]. Various predictors, such as age, gender, education, socioeconomic status, health status, and internet usage, were shown to affect the prevalence and extent of online health information seeking in previous studies [10-14]. The Pew Internet and American Life Project has shown that female gender, young age, and high education level were associated with more online health information seeking [12]; whereas, in a survey of primary care patients in Scotland, employment status, educational attainment, geographic location, age, and gender were associated with online health information seeking [10].

A bibliometric study in 2015 identified 533 publications on the topic of online health information seeking [15]. Although the topic has become an increasingly important research focus, the majority of the research was from the United States, and either focused on internet users in general (eg, consumers, students, parents) or patients with specific diseases (eg, human immunodeficiency virus, sexually transmitted diseases, cancer). In Hong Kong, only one exploratory study on the prevalence and patterns of online health information seeking was done in 2006; a convenience sample of 443 members of the general public was surveyed in shopping malls and subway stations, and 44% of respondents had looked for health information online in the past 6 months. Among the health surfers, the majority (78%) visited websites from the government, hospitals, or nonprofit organizations [16].

With the rapid development of the internet and its widespread use, doctors will encounter more "internet-informed" patients. Government statistics showed a 26.5% increase in the proportion of the Hong Kong population using the internet in the past decade (from 62.9% in 2006 to 89.4% in 2017) [17], and the prevalence and patterns of online health information seeking will likely have changed since the previous study by Yan [16]. This study aims to map an up-to-date picture of online health information-seeking behaviors among primary care patients in Hong Kong. The results can better inform frontline doctors, public health professionals, and health educators on this issue, and better prepare them to collaborate, facilitate, or empower patients to access and manage health information online.

The study objectives were to (1) determine the prevalence and pattern of online health information seeking, (2) explore the contents, sources, and reasons for online health information seeking, (3) explore the predictors of online health information seeking, and (4) explore the predictors of electronic health (eHealth) literacy.

Methods

Study Design

The study was a cross-sectional, anonymous, self-administered questionnaire survey. Consecutive sampling was used. All patients attending a university primary care clinic in Hong Kong for consultation during the data collection period (March to April 2017) were invited to complete the questionnaire. The following groups of patients were excluded from the study: (1) age younger than 16 years, (2) unable to consent (eg, mentally incapacitated), (3) unable to read or understand the questionnaire, and (4) those who had filled in the questionnaire previously.

Questionnaires were distributed to patients by nursing staff during preconsultation health assessments. Patients could opt for either the Chinese or English version of the questionnaire, which was to be filled in while awaiting consultation and returned to collection boxes in the waiting hall. The survey required approximately 5 minutes to complete.

Study Population and Sampling

The study population was patients attending the primary care clinic of a university in Hong Kong. The clinic serves university students, staff, and their dependents, as well as certain retirees. The clinic population is approximately 47,000.

For sample size estimation, the formula for cross-sectional studies was used, where sample size $n = Np(1-p)/[(d^2/Z^2_{1-\alpha/2}*(N-1)+p(1-p)]$ [18]. Given a population size (N) of 47,000, a hypothesized proportion (p) of 0.5, and a margin of error (d) of 0.05, the minimal sample size required ranged from 382 (95% confidence level) to 655 (99% confidence level).

Survey Instrument

The questionnaire was developed based on a review of the literature on online health information seeking and eHealth literacy. Items from previously validated instruments were included where appropriate, including one item on self-rated health status (from SF-12 version 2 health survey [19]), and the full eHealth Literacy Scale (eHEALS) [20]. The final questionnaire consisted of 25 items, covering demographic backgrounds, health status, online health information-seeking behavior, and eHealth literacy.

The eHEALS was used because it is the most widely used validated measure of eHealth literacy; it has been validated with various population groups [20-22]. eHEALS contains eight questions on a 5-point Likert scale, of which various aspects of self-perceived eHealth literacy were measured. The sum of all items is a composite measure, with high scores indicating greater

XSL•FO
RenderX

literacy. Permission was obtained from the original author for reuse and translation of eHEALS in this study.

The questionnaire was reviewed by five domain experts (two family physicians, one community pharmacist, one health education nurse, and one public health researcher), and content validity of each question was rated on a 4-point Likert scale (not relevant, somewhat relevant, quite relevant, highly relevant). The item-level content validity index (CVI) was computed as the proportion of experts who rated a question as quite or highly relevant [23]. The item-level CVIs of all questions were rated 1.00, and the scale-level CVI thus computed was also 1.00.

The Chinese version was translated by the principal investigator with feedback from the domain experts; back-translation was done by a professional translator to ensure the two language versions were conceptually equivalent [24]. In this study, Cronbach alpha of the Chinese version of eHEALS was .891, and that of the English version was .918, which indicates a high level of internal consistency, and matched the Cronbach alpha of .88 in the original study [20].

The questionnaires were pilot-tested on 52 patients of different gender, age, and education level, and they were individually debriefed by the principal investigator. Some minor rewordings on the Chinese version were done based on the feedback received, and the questionnaires were finalized after a second round of back-translation (Multimedia Appendices 1 and 2).

Statistical Analysis

Data were analyzed using IBM SPSS Statistics version 24.0 (IBM Corp, Armonk, NY, USA). Frequency tables were computed to check for completeness, range, and consistency. Descriptive statistics were computed to summarize the data, with means and standard deviations calculated where applicable. A Poisson regression analysis explored the demographic and behavioral predictors to the extent of online health information seeking. A multiple regression analysis explored the demographic and behavioral predictors to eHealth literacy. Statistical significance was established at $P<.05$ for all tests.

Ethical Consideration

The study was an anonymous survey, with no personal information collected, and involved minimal risk. Participation was completely voluntarily; patients could refuse to participate without any negative consequences. The purpose of the study was explained in a cover letter; informed consent was implied by completing the questionnaire. Ethics approval was received from the Human Research Ethics Committee of The University of Hong Kong (ref: EA1702020) before the study commenced.

Results

A total of 1291 questionnaires were distributed, which yielded a response rate of 94.50% (1220/1291). Of the returned questionnaires, 26 were excluded from analysis due to grossly incomplete data (eg, missing most demographic data or a whole section of questions). This sample of 1194 respondents represented 2.54% of the total clinic population (N=47,000). Overall, 91.96% (1098/1194) of respondents opted for the Chinese questionnaire.

Demographics of Study Sample

The demographic characteristics of the study sample are listed in Table 1. Approximately 60% (717/1179) of respondents were female. The respondents spanned all age groups and occupation ranks. Approximately half of the respondents were students with a tertiary education level. Most respondents rated their own health as good (40.35%, 481/1192) or very good (29.78%, 355/1192), whereas 26.17% (312/1192) rated their health as fair or poor. In addition, 19.10% (225/1178) had chronic medical conditions requiring regular follow-up or treatment. Almost all respondents (97.32%, 1162/1194) used the internet; the majority of them used the internet several times a day (74.46%, 863/1159) and spent more than 3 hours per day on the internet (51.42%, 596/1159).

Prevalence and Pattern of Online Health Information Seeking

Of the respondents who used the internet, 87.44% (1016/1162) had found health information online in the past (Table 2). Most of them reported a frequency of online health information seeking from once every few months (29.27%, 295/1008), once a month (16.87%, 170/1008), to several times a month (24.01%, 242/1008). Other than finding information for oneself (94.96%, 961/1012), most also searched on behalf of family members (69.17%, 700/1012), and some searched for friends and colleagues (29.25%, 296/1012). The majority used mobile phones (74.11%, 747/1008) and laptop computers (58.43%, 589/1008) for online health information seeking.

XSL•FO
RenderX

Table 1. Characteristics of the study sample (N=1194).

Characteristic	n (%)	Missing data, n
Age (years)		**5**
16-24	572 (48.11)	
25-34	220 (18.50)	
35-44	143 (12.03)	
45-54	124 (10.43)	
≥55	130 (10.93)	
Gender		**15**
Male	462 (39.19)	
Female	717 (60.81)	
Education level		**6**
Secondary and below	144 (12.12)	
Tertiary	646 (54.38)	
Postgraduate	398 (33.50)	
Occupation		**8**
Managers, professionals, and academic staff	244 (20.57)	
Technicians and associate professionals	70 (5.90)	
Clerical, services, and sales workers	161 (13.58)	
Craft workers and laborers	10 (0.84)	
Student	640 (53.96)	
Housewife, retired, and unemployed	61 (5.14)	
Self-rated health status		**2**
Excellent	46 (3.86)	
Very good	355 (29.78)	
Good	481 (40.35)	
Fair	290 (24.33)	
Poor	20 (1.68)	
Have chronic medical condition	225 (19.10)	16
Habit of using the internet	1162 (97.32)	
Access the internet with: [a,b]		
Desktop computer	533 (45.87)	
Laptop computer	801 (68.93)	
Tablet	322 (27.71)	
Mobile phone	941 (80.98)	
Frequency of using the internet[b]		**3**
Several times a week or less	46 (3.97)	
Every day	250 (21.57)	
Several times a day	863 (74.46)	
Time spent using the internet per day (hours)[b]		**3**
<1	64 (5.52)	
1 to <2	231 (19.93)	
2 to <3	268 (23.12)	

XSL•FO
RenderX

Characteristic	n (%)	Missing data, n
3 to <4	187 (16.13)	
≥4	409 (35.29)	
Use wearable health-monitoring devices[b]	189 (16.31)	

[a]Multiple responses allowed.

[b]Questions only for respondents who used the internet (n=1162).

Table 2. Prevalence and pattern of online health information seeking for patients who had used the internet to find health information (n=1016).

Characteristic	n (%)	Missing data, n
Frequency		**8**
Once a year or less	48 (4.76)	
Every few months	295 (29.27)	
Once a month	170 (16.87)	
Several times a month	242 (24.01)	
Once a week	92 (9.13)	
Several times a week	118 (11.71)	
Every day	43 (4.27)	
Finding information for:[a]		**4**
Myself	961 (94.96)	
Family members	700 (69.17)	
Friends or coworkers	296 (29.25)	
Devices used[a]		**8**
Desktop	410 (40.67)	
Laptop	589 (58.43)	
Tablet	192 (19.05)	
Mobile phone	747 (74.11)	
Asked doctor about online health information	271 (26.88)	8
Shared online health information with doctor[b]	124 (45.8)	
Asked about specific disease[b]	222 (81.9)	
Asked for specific treatment, test, or referral[b]	143 (52.8)	
Doctors' interest about online health information[b]		
Very interested	6 (2.2)	
Quite interested	47 (17.3)	
Slightly interested	106 (39.1)	
Not at all interested	46 (17.0)	
Don't know or can't remember	66 (24.4)	

[a]Multiple responses allowed.

[b]Questions for respondents who asked a doctor about online health information only (n=271).

Only a minority of respondents (26.88%, 271/1008) had ever asked a doctor about health information they found online. Of those who asked, 45.8% (124/271) shared the information with the doctor through email, printout, or mobile phone screenshots. The majority of them asked about a specific disease (81.9%, 222/271), or specific treatment, test, or referral (52.8%, 143/271). When asked about health information found online, the perceived responses from doctors were uninspiring: 39.1% (106/271) of doctors were only slightly interested, and 17.0% (46/271) were not at all interested about the health information respondents found online.

XSL•FO
RenderX

Contents, Sources, and Reasons for Online Health Information Seeking

A variety of online health information was sought by the respondents (Table 3). The top three types of information sought were symptoms (81.59%, 829/1016), a disease or condition (70.47%, 716/1016), and medication (57.19%, 581/1016). A total of 44.49% (452/1016) of respondents sought information on healthy behaviors. A minority of respondents sought information on Chinese medicine (20.77%, 211/1016), alternative medicine (11.52%, 117/1016), and health insurance (10.93%, 111/1016). On average, 4.36 types of information were sought per respondent (SD 2.11).

Regarding the reasons for seeking health information online, the majority of respondents cited noticing new symptoms or a change in health (70.08%, 712/1016), for knowledge or curiosity (51.57%, 524/1016), and deciding to change behaviors or daily routine (50.98%, 518/1016). Less than a quarter of respondents sought information to prepare for a doctor's consultation (23.62%, 240/1016); being prescribed a new medication, test, or treatment (21.65%, 220/1016); being diagnosed a new medical condition (19.98%, 203/1016); or having doubts about information given by doctor (11.42%, 116/1016).

For the source of health information, most respondents consulted online encyclopedias (eg, Wikipedia; 69.98%, 711/1016), health portals or medical encyclopedias (eg, MIMS, MedlinePlus, WebMD; 41.83%, 425/1016), or Q&A sites (eg, Yahoo! Answers, Baidu Knows; 40.85%, 415/1016). Official websites were less consulted: 36.61% (372/1016) visited hospital or clinic websites, 34.55% (351/1016) visited government websites, 27.85% (283/1016) visited university websites, and 26.18% (266/1016) consulted nonprofit organization websites. The mean number of sources sought per respondent was 4.22 (SD 2.35).

Convenience was the top reason (55.41%, 563/1016) for choosing a particular website for health information, followed by easy to understand (51.97%, 528/1016) and top results from search engines (41.14%, 418/1016); only a minority of respondents cited recommendations from health care professionals (15.26%, 155/1016) or family and friends (9.35%, 95/1016) as the reason for choosing a website.

XSL•FO
RenderX

Table 3. Contents, sources, and reasons for online health information seeking (n=1016).

Question	n (%)
Online health information sought[a]	
Symptom	829 (81.59)
Disease or condition	716 (70.47)
Medication	581 (57.19)
Service info	470 (46.26)
Healthy behaviors	452 (44.49)
Treatment and procedure	367 (36.12)
Tests and investigations	288 (28.35)
Vitamins and supplements	282 (27.76)
Chinese medicine	211 (20.77)
Alternative medicine	117 (11.52)
Health insurance	111 (10.93)
Reason for seeking health information online[a]	
Noticing new symptoms or change in health	712 (70.08)
For knowledge or curiosity	524 (51.57)
Deciding to change behaviors or daily routine	518 (50.98)
Hearing or seeing something in the news wanted to learn more about	446 (43.90)
Finding or selecting a doctor or health facility	342 (33.66)
Dealing with an ongoing medical condition	243 (23.92)
Preparing for a doctor's consultation	240 (23.62)
Being prescribed with a new medication, test, or treatment	220 (21.65)
Being diagnosed with a new medical condition	203 (19.98)
Having doubts about information given by doctor	116 (11.42)
Source of online health information[a]	
Online encyclopedia (eg, Wikipedia)	711 (69.98)
Health portal and medical encyclopedia (eg, MIMS, MedlinePlus, WebMD)	425 (41.83)
Q&A sites (eg, Yahoo! Answers, Baidu Knows)	415 (40.85)
Hospital or clinic	372 (36.61)
Government	351 (34.55)
News sites	317 (31.20)
Internet forums and message boards	295 (29.04)
University	283 (27.85)
Nonprofit organization	266 (26.18)
Social media (eg, Facebook, Twitter)	259 (25.49)
Video-sharing sites (eg, YouTube)	248 (24.41)
Commercial sites	159 (15.65)
Blogs	155 (15.26)
Reason for choosing the source[a]	
Convenience	563 (55.41)
Easy to understand	528 (51.97)
Top results from search engines	418 (41.14)

XSL•FO
RenderX

Question	n (%)
I think it's trustworthy	323 (31.79)
Usual habit	303 (29.82)
Recommended by professionals	155 (15.26)
Recommended by family or friends	95 (9.35)

[a]Multiple responses allowed.

Predictors to Online Health Information Seeking

Prior to regression analyses, missing values analysis was performed. Overall, only 0.422% of items were missing from the dataset. Missing data were shown to be missing completely at random (MCAR) as evidenced by a nonsignificant Little MCAR test (P=.72). Single imputation of the missing values was performed using the expectation-maximization algorithm.

A Poisson regression using a logarithmic link function was performed to explore various demographic and behavioral factors that impact the extent of online health information seeking (Table 4). The number of types of health information sought (online health information-seeking info score) was chosen to represent the extent of online health information seeking as supported by previous research [10]. Poisson regression was used because the online health information-seeking info score is a count variable. After adjusting for other covariates, only fair or poor health status, having a chronic medical condition, using the internet several times a day, and a higher eHEALS score were significant positive predictors of online health information seeking (df=992, log likelihood=–2109.024).

Predictors of eHealth Literacy

A multiple regression analysis was performed to explore the various demographic and behavioral factors that impact eHealth literacy (Table 5). The regression model was statistically significant and accounted for 11.5% of eHealth literacy (R^2=.126, adjusted R^2=.115, $F_{12,1003}$=12.038, P<.001). After adjusting for other covariates, more frequent internet use, more frequent online health information seeking, and more types of health information sought were significant positive predictors of eHealth literacy, whereas age and poorer health status were significant negative predictors.

XSL•FO
RenderX

Table 4. Summary of Poisson regression analysis for variables predicting online health information seeking (n=1016).

Variable	Regression coefficient (SE)	Wald χ^2_1	*P* value	AOR[a] (95% CI)
(Intercept)	0.557 (0.187)			
Age (years)				
16-24 (Ref[b])				
25-34	0.006 (0.057)	0.013	.91	1.006 (0.899-1.126)
35-44	0.063 (0.076)	0.686	.41	1.065 (0.917-1.237)
45-54	0.022 (0.081)	0.074	.79	1.022 (0.872-1.199)
≥55	0.041 (0.087)	0.227	.63	1.042 (0.879-1.236)
Gender				
Male (Ref)				
Female	0.045 (0.033)	1.913	.17	1.046 (0.981-1.116)
Education level				
Secondary and below (Ref)				
Tertiary	0.060 (0.065)	0.874	.35	1.062 (0.936-1.205)
Postgraduate	0.064 (0.070)	0.836	.36	1.066 (0.930-1.222)
Occupation				
Managers, executives, and officials (Ref)				
Professionals and academic staff	0.106 (0.077)	1.923	.17	1.112 (0.957-1.292)
Technicians and associate professionals	–0.048 (0.092)	0.275	.60	0.953 (0.795-1.142)
Clerical and office workers	–0.013 (0.084)	0.026	.87	0.987 (0.837-1.163)
Other workers or not currently employed	–0.150 (0.101)	2.212	.14	0.861 (0.707-1.049)
Student	–0.011 (0.089)	0.016	.90	0.989 (0.831-1.177)
Self-rated health status				
Excellent, very good, or good (Ref)				
Fair or poor	0.151 (0.035)	18.510	<.001	1.163 (1.086-1.247)
Have chronic medical condition	0.088 (0.040)	4.790	.03	1.092 (1.009-1.181)
Frequency of using the Internet				
Several times a week or less (Ref)				
Every day	0.172 (0.101)	2.903	.09	1.188 (0.974-1.448)
Several times a day	0.208 (0.099)	4.399	.04	1.231 (1.014-1.496)
Time using the internet per day (hours)				
<1 (Ref)				
1 to <2	0.020 (0.084)	0.059	.81	1.021 (0.866-1.204)
2 to <3	0.020 (0.085)	0.056	.81	1.020 (0.863-1.206)
3 to <4	0.092 (0.088)	1.094	.30	1.097 (0.922-1.304)
≥4	0.083 (0.084)	0.981	.32	1.087 (0.922-1.282)
Uses wearable health-monitoring devices	0.037 (0.040)	0.872	.35	1.038 (0.960-1.121)
eHEALS[c] score	0.018 (0.004)	25.796	<.001	1.018 (1.011-1.025)

[a]AOR: adjusted odds ratio.

[b]Ref: reference group.

[c]eHEALS: eHealth Literacy Scale.

XSL•FO
RenderX

Table 5. Summary of multiple regression analysis for variables predicting eHealth literacy (n=1016).

Variable	Unstandardized regression coefficient (SE)	Standardized regression coefficient	t_{1003}	*P* value
(Intercept)	27.188 (1.564)			
Age	–0.484 (0.149)	–0.145	–3.249	<.001
Gender (male=0)	–0.450 (0.286)	–0.047	–1.574	.12
Education level	–0.011 (0.246)	–0.002	–0.046	.96
Occupation	–0.184 (0.111)	–0.070	–1.656	.10
Self-rated health status (excellent to poor)	–0.918 (0.187)	–0.153	–4.919	<.001
Have chronic medical condition	0.645 (0.374)	0.056	1.726	.09
Frequency of internet use	0.849 (0.286)	0.096	2.967	.003
Hours of internet use per day	–0.008 (0.116)	–0.002	–0.070	.94
Use wearable health-monitoring devices	0.393 (0.362)	0.033	1.087	.28
Frequency of online health information seeking	0.493 (0.093)	0.170	5.328	<.001
Number of types of health information sought	0.230 (0.081)	0.106	2.831	.01
Number of sources of health information sought	0.038 (0.073)	0.019	0.521	.60

Discussion

Principal Findings

Results from this study suggest that among the study population, consisting mostly of university students and staff, internet use is extensive and online health information seeking is highly prevalent. For those who did seek health information online, most searched at least once every few months, for themselves as well as for family and friends. The mobile phone was the most used device to access the internet in general as well as for seeking online health information, which implies a desire for instant information. Rates of online health information seeking are much more prevalent compared to a local study done a decade ago [16], but our figures are comparable to more recent studies done overseas [4,10,25].

In contrast with the high prevalence of online health information seeking, fewer respondents asked their doctors about the health information they found online. This could be due to a perceived lack of interest from the doctors; indeed, for most of those who did ask, their doctors were apparently not interested. This could have led to a vicious cycle of poor patient-doctor communication.

A large variety of health information was sought for different reasons, but the vast majority of respondents searched for symptoms and diseases, for preconsultation self-diagnosis, and to decide whether to consult a doctor or not. Approximately half of the respondents sought information about healthy behaviors and were deciding on changing their daily habits, which indicates health consciousness. Few respondents searched for information postconsultation or when given a new diagnosis, medication, or treatment. Even fewer searched because of doubts about doctors' information—this might be explained by a strong patient-doctor trust or less desire to challenge the doctors' authority.

Compared to the previous local study [16], fewer respondents sought health information at official websites such as hospitals, government, universities, or nonprofit organizations. Coupled with the finding that the top reasons for choosing a website were convenience and that it was easy to understand, this suggests that the respondents were less concerned about the accuracy or quality of health information online, or unaware of the possible consequences of receiving inaccurate health information online.

Fair or poor health status and having a chronic medical condition were found to be predictors of online health information seeking, which implies that online health information seeking is need-based. The eHEALS score was also found to be a significant predictor, which is consistent with previous studies [11,14], suggesting that eHealth literacy serves as an enabler to online health information seeking. In contrast with previous studies [15,25], age, gender, education, and occupation were not shown to be predictors to online health information seeking. Although there was no significant generation gap in the extent of online health information seeking, older age was associated with lower eHealth literacy.

It is also worth noting that although patients with poorer health status tended to search for health information online more often, they also had lower eHealth literacy—this group of patients could be more vulnerable to unreliable online health information and the risk to health it entails.

Relevance to Clinical Practice

According to McMullan [26], doctors may respond in three different ways toward patients' online health information-seeking behaviors: (1) feel threatened and respond defensively, (2) collaborate with patients to obtain and analyze the information, and (3) guide patients to reliable health websites (ie, internet prescription). It is obvious that a defensive response is unfavorable to patient-doctor rapport; indeed, a recent study has shown that failure to acknowledge and understand patients' online health information-seeking behaviors will become a

XSL•FO
RenderX

barrier to patient-doctor communication [27]. Therefore, for better rapport and patient-doctor communication, primary care doctors should facilitate and empower patients to search for health information online correctly. Thus, doctors should also be educated about the variety of and be able to assess the quality of online health information so they can better teach their patients.

Doctors, public health professionals, and health care organizations should work together to offer high-quality health information online, to improve the ease of use and readability of health care websites, and to educate patients and the general public on the importance of assessing the quality of online health information.

Strengths and Limitations

This is the first study exploring primary care patients' online health information-seeking behavior in Hong Kong. It provides a comprehensive and up-to-date quantitative picture of online health information seeking for primary care doctors to understand their patients' health information needs, which they might not disclose to doctors.

Although the study sample was representative of the whole clinic population, it is skewed toward younger and more educated patients. Thus, the study's external validity is reduced. Moreover, self-reported health status and chronic medical conditions were used in this study, as compared to physician-reported health status (eg, by chart review), which could obtain a more accurate and detailed measure of respondents' health status.

Conclusions

Online health information seeking is prevalent among primary care patients in Hong Kong. Most searched for information preconsultation, but only a minority shared the information with doctors. Websites were chosen more for convenience than for accuracy or authoritativeness. Doctors should recognize patients' online health information-seeking behavior, and facilitate and empower them to search for high-quality online health information.

Acknowledgments

We thank Dr Winne Ko for assistance with survey instrument validation and statistical analyses, as well as all University Health Service staff for assistance on questionnaire distribution and collection throughout the study period.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire (English version).

[PDF File (Adobe PDF File), 618KB - jmir_v21i3e10831_app1.pdf]

Multimedia Appendix 2

Questionnaire (Chinese version).

[PDF File (Adobe PDF File), 706KB - jmir_v21i3e10831_app2.pdf]

References

1. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer use of "Dr Google": a survey on health information-seeking behaviors and navigational needs. J Med Internet Res 2015 Dec 29;17(12):e288 [FREE Full text] [doi: 10.2196/jmir.4345] [Medline: 26715363]
2. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
3. 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]
4. Song H, Omori K, Kim J, Tenzek KE, Morey HJ, Lin W, et al. Trusting social media as a source of health information: online surveys comparing the United States, Korea, and Hong Kong. J Med Internet Res 2016 Mar 14;18(3):e25 [FREE Full text] [doi: 10.2196/jmir.4193] [Medline: 26976273]
5. Anderson JG, Rainey MR, Eysenbach G. The impact of CyberHealthcare on the physician-patient relationship. J Med Syst 2003 Feb;27(1):67-84. [Medline: 12617199]
6. Akerkar SM, Bichile LS. Doctor patient relationship: changing dynamics in the information age. J Postgrad Med 2004;50(2):120-122 [FREE Full text] [Medline: 15235209]

XSL•FO
RenderX

7. Caiata-Zufferey M, Abraham A, Sommerhalder K, Schulz PJ. Online health information seeking in the context of the medical consultation in Switzerland. Qual Health Res 2010 Aug;20(8):1050-1061. [doi: 10.1177/1049732310368404] [Medline: 20442347]
8. Moreland J, French T, Cumming G. Exploring online health information seeking in Scotland. 2015 Presented at: The 3rd European Workshop on Practical Aspects of Health Informatics; Oct 27, 2015; Elgin, Scotland URL: http://ceur-ws.org/Vol-1574/paper8.pdf [WebCite Cache ID 6yoJG7ymR]
9. Attfield SJ, Adams A, Blandford A. Patient information needs: pre- and post-consultation. Health Informatics J 2006 Jun;12(2):165-177. [doi: 10.1177/1460458206063811] [Medline: 17023406]
10. Moreland J, French TL, Cumming GP. The prevalence of online health information seeking among patients in Scotland: a cross-sectional exploratory study. JMIR Res Protoc 2015;4(3):e85 [FREE Full text] [doi: 10.2196/resprot.4010] [Medline: 26177562]
11. 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]
12. Li J, Theng Y, Foo S. Predictors of online health information seeking behavior: changes between 2002 and 2012. Health Informatics J 2015 Aug 10;22(4):804-814. [doi: 10.1177/1460458215595851] [Medline: 26261218]
13. Oh YS, Cho Y. Examining the relationships between resources and online health information seeking among patients with chronic diseases and healthy people. Soc Work Health Care 2015;54(2):83-100. [doi: 10.1080/00981389.2014.987940] [Medline: 25674723]
14. Li N, Orrange S, Kravitz RL, Bell RA. Reasons for and predictors of patients' online health information seeking following a medical appointment. Fam Pract 2014 Oct;31(5):550-556. [doi: 10.1093/fampra/cmu034] [Medline: 24963151]
15. Li F, Li M, Guan P, Ma S, Cui L. Mapping publication trends and identifying hot spots of research on Internet health information seeking behavior: a quantitative and co-word biclustering analysis. J Med Internet Res 2015 Mar 25;17(3):e81 [FREE Full text] [doi: 10.2196/jmir.3326] [Medline: 25830358]
16. Yan YY. Online health information seeking behavior in Hong Kong: an exploratory study. J Med Syst 2010 Apr;34(2):147-153. [Medline: 20433053]
17. Census and Statistics Department, Hong Kong Special Administrative Region. Hong Kong as an information society URL: https://www.censtatd.gov.hk/hkstat/sub/sp120.jsp?productCode=B1110006 [accessed 2018-10-29] [WebCite Cache ID 73WJAWVd1]
18. Dean A, Sullivan K, Soe M. OpenEpi. 2013. Open source epidemiologic statistics for public health URL: http://www.openepi.com/ [accessed 2018-04-20] [WebCite Cache ID 6yoKZeKB3]
19. Lam ET, Lam CL, Fong DY, Huang WW. Is the SF-12 version 2 Health Survey a valid and equivalent substitute for the SF-36 version 2 Health Survey for the Chinese? J Eval Clin Pract 2013 Feb;19(1):200-208. [doi: 10.1111/j.1365-2753.2011.01800.x] [Medline: 22128754]
20. 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]
21. Nguyen J, Moorhouse M, Curbow B, Christie J, Walsh-Childers K, Islam S. Construct validity of the eHealth Literacy Scale (eHEALS) among two adult populations: a Rasch analysis. JMIR Public Health Surveill 2016 May;2(1):e24 [FREE Full text] [doi: 10.2196/publichealth.4967] [Medline: 27244771]
22. 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]
23. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006 Oct;29(5):489-497. [doi: 10.1002/nur.20147] [Medline: 16977646]
24. World Health Organization. 2017. Process of translation and adaptation of instruments URL: http://www.who.int/substance_abuse/research_tools/translation/en/ [WebCite Cache ID 6yoKlE1of]
25. 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]
26. 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]
27. 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;17(1):e19 [FREE Full text] [doi: 10.2196/jmir.3588] [Medline: 25586865]

Abbreviations

CVI: content validity index
eHEALS: eHealth Literacy Scale
eHealth: electronic health

XSL•FO
RenderX

MCAR: missing completely at random

Edited by G Eysenbach; submitted 23.04.18; peer-reviewed by G Cumming, JR Bautista, Z Ma, Y Yang; comments to author 28.08.18; revised version received 30.10.18; accepted 27.01.19; published 27.03.19.

Please cite as:
Wong DKK, 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;21(3):e10831
URL: http://www.jmir.org/2019/3/e10831/
doi:10.2196/10831
PMID:30916666

©David Ka-Ki Wong, Man-Kuen Cheung. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Genomic Common Data Model for Seamless Interoperation of Biomedical Data in Clinical Practice: Retrospective Study

Seo Jeong Shin[1*], MS; Seng Chan You[2*], MS, MD; Yu Rang Park[3*], PhD; Jin Roh[4], MD, PhD; Jang-Hee Kim[4], MD; Seokjin Haam[5], MD; Christian G Reich[6], MD, PhD; Clair Blacketer[7], MPH; Dae-Soon Son[8], PhD; Seungbin Oh[9], PharmD; Rae Woong Park[1,2], MD, PhD

[1]Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Republic of Korea
[2]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea
[3]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea
[4]Department of Pathology, Ajou University Hospital, Suwon, Republic of Korea
[5]Department of Thoracic & Cardiovascular Surgery, Ajou University Hospital, Suwon, Republic of Korea
[6]IQVIA, Durham, NC, United States
[7]Department of Epidemiology, Janssen Research and Development, Titusville, NJ, United States
[8]Samsung Genome Institute, Samsung Medical Center, Seoul, Republic of Korea
[9]Department of Pharmacy, Kangwon University, Chuncheon, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Rae Woong Park, MD, PhD
Department of Biomedical Informatics
Ajou University School of Medicine
206 World cup-ro, Yeongtong-gu
Suwon, 16499
Republic of Korea
Phone: 82 312194471
Fax: 82 312194472
Email: veritas@ajou.ac.kr

Abstract

Background: Clinical sequencing data should be shared in order to achieve the sufficient scale and diversity required to provide strong evidence for improving patient care. A distributed research network allows researchers to share this evidence rather than the patient-level data across centers, thereby avoiding privacy issues. The Observational Medical Outcomes Partnership (OMOP) common data model (CDM) used in distributed research networks has low coverage of sequencing data and does not reflect the latest trends of precision medicine.

Objective: The aim of this study was to develop and evaluate the feasibility of a genomic CDM (G-CDM), as an extension of the OMOP-CDM, for application of genomic data in clinical practice.

Methods: Existing genomic data models and sequencing reports were reviewed to extend the OMOP-CDM to cover genomic data. The Human Genome Organisation Gene Nomenclature Committee and Human Genome Variation Society nomenclature were adopted to standardize the terminology in the model. Sequencing data of 114 and 1060 patients with lung cancer were obtained from the Ajou University School of Medicine database of Ajou University Hospital and The Cancer Genome Atlas, respectively, which were transformed to a format appropriate for the G-CDM. The data were compared with respect to gene name, variant type, and actionable mutations.

Results: The G-CDM was extended into four tables linked to tables of the OMOP-CDM. Upon comparison with The Cancer Genome Atlas data, a clinically actionable mutation, p.Leu858Arg, in the *EGFR* gene was 6.64 times more frequent in the Ajou University School of Medicine database, while the p.Gly12Xaa mutation in the *KRAS* gene was 2.02 times more frequent in The Cancer Genome Atlas dataset. The data-exploring tool GeneProfiler was further developed to conduct descriptive analyses automatically using the G-CDM, which provides the proportions of genes, variant types, and actionable mutations. GeneProfiler also allows for querying the specific gene name and Human Genome Variation Society nomenclature to calculate the proportion of patients with a given mutation.

XSL•FO
RenderX

Conclusions: We developed the G-CDM for effective integration of genomic data with standardized clinical data, allowing for data sharing across institutes. The feasibility of the G-CDM was validated by assessing the differences in data characteristics between two different genomic databases through the proposed data-exploring tool GeneProfiler. The G-CDM may facilitate analyses of interoperating clinical and genomic datasets across multiple institutions, minimizing privacy issues and enabling researchers to better understand the characteristics of patients and promote personalized medicine in clinical practice.

(J Med Internet Res 2019;21(3):e13249) doi:10.2196/13249

KEYWORDS

high-throughput nucleotide sequencing; databases, genetic; multicenter study; patient privacy; data visualization

Introduction

Background

Recognition of the importance of clinical next-generation sequencing (NGS) in precision medicine has had a profound impact on improving medical care [1-3]. Patients' sequencing data are currently generated through relatively large-scale projects aimed at exploring the role of clinical NGS in precision medicine conducted by organizations such as the American Association for Cancer Research Project GENIE [4] and the China Precision Medicine Initiative [5]. However, genomic data are considered to be privacy sensitive and potentially reidentifiable, which raises concerns about transmitting and sharing patient-level data outside of host institutions for collaborative research [6]. In addition, genomic sequencing data of subjects in a predefined cohort cannot reflect the full diversity of the entire population at the point of care, which limits the practical application of the data for research purposes [7].

There has been a recent widespread effort to collect genomic information on patients in clinical practice through routine laboratory tests by the UK Biobank [8] and Geisinger Health System [9]. Since March 2017, the South Korea government has provided conditional insurance for an NGS technology-based cancer gene panel [10], which is expected to lead to rapid accumulation of clinical sequencing data in each hospital. However, the vocabulary and structure of these datasets are not standardized, which makes it difficult to conduct appropriate multicenter or comparative analyses for clinical decision making [11]. This lack of standardization can be overcome by using the common data model (CDM), which applies the same data structure to run an identical analysis code for each data holder [12]. For example, the Informatics for Integrating Biology and the Bedside is a clinical data warehouse platform comprising genetic data that adopts the CDM to support the distributed research network [13,14], an infrastructure for novel internet-based strategies that allows researchers to use retrospective multicenter data in a CDM (in contrast to single-center or cloud-based research) without exporting the protected personal health information. Researchers can combine the results of an analysis code run over the network to generate a refined clinical hypothesis [12,15]. To date, the distributed research network has been adopted by global research collaboration groups, including the Observational Health Data Sciences and Informatics (OHDSI) consortium [16]. The Observational Medical Outcomes Partnership (OMOP) CDM, now in version 6.0, was developed by the OHDSI consortium and includes clinical data from over 20 countries, with information of 1.5 billion patients transformed to date.

Prior Work

Due to the nature and extraordinary complexity of sequencing data, it is challenging to effectively describe and interpret the status of sequence alterations [17]. Furthermore, sequencing data were applied in the clinical domain of NGS relatively later than other types of genomic tests; hence, the analytical process has not been standardized [18]. To improve the efficiency of data processing, sequencing data should be managed using standardized structures and semantics. Although several standard models for genomic data have been introduced to date, they have limited applicability. For example, the standard for non-NGS–specific data models, including the minimum information about a microarray experiment [19] for DNA microarray analysis, the tissue microarray object model [20] for tissue microarray analysis, and the proteomics experiment data repository [21] for proteomics, cannot be properly adopted for sequencing data. Although the minimum information about a high-throughput nucleotide sequencing experiment was developed as a data model specific for sequencing data, it requires experimental processing data and detailed analytical protocols to enable researchers to reproduce the analysis [22].

Aim

Given the limitations outlined above, the objective of this study was to create a genomic data CDM (G-CDM) for use in the distributed research network. To address patient privacy issues and support the diversity of genomic data such as ethnicity, the OMOP-CDM used in the OHDSI consortium was chosen for this study for expansion. Furthermore, we validated the feasibility of the model by exploring the difference in genomic data retrieved from public databases and clinical practice.

Methods

Construction of the Genomic Data Common Data Model

The proposed G-CDM was developed by extending the OMOP-CDM to achieve the seamless management of clinical sequencing data through a structured database model. Clinical information such as basic patient background (eg, sex and age), clinical diagnosis, procedures, or specimen type was stored in existing tables of the OMOP-CDM. We further reviewed other genomic data models and clinical sequencing reports to design additional tables for describing and interpreting sequence alterations occurring in target genes. There are various types

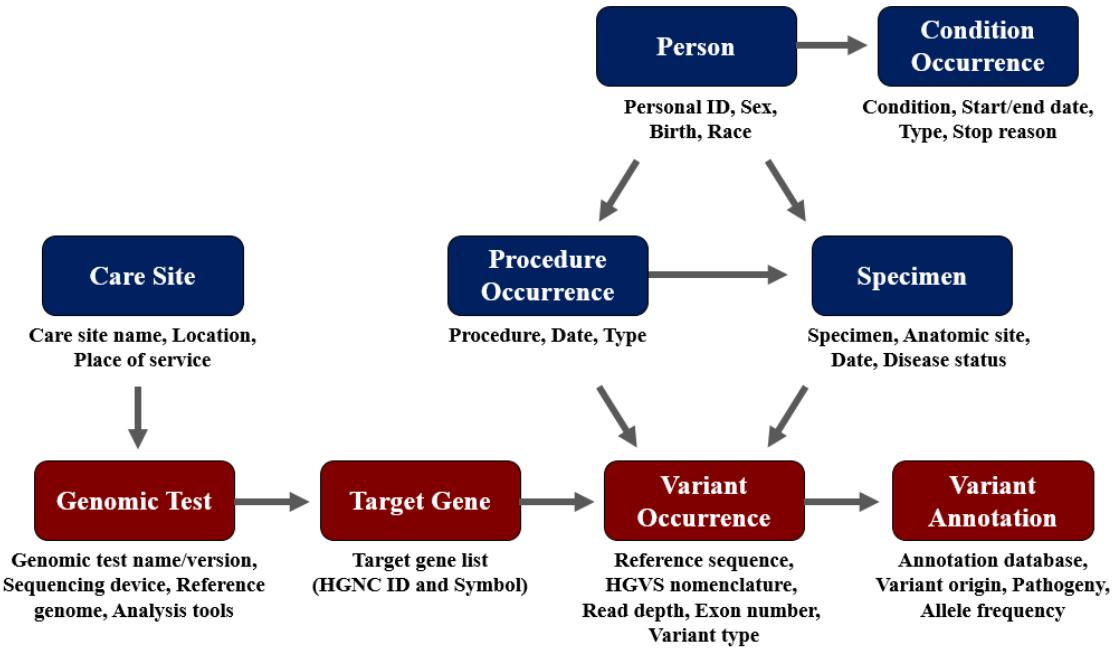
XSL•FO
RenderX

(>50) of public cancer databases describing variants, including comprehensive cancer projects, resources, and cancer type-specific databases [23]. According to our inclusion and exclusion criteria (Multimedia Appendix 1), we selected datasets from The Cancer Genome Atlas (TCGA), Catalogue of Somatic Mutations in Cancer, and International Cancer Genome Consortium for review and reference, to define the method of sequence alteration description. The data quality of these representative databases has been validated through many studies and papers. The database TCGA provides large-scale datasets of genomic alterations, including insertions/deletions (INDELs) or single nucleotide polymorphisms (SNPs), discovered in over 30 human tumor types to generate comprehensive profiles of cancer genomics [24]. The database Catalogue of Somatic Mutations in Cancer provides somatic mutations across 1,391,372 tumor samples encompassing 5,977,977 coding mutations as of August 2018 [25], while the database International Cancer Genome Consortium provides the datasets of oncogenic mutations of 50 different cancer types to support large-scale studies [26,27]. We excluded the databases built based on non-NGS techniques or cancer type–specific databases from referencing. The ISO20428 document, which is a standard format for reporting sequencing results, was reviewed to design columns for variant annotation (Multimedia Appendix 2). To guarantee interoperability of the data, standard terminologies were adopted in the G-CDM [28,29]. The name of a human gene, a key factor in sequencing data, was fixed according to the nomenclature of the Human Genome Organisation Gene Nomenclature Committee, which currently contains and maintains approximately 41,000 unique gene symbols. In addition, the Human Genome Variation Society nomenclature was adopted to standardize the manner of describing sequence alterations in each gene at both the DNA and protein level. Although either one- or three-letter abbreviations are permitted in the Human Genome Variation Society nomenclature, we propose expressing the amino acid by its three-letter code only to permit seamless data analysis for widespread research (Multimedia Appendix 2).

Data Structure of the Genomic Data Common Data Model

To link clinical data in the OMOP-CDM, the following information on each patient with NGS data was stored in a separate corresponding table: Person, Condition_Occurrence (diagnosis), Procedure_Occurrence, Specimen, and Care_Site (Figure 1). The Person table included personal patient information such as individual identification, sex, age, and race. The Condition_Occurrence table contained information on the patient's condition or diagnosis, including the disease such as "lung cancer" or condition type such as "primary condition." The Procedure_Occurrence table included information on how the specimen used for NGS was obtained and the name of the genomic test conducted for a patient. The Specimen table included information on the specimen used for the genomic test, such as "target" (tumor tissue) and "reference" (normal tissue), along with specimen type, including paraffin-embedded slide, the date the specimen was obtained, and the anatomical site of the specimen. The Care_Site table included information on the site at which the genomic test was conducted.

Figure 1. Schematic diagram of the relationship between tables composing the genomic common data model. Tables in red ("Genomic_Test," "Target_Gene," "Variant_Occurrence," and "Variant_Annotation") are those storing genomic sequencing data and processes, whereas tables in blue ("Person," "Condition_Occurrence," "Procedure_Occurrence," "Specimen," and "Care_Site") are those already existing in the Observational Medical Outcomes Partnership-common data model and store clinical data directly linked to the "Variant_Occurrence" and "Genomic_Test" tables. ID: identification; HGVS: Human Genome Variation Society; HGNC: Human Genome Organisation Gene Nomenclature Committee.



XSL•FO
RenderX

In addition to these five tables, we expanded the model to be linked to four other tables containing information related to the sequencing data: (1) the Genomic_Test table included the test name, version, sequencing device, analytical tools, and reference databases, with a care site identification column; (2) the Target_Gene table contained a list of genes targeted by the genomic test following Human Genome Organisation Gene Nomenclature Committee nomenclature for standardized gene symbols; (3) the Variant_Occurrence table included descriptive information about the variants of target genes; and (4) the Variant_Annotation table included information on each variant and the clinical interpretation thereof, such as annotation database name, variant origin such as somatic or germline, pathogenicity of the variant, allele frequency, and medication.

Procedure identification for conducting sequencing, specimen identification of both the target and reference specimens, and target gene identification were included as foreign keys to link the information in the Procedure, Specimen, and Target_Gene tables. Data on reference sequence, reference SNP identification, Human Genome Variation Society nomenclature at both the DNA and protein levels, read depth, exon number, and variant type of both structural DNA and functional proteins were stored as variant description parameters. Detailed schemes and descriptions of each column and table used in the genomic extension model are provided in Multimedia Appendices 3 and 4.

Data Description

The Ajou University School of Medicine (AUSOM) database consists of electronic medical record data of patients who underwent NGS-based cancer panel screening of the tumor tissue between June 2017 and August 2018 at Ajou University Hospital, including 92 patients with lung adenocarcinoma and 22 patients with lung squamous cell carcinoma. Public sequence alteration data of the lung cancer cohort Pan-Lung Cancer study of TCGA [30] were obtained from the Memorial Sloan-Kettering Cancer Center cBioPortal [31].

The overall processes of NGS conducted at Ajou University Hospital and the TCGA database are detailed in Multimedia Appendix 5. Two representative differences between the sequencing pipelines of the two databases are the number of genes and the composition of variant types targeted in the test. For example, in the cancer panel of AUSOM, 49 cancer-related genes were targeted for sequencing, while the TCGA data were harvested using whole-exome sequencing with 16,896 genes. Thus, for development and testing of the proposed G-CDM, we selected 1060 patients from TCGA with available variant data of the 49 target genes selected in the AUSOM panel (Table 1).

Table 1. Description of data used to build the genomic common data model and to validate the data model.

Variable	AUSOM[a] (N=114), n (%)	TCGA[b] (N=1060), n (%)
Age (years)		
≤49	7 (6.1)	44 (4.2)
50-59	26 (22.8)	163 (15.4)
60-69	41 (36.0)	310 (29.2)
70-79	35 (30.7)	317 (29.9)
≥80	5 (4.4)	56 (5.2)
Unknown	0 (0.0)	170 (16.0)
Gender		
Male	64 (56.1)	628 (59.0)
Female	50 (43.9)	429 (41.0)
Unknown	0 (0.0)	3 (0.2)
Pathology		
Lung adenocarcinoma	92 (80.7)	603 (56.9)
Lung squamous carcinoma	22 (19.3)	457 (43.1)
Cancer stage		
Stage I	78 (68.4)	526 (49.6)
Stage II	16 (14.0)	286 (27.0)
Stage III	18 (15.8)	184 (17.4)
Stage IV	0 (0.0)	36 (3.4)
Unknown	2 (1.8)	28 (2.6)

[a]AUSOM: Ajou University School of Medicine.

[b]TCGA: The Cancer Genome Atlas.

XSL•FO
RenderX

The variant types, including SNPs, INDELs, multinucleotide polymorphisms (MNPs), copy number variants (CNVs), and translocations, were explored in the AUSOM database, whereas only SNPs and INDELs were identified in the TCGA database. Information on clinical characteristics such as age, sex, and disease status and genomic alterations such as variant type, DNA and protein level changes, and functional impact were used to compare the AUSOM and TCGA databases.

Study Design

Sequencing data of the TCGA database, which was licensed by Yonsei University for use, and of the AUSOM database were transformed into the G-CDM at Yonsei University and Ajou University, respectively. To execute the transformation process, the Structured Query Language (SQL) script in Microsoft SQL Server 2017 was used as the relational database backend for storage and querying the sequencing data. The G-CDM database was built using the Intel Xeon CPU E5-2596 v4 2.20 GHz, Java v.1.8.0, R v.3.5.1, and DBMS SQL Server 2017 at Ajou University, while the Intel Xeon Gold 6132 CPU 2.60 GHz, Java v.1.8.0, R v.3.4.4, and DBMS SQL Server 2017 were used at Yonsei University.

After extracting parameters of interest for a cohort of patients by using a Condition_Occurrence table, the genetic information of the patients was summarized in each of the two institutions. Owing to the restrictions on exporting the original clinical sequencing data in the AUSOM database outside the hospital, the two institutions gathered and compared only the descriptive statistical analysis results to compare the two sequencing databases in further research.

The data visualization tool "GeneProfiler" was developed to run based on the G-CDM as a demonstration that the standardized structure and vocabulary system can serve as a usable medium for performing distributed research by allowing genomic analysis with an identical code. To validate the feasibility of the G-CDM as a storage system and analysis medium, the differences in sequencing data between the AUSOM and TCGA databases were explored. The background profile of variants was described based on several aspects such as gene names, variant types, and disease subtypes. Representative actionable mutations for patients with non-small cell lung cancer (NSCLC) tend to occur in the *EGFR*, *KRAS*, *PIK3CA*, *BRAF*, and *NRAS* genes according to National Comprehensive Cancer Network guidelines [32,33]. Therefore, the proportions of actionable mutations in these five genes were compared between the two databases and between the subtypes of lung cancer.

Data Visualization Tool

We developed a new data visualizing tool called "GeneProfiler" using the R Shiny package to facilitate the utility and accessibility of the G-CDM. After converting genomic data into the G-CDM, the data can be visualized by simply connecting the database with the graphic user interface (Figure 2). As users link their database into "GeneProfiler," this tool automatically provides the descriptive statistics as several plots and tables. "GeneProfiler" includes action buttons to generate plots of overall variant profiles, proportion of certain mutation types, and proportion of genes with actionable mutations. Users can also freely explore the proportion of patients with mutations in specific genes or specific variants and can download the results as a plot or table to conduct distributed research. After downloading result tables of several databases from GeneProfiler, users can generate graphs comparing these databases by uploading the merged tables (Multimedia Appendix 6). The R Shiny code of "GeneProfiler" was uploaded and is open to the public in GitHub [34].

Statistical Analysis

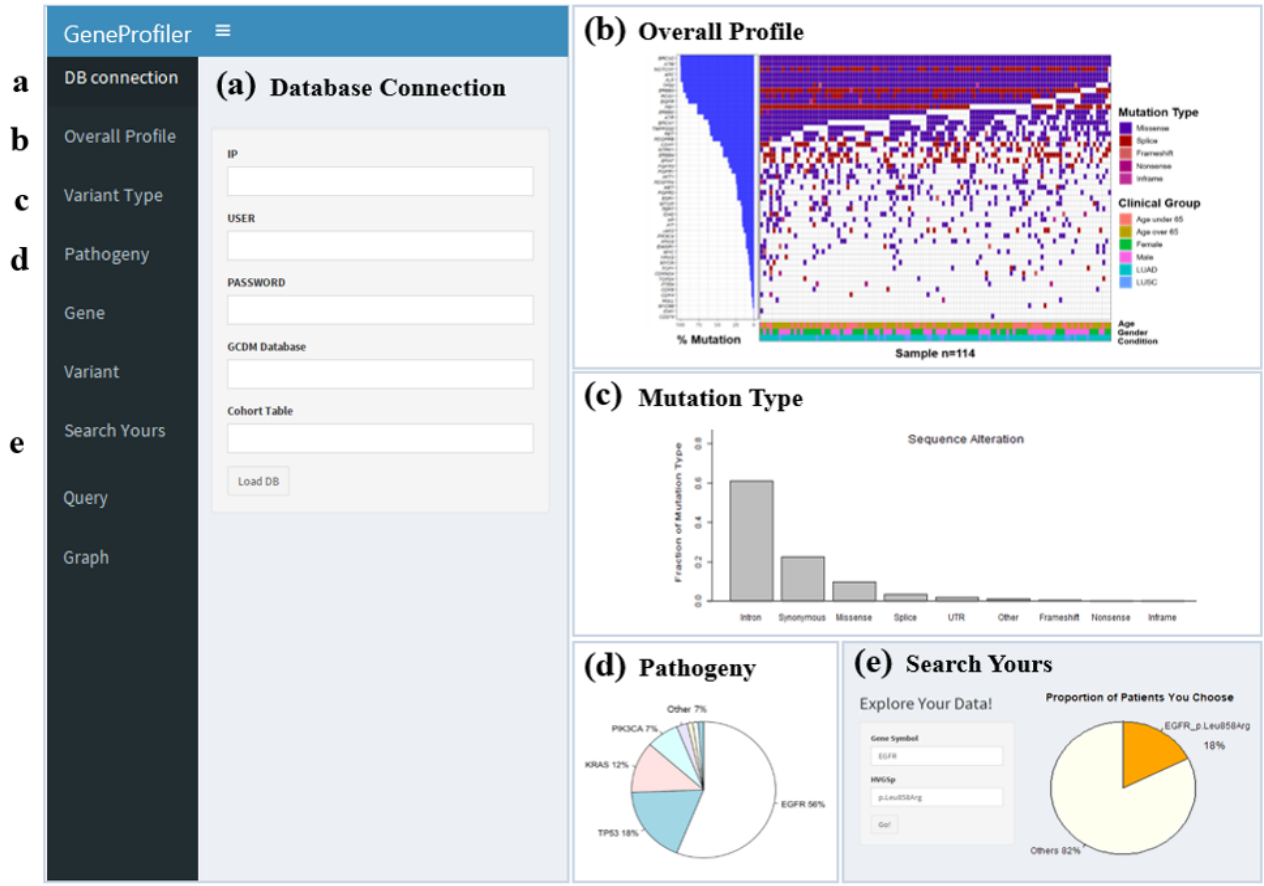
Descriptive analysis was performed using frequencies for categorical variables. Genomic characteristics were compared between the two databases using a chi-squared test, and values of $P<.05$ were considered statistically significant. The R program version 3.5.1 was used for data preprocessing and statistical analysis. A mutation waterfall plot was created using "GenVisR," an R package available via Bioconductor [35], which also provided the proportions of genes, variant types, and specific variants using the R Shiny tool developed in this study.

Ethics Statement

This study was approved by the institutional review board at Ajou University Hospital of Korea (IRB approval number: AJIRB-MED-MDB-18-390).

XSL•FO
RenderX

Figure 2. Data visualization tool for clinical sequencing data holders who converted their genomic data into genomic CDM. Users can (a) connect their genomic CDM database; (b) get analysis plots such as an overall profile, (c) mutation type, and (d) pathogeny of variants; and (e) search the proportion of patients with gene name and variant information. CDM: common data model.



Results

Data Comparison for Model Validation

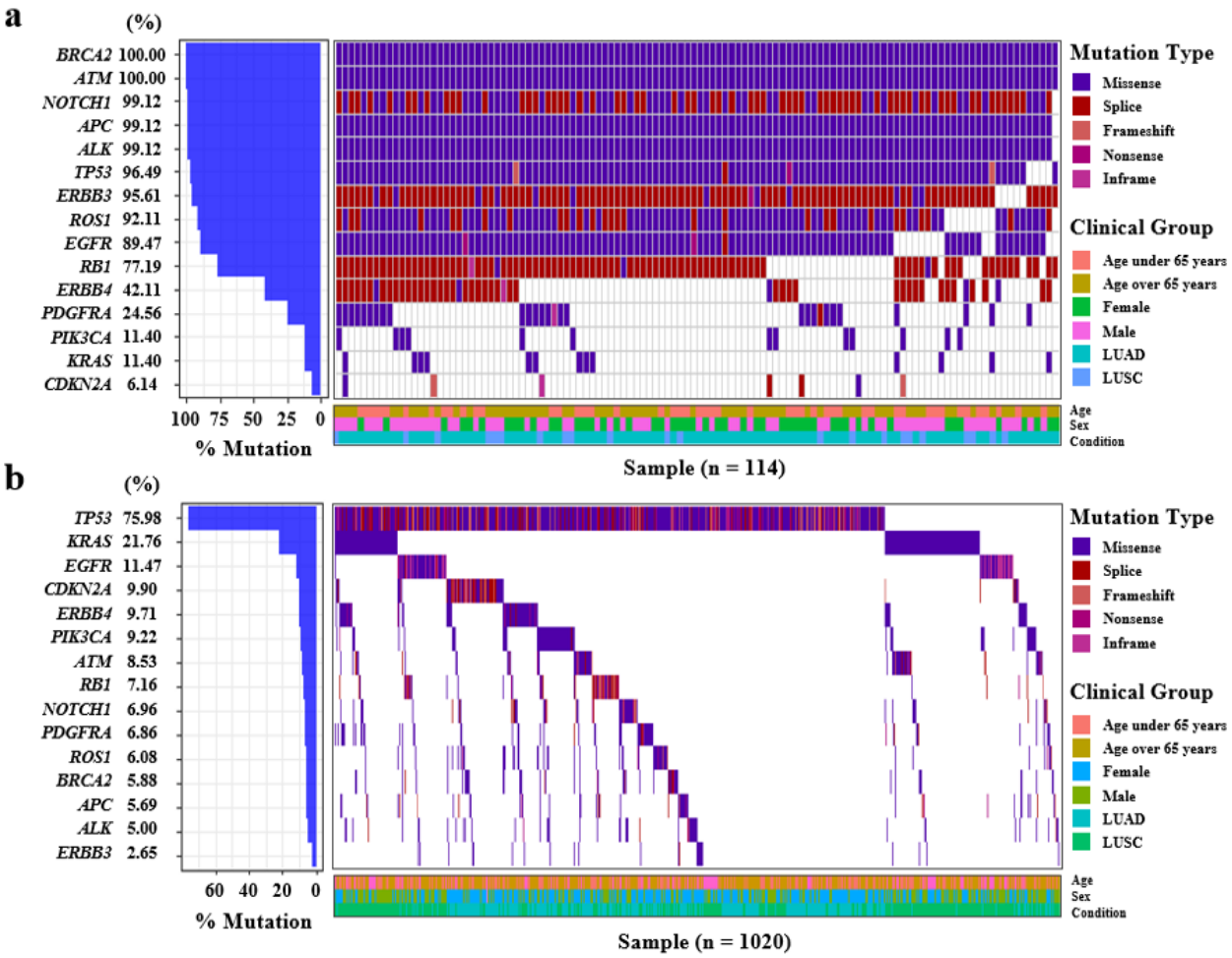
To confirm the differences between the AUSOM and TCGA databases, the summary results of the sequencing data such as the gene, variant type, and disease subtypes were gathered and compared. We characterized the biological background of total variants in both databases for variant types, with DNA-level structural variants classified as “sequence alteration” and protein functional types classified as “variant feature.” Among the SNPs, insertions, and deletions, the most frequent structural variant type was SNPs, accounting for >80% of total variants in both databases (Multimedia Appendix 7). However, the functional types of the variants, including missense, nonsense, frameshift, inframe, and splice variants, showed different frequencies between the databases (all P<.001), with intron and synonymous variants being most frequent in the AUSOM database (combined frequency of 83%) and missense variants being the most frequent in the TCGA database (73%; Multimedia Appendix 7).

A waterfall plot was created in both the AUSOM and TCGA databases, which focused only on protein-altering variants such as missense, nonsense, frameshift, inframe, and splicing variants to obtain a variant profile (Figure 3; Multimedia Appendix 8). The 15 genes as a union of the top 10 genes in each database were selected as targets for overall profiling. In the AUSOM database, the top 10 genes had a variant frequency > 75% among patients with lung cancer, whereas only one gene, *TP53*, had a variant frequency > 25% in the TCGA database. In particular, *EGFR* variants showed very different frequencies in the AUSOM and TCGA databases (89.5% and 11.5%, respectively). All 15 genes had different proportions of variants in the two databases (all P<.001). Although the ranking of genes with high frequencies of variants differed between databases; the most frequent variant type was a missense variant in both databases (Figure 3).

In contrast, comparison of the waterfall plot of all 49 genes targeted in the cancer panel of the AUSOM database to that of the same gene set of the TCGA database showed a higher frequency of frameshift and nonsense type variants than splice type variants in the TCGA data, although the ranking of genes with more variants still differed between the two databases (Multimedia Appendix 8). Exploration of the CNVs in AUSOM showed that *RET* was the gene with the most frequent CNVs, specifically due to amplification (Multimedia Appendix 8).

XSL•FO
RenderX

Figure 3. Waterfall plot describing the variant profile of the top 10 genes in (a) Ajou University School of Medicine and (b) The Cancer Genome Atlas databases. Each row represents gene symbols ordered by their frequency of variants with different colors indicating different variant types. Columns represent each patient with only one sample per patient. The bar graph on the left corresponds to the frequency of variants in each gene. Clinical groups such as age, sex, and condition are shown in the bottom box. LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma.

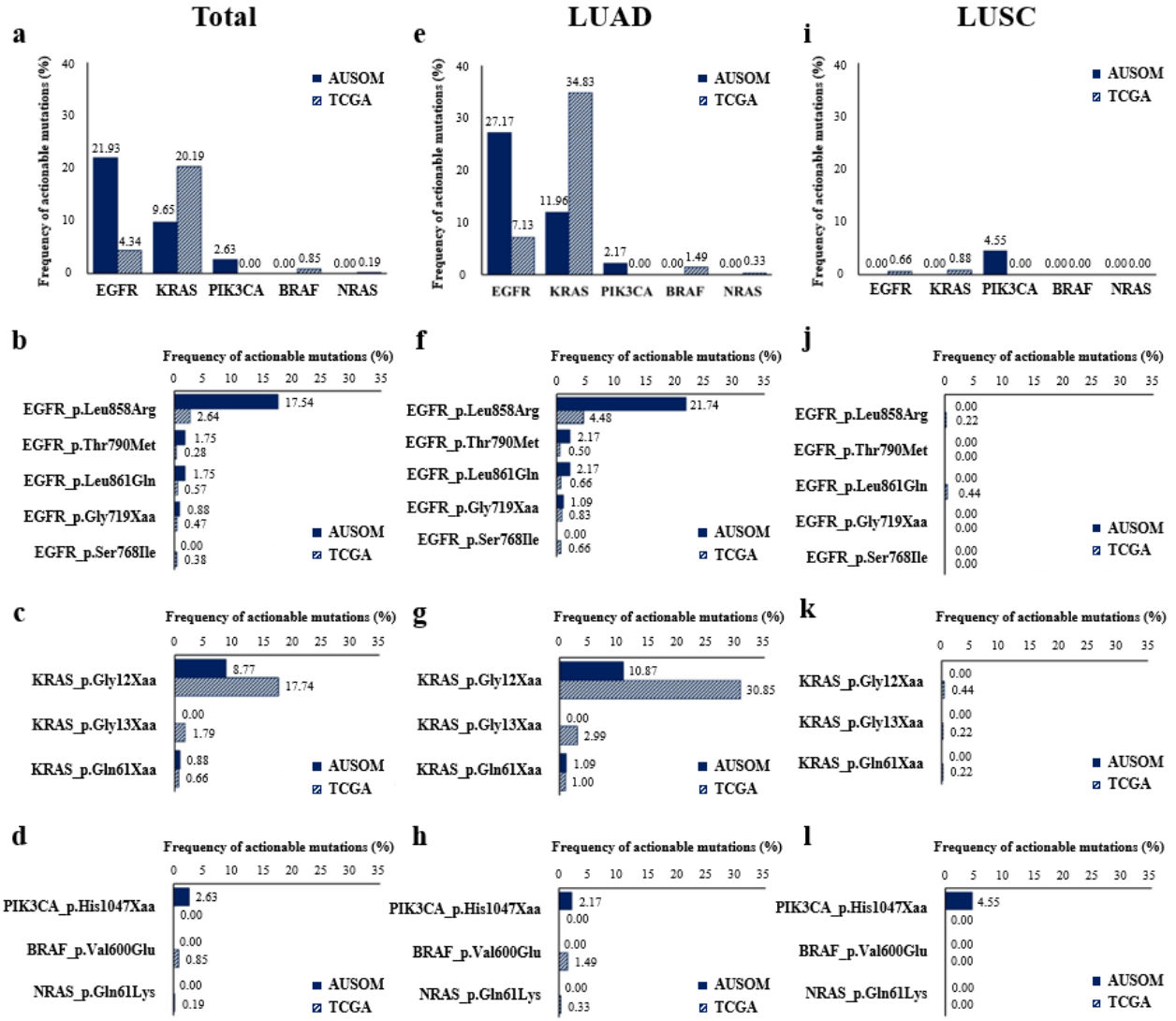


Comparison of Actionable Mutations for Model Validation

An actionable mutation is a specific genomic event that potentially affects a patient's response to a targeted therapy [36]. Of the five representative actionable mutations for NSCLC examined (*EGFR*, *KRAS*, *PIK3CA*, *BRAF*, and *NRAS*), *EGFR* showed the greatest frequency of variants in the AUSOM database (21.9%), while *KRAS* showed the greatest frequency of variants in the TCGA database (20.2%; Figure 4a). In particular, the point mutation p.Leu858Arg in *EGFR* was found in 17.5% of the patients, followed by p.Thr790Met (1.8%) in the AUSOM database (Figure 4b). Point mutations in the *KRAS* gene, such as p.Gly12Xaa and p.Gly13Xaa, were more frequent in the TCGA database (20.2%) than in the AUSOM database (9.7%; Figure 4a,c). In addition, patients with lung adenocarcinoma (Figure 4e-h) tended to have more actionable mutations than those with lung squamous cell carcinoma (Figure 4i-l).

XSL•FO
RenderX

Figure 4. Frequencies of actionable mutations detected in the sequencing process between the AUSOM and TCGA databases. Frequency is shown according to the (a, e, i) level of five selected genes and (b, f, j) actionable mutations in EGFR, (c, g, k) KRAS, and (d, h, l) others such as PIK3CA, BRAF, and NRAS. Frequency is also shown according to patient groups: (a-d) total, (e-h) lung adenocarcinoma, and (k-l) lung squamous cell carcinoma. AUSOM: Ajou University School of Medicine; TCGA: The Cancer Genome Atlas; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma.



Discussion

Overview

We developed a new data model for clinical sequencing data, which was applied using sequencing data of patients with lung cancer from two different databases, AUSOM and TCGA, which were transformed into an identical format for the G-CDM. To evaluate the feasibility of the G-CDM, the composition of the datasets was compared with regard to the frequency of a gene name and variant types in which a sequence alteration occurred and to the prevalence of actionable mutations. Moreover, we developed novel user-friendly software—GeneProfiler—for visualization of clinical sequencing data.

Interpretation of the Principal Results

The first result obtained by comparison of the databases transformed in a standardized form for the G-CDM was the clear difference in the composition of the sequencing data between TCGA, a controlled research-oriented database, and AUSOM, an actual clinical practice database. This difference suggested a difference in variant frequencies and types between the two databases. Indeed, the total number of variants per patient was much higher for the AUSOM database than for the TCGA database, whereas the frequency of variants differed according to the variant type considered. Comparison of actionable mutations in five genes of NSCLC showed a much higher mutation frequency of *EGFR* in the AUSOM database (a cohort of Asian patients) than in the TCGA database (a cohort of American patients). This finding is in line with previous knowledge that Asian patients with NSCLC have a higher prevalence of *EGFR* mutations than Americans [32,37]. In contrast, actionable mutations in the *KRAS* gene were less prevalent in patients in the AUSOM database than in those in the TCGA database, which is also consistent with previous knowledge that Asian populations have a much lower rate of mutations in *KRAS* than non-Asian populations with NSCLC [32,37].

XSL•FO
RenderX

The second key result of this study is conduct of a multicenter research through internet-based sharing of analysis codes with CDM-based conversion of databases from different institutions. This is meaningful because the distributed research was conducted with genomic data that had not been previously verified. Such distributed research would be a useful strategy to address the problem of limited data integration due to privacy issues of clinical sequencing data.

Moreover, because data from the TCGA database were generated relatively earlier than those in the AUSOM database, the sequencing equipment or bioinformatics method may have caused the observed differences. These differences between the databases further emphasize the importance of analyzing data obtained from multiple clinical sites together with research-driven public data to obtain a higher level of representative evidence from diverse populations. Both genomic data models and intermediate results should be shared as widely as possible to promote clinical advances by overcoming the current challenges of unstructured and siloed data environments that lead to a lack of interoperability [38]. Our proposed OMOP-CDM extension model was developed by referencing the OHDSI distributed research network, because existing models such as the HL7 reference information model are not suitable for internet-based research and have limited practical use [39,40].

In the process of modeling the structure of the G-CDM, two specimen identifications were allocated in the Variant_Occurrence table, because recent methods of NGS testing in cancer patients tend to be based on a comparison of normal and tumor tissues simultaneously from the same individual. In cases of patients with a congenital disease, there is an option to fill out this field with only single-specimen identification. The contents of annotation to a variant can also differ according to the type or version of the annotation databases used in the annotation process. For this reason, the Variant_Annotation table was separated from the Variant_Occurrence table to allow for subsequent updating of diverse or new interpretations.

Limitations

Genomic data are generated using highly complicated sequencing pipelines and analytical processes; consequently, NGS data have inherent limitations in terms of data quality and reliability. Although we compared the sequencing pipelines and analytical processes used to generate the sequencing data of both the AUSOM and TCGA databases, we were unable to confirm the detailed parameters and options used in each process. Thus, the differences between the two databases found in this study should be interpreted considering the possibility that the data may have been generated by dissimilar methods and criteria.

Moreover, the clinical NGS data used in this study were generated in the clinical practice of Ajou University Hospital within the last 2 years. Given the recent time frame, mortality was rare among these patients; thus, we were not able to perform survival analysis by leveraging both genomic data and clinical data.

The G-CDM, as a common data structure and vocabulary system, minimizes privacy issues when conducting multicenter studies by integrating statistical results of the same analysis code rather than sharing the clinical sequencing data directly. However, when the G-CDM is used for repeated queries with a malicious purpose, there is concern for compromising the privacy of the individual, even if the queries target only the aggregated statistics. The G-CDM can be complemented by inhibiting reidentification attacks, as proposed in previous studies related to the mitigation of privacy risks, through limiting response to a query targeting a unique individual or through introduction of noise into the original data [41,42].

Conclusions

We propose the distributed research network–based G-CDM as a starting point for a broad community discussion on genomic data–based precision medicine. Based on the G-CDM developed in this study, the data validation process identified differences between the clinical NGS data derived from a clinical practice and those derived from prospective research. We believe that the construction and adoption of this standard data model will facilitate the usefulness of clinical NGS data.

Acknowledgments

This work was supported by the Bio Industrial Strategic Technology Development Program (20001234) funded by the Ministry of Trade, Industry & Energy (Republic of Korea); a grant from the Korea Health Technology R&D Project through the Korean Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI16C0992); and by the Technology Innovation Program (20002289), funded by the Ministry of Trade, Industry & Energy (Republic of Korea).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Inclusion and exclusion criteria for databases used for review and reference to define the method of sequence alteration description.

[XLSX File (Microsoft Excel File), 9KB - jmir_v21i3e13249_app1.xlsx]

XSL•FO
RenderX

Multimedia Appendix 2

Architecture of the genomic common data model (G-CDM). Conceptualized description of the genomic models, databases, and nomenclature referenced during the G-CDM development process.

[PNG File, 159KB - jmir_v21i3e13249_app2.png]

Multimedia Appendix 3

Genomic common data model (G-CDM) entity-relationship diagram as an extension of the Observational Medical Outcomes Partnership (OMOP)-CDM. Tables with genomic data (red) and clinical data (blue) are linked. Not all columns composing each table are shown for clarity and conciseness.

[PNG File, 222KB - jmir_v21i3e13249_app3.PNG]

Multimedia Appendix 4

Description of the genomic common data model (G-CDM) specifications as an extension of the Observational Medical Outcomes Partnership (OMOP)-CDM.

[PDF File (Adobe PDF File), 292KB - jmir_v21i3e13249_app4.pdf]

Multimedia Appendix 5

Next-generation sequencing pipelines used at Ajou University School of Medicine (AUSOM) and for the data collected in The Cancer Genome Atlas (TCGA).

[PNG File, 115KB - jmir_v21i3e13249_app5.PNG]

Multimedia Appendix 6

GeneProfiler for conducting distributed research. (a) Using the "Query" tab, an identical analysis query can be submitted to datasets of different institutions. (b) Using the "Graph" tab, a comparative graph can be generated by inputting a merged table containing the analysis results of several institutions.

[PNG File, 318KB - jmir_v21i3e13249_app6.png]

Multimedia Appendix 7

Characterization of the biological background of lung cancer, including lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC). The fractions of (a) structural mutations and (b) functional mutations are shown according to mutation type between the Ajou University School of Medicine (AUSOM) and The Cancer Genome Atlas (TCGA) datasets.

[PNG File, 111KB - jmir_v21i3e13249_app7.png]

Multimedia Appendix 8

Overall mutation profile of lung cancer patients for (a-b) total targeted genes in the Ajou University School of Medicine (AUSOM) and The Cancer Genome Atlas (TCGA) databases and (c) copy number variations in AUSOM.

[PNG File, 656KB - jmir_v21i3e13249_app8.PNG]

References

1. Schwaederle M, Parker BA, Schwab RB, Daniels GA, Piccioni DE, Kesari S, et al. Precision Oncology: The UC San Diego Moores Cancer Center PREDICT Experience. Mol Cancer Ther 2016 Dec;15(4):743-752 [FREE Full text] [doi: 10.1158/1535-7163.MCT-15-0795] [Medline: 26873727]
2. Kris MG, Johnson BE, Berry LD, Kwiatkowski DJ, Iafrate AJ, Wistuba II, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. JAMA 2014 May 21;311(19):1998-2006 [FREE Full text] [doi: 10.1001/jama.2014.3741] [Medline: 24846037]
3. Stockley TL, Oza AM, Berman HK, Leighl NB, Knox JJ, Shepherd FA, et al. Molecular profiling of advanced solid tumors and patient outcomes with genotype-matched clinical trials: the Princess Margaret IMPACT/COMPACT trial. Genome Med 2016 Dec 25;8(1):109 [FREE Full text] [doi: 10.1186/s13073-016-0364-2] [Medline: 27782854]
4. AACR Project GENIE Consortium. AACR Project GENIE: Powering Precision Medicine through an International Consortium. Cancer Discov 2017 Dec;7(8):818-831 [FREE Full text] [doi: 10.1158/2159-8290.CD-17-0151] [Medline: 28572459]

XSL•FO
RenderX

5. Cyranoski D. China embraces precision medicine on a massive scale. Nature 2016 Jan 07;529(7584):9-10. [doi: 10.1038/529009a] [Medline: 26738574]
6. Wang S, Jiang X, Tang H, Wang X, Bu D, Carey K, et al. A community effort to protect genomic data sharing, collaboration and outsourcing. NPJ Genom Med 2017;2:33 [FREE Full text] [doi: 10.1038/s41525-017-0036-1] [Medline: 29263842]
7. Scollen S, Page A, Wilson J. From the data on many, precision medicine for "one": The case for widespread genomic data sharing. Biomed Hub 2017 Nov 21;2(1):21-21. [doi: 10.1159/000481682]
8. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015 Mar;12(3):e1001779 [FREE Full text] [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
9. Murray MF. The path to routine genomic screening in health care. Ann Intern Med 2018 Sep 18;169(6):407-408. [doi: 10.7326/M18-1722] [Medline: 30073309]
10. Park YR, Shin S. Status and direction of healthcare data in Korea for artificial intelligence. Hanyang Med Rev 2017;37(2):86. [doi: 10.7599/hmr.2017.37.2.86]
11. Nam DJ, Kwon HW, Lee H, Ahn EK. National healthcare service and its big data analytics. Healthc Inform Res 2018 Jul;24(3):247-249 [FREE Full text] [doi: 10.4258/hir.2018.24.3.247] [Medline: 30109158]
12. Park RW. Sharing clinical big data while protecting confidentiality and security: Observational health data sciences and informatics. Healthc Inform Res 2017 Jan;23(1):1-3 [FREE Full text] [doi: 10.4258/hir.2017.23.1.1] [Medline: 28261525]
13. Murphy SN, Mendis ME, Berkowitz DA, Kohane I, Chueh HC. Integration of clinical and genetic data in the i2b2 architecture. 2006 Nov 11 Presented at: AMIA Annu Symp Proc; 2006; Washington, DC p. 1040.
14. Klann JG, Abend A, Raghavan VA, Mandl KD, Murphy SN. Data interchange using i2b2. J Am Med Inform Assoc 2016 Dec;23(5):909-915 [FREE Full text] [doi: 10.1093/jamia/ocv188] [Medline: 26911824]
15. Kim KK, Browe DK, Logan HC, Holm R, Hack L, Ohno-Machado L. Data governance requirements for distributed clinical research networks: triangulating perspectives of diverse stakeholders. J Am Med Inform Assoc 2014;21(4):714-719 [FREE Full text] [doi: 10.1136/amiajnl-2013-002308] [Medline: 24302285]
16. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Stud Health Technol Inform 2015;216:574-578 [FREE Full text] [Medline: 26262116]
17. Green R, Rehm H, Kohane I. In: Ginsburg H, Willard H, editors. Clinical genome sequencing. San Diego: Genomic and personalized medicine; 2013.
18. Ibrahim MM, Lacadie SA, Ohler U. JAMM: a peak finder for joint analysis of NGS replicates. Bioinformatics 2015 Jan 01;31(1):48-55. [doi: 10.1093/bioinformatics/btu568] [Medline: 25223640]
19. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2001 Dec;29(4):365-371. [doi: 10.1038/ng1201-365] [Medline: 11726920]
20. Lee HW, Park YR, Sim J, Park RW, Kim WH, Kim JH. The tissue microarray object model: a data model for storage, analysis, and exchange of tissue microarray experimental data. Arch Pathol Lab Med 2006 Jul;130(7):1004-1013. [doi: 10.1043/1543-2165(2006)130[1004:TTMOMA]2.0.CO;2] [Medline: 16831026]
21. Garwood K, McLaughlin T, Garwood C, Joens S, Morrison N, Taylor CF, et al. PEDRo: a database for storing, searching and disseminating experimental proteomics data. BMC Genomics 2004 Sep 17;5:68 [FREE Full text] [doi: 10.1186/1471-2164-5-68] [Medline: 15377392]
22. Kahl G. Minimum information about a high-throughput nucleotide sequencing experiment (MINSEQE). In: The dictionary of genomics, transcriptomics and proteomics. Hoboken, NJ: Wiley-Blackwell; 2015.
23. Pavlopoulou A, Spandidos DA, Michalopoulos I. Human cancer databases (review). Oncol Rep 2015 Jan;33(1):3-18 [FREE Full text] [doi: 10.3892/or.2014.3579] [Medline: 25369839]
24. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015;19(1A):A68-A77 [FREE Full text] [doi: 10.5114/wo.2014.47136] [Medline: 25691825]
25. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res 2019 Jan 08;47(D1):D941-D947 [FREE Full text] [doi: 10.1093/nar/gky1015] [Medline: 30371878]
26. International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. Nature 2010 Apr 15;464(7291):993-998 [FREE Full text] [doi: 10.1038/nature08987] [Medline: 20393554]
27. Ewing AD, Houlahan KE, Hu Y, Ellrott K, Caloian C, Yamaguchi TN, ICGC-TCGA DREAM Somatic Mutation Calling Challenge participants, et al. Combining tumor genome simulation with crowdsourcing to benchmark somatic single-nucleotide-variant detection. Nat Methods 2015 Jul;12(7):623-630 [FREE Full text] [doi: 10.1038/nmeth.3407] [Medline: 25984700]
28. Olaronke I, Gambo I, Soriyan A, Olaleke J. Interoperability in healthcare: benefits, challenges and resolutions. Int J Innov Appl Stud 2013;38(1):64-73.

XSL•FO
RenderX

29. Bicer V, Laleci GB, Dogac A, Kabak Y. Artemis message exchange framework: semantic interoperability of exchanged messages in the healthcare domain. SIGMOD Rec 2005 Sep 01;34(3):71. [doi: 10.1145/1084805.1084819]
30. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, Cancer Genome Atlas Research Network, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet 2016 Dec;48(6):607-616. [doi: 10.1038/ng.3564] [Medline: 27158780]
31. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012 May;2(5):401-404 [FREE Full text] [doi: 10.1158/2159-8290.CD-12-0095] [Medline: 22588877]
32. Liu L, Liu J, Shao D, Deng Q, Tang H, Liu Z, et al. Comprehensive genomic profiling of lung cancer using a validated panel to explore therapeutic targets in East Asian patients. Cancer Sci 2017 Dec;108(12):2487-2494 [FREE Full text] [doi: 10.1111/cas.13410] [Medline: 28949084]
33. National Comprehensive Cancer Network. 2019. NCCN Guidelines URL: https://www.nccn.org/professionals/physician_gls/default.aspx [accessed 2019-02-08] [WebCite Cache ID 761zJMg3W]
34. Shin SJ, You SC, Kim JH. GitHub. 2019. GeneProfiler URL: https://github.com/ABMI/GeneProfiler [accessed 2019-02-08] [WebCite Cache ID 761y9H7gJ]
35. Skidmore ZL, Wagner AH, Lesurf R, Campbell KM, Kunisaki J, Griffith OL, et al. GenVisR: Genomic visualizations in R. Bioinformatics 2016 Dec 01;32(19):3012-3014 [FREE Full text] [doi: 10.1093/bioinformatics/btw325] [Medline: 27288499]
36. Carr TH, McEwen R, Dougherty B, Johnson JH, Dry JR, Lai Z, et al. Defining actionable mutations for oncology therapeutic development. Nat Rev Cancer 2016 Dec 26;16(5):319-329. [doi: 10.1038/nrc.2016.35] [Medline: 27112209]
37. Arrieta O, Cardona AF, Martín C, Más-López L, Corrales-Rodríguez L, Bramuglia G, et al. Updated frequency of EGFR and KRAS mutations in nonsmall-cell lung cancer in Latin America: The Latin-American consortium for the investigation of lung cancer (CLICaP). J Thorac Oncol 2015 May;10(5):838-843 [FREE Full text] [doi: 10.1097/JTO.0000000000000481] [Medline: 25634006]
38. Tang C, Plasek JM, Bates DW. Rethinking data sharing at the dawn of a health data economy: A viewpoint. J Med Internet Res 2018 Nov 22;20(11):e11519 [FREE Full text] [doi: 10.2196/11519] [Medline: 30467103]
39. Schadow G, Mead CN, Walker DM. The HL7 reference information model under scrutiny. Stud Health Technol Inform 2006;124:151-156. [Medline: 17108519]
40. Beeler GW. HL7 version 3--an object-oriented methodology for collaborative standards development. Int J Med Inform 1998 Feb;48(1-3):151-161. [Medline: 9600415]
41. Shringarpure SS, Bustamante CD. Privacy risks from genomic data-sharing beacons. Am J Hum Genet 2015 Nov 05;97(5):631-646 [FREE Full text] [doi: 10.1016/j.ajhg.2015.09.010] [Medline: 26522470]
42. Raisaro JL, Tramèr F, Ji Z, Bu D, Zhao Y, Carey K, et al. Addressing Beacon re-identification attacks: quantification and mitigation of privacy risks. J Am Med Inform Assoc 2017 Jul 01;24(4):799-805 [FREE Full text] [doi: 10.1093/jamia/ocw167] [Medline: 28339683]

Abbreviations

AUSOM: Ajou University School of Medicine
CDM: common data model
CNV: copy number variant
G-CDM: genomic common data model
INDEL: insertion/deletion
NGS: next-generation sequencing
OHDSI: Observational Health Data Sciences and Informatics
OMOP: Observational Medical Outcomes Partnership
SNP: single nucleotide polymorphism
SQL: structured query language
TCGA: The Cancer Genome Atlas

XSL•FO
RenderX

Edited by P Zhang; submitted 31.12.18; peer-reviewed by S Wang; comments to author 21.01.19; revised version received 11.02.19; accepted 25.02.19; published 26.03.19.

Please cite as:
Shin SJ, You SC, Park YR, Roh J, Kim JH, Haam S, Reich CG, Blacketer C, Son DS, Oh S, Park RW
Genomic Common Data Model for Seamless Interoperation of Biomedical Data in Clinical Practice: Retrospective Study
J Med Internet Res 2019;21(3):e13249
URL: http://www.jmir.org/2019/3/e13249/
doi:10.2196/13249
PMID:30912749

©Seo Jeong Shin, Seng Chan You, Yu Rang Park, Jin Roh, Jang-Hee Kim, Seokjin Haam, Christian G Reich, Clair Blacketer, Dae-Soon Son, Seungbin Oh, Rae Woong Park. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effect of Online Effort and Reputation of Physicians on Patients' Choice: 3-Wave Data Analysis of China's Good Doctor Website

Zhaohua Deng[1], PhD; Ziying Hong[1], MA; Wei Zhang[1], PhD; Richard Evans[2], PhD; Yanyan Chen[3], PhD

[1]Smart Health Institute, School of Medicine and Health Management, Huazhong University of Science and Technology, Wuhan, China

[2]College of Engineering, Design and Physical Sciences, Brunel University, London, United Kingdom

[3]Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Corresponding Author:
Wei Zhang, PhD
Smart Health Institute
School of Medicine and Health Management
Huazhong University of Science and Technology
No. 13 Hangkong Road, Qiaokou District
Wuhan, 430030
China
Phone: 86 13397110378
Email: weizhanghust@hust.edu.cn

Abstract

Background: Nowadays, patients are seeking physician information more frequently via the internet. Physician-rating websites (PRWs) have been recognized as the most convenient way to gain insight and detailed information about specific physicians before receiving consultation. However, little is known about how the information provided on PRWs may affect patients' decisions to seek medical advice.

Objective: This study aimed to examine whether the physicians' online efforts and their reputation have a relationship with patients' choice of physician on PRWs.

Methods: A model, based on social exchange theory, was developed to analyze the factors associated with the number of online patients. A 3-wave data collection exercise, covering 4037 physicians on China's Good Doctor website, was conducted during the months of February, April, and June 2017. Increases in consultation in a 60-day period were used as the dependent variable, whereas 2 series of data were analyzed using linear regression modeling. The fixed-effect model was used to analyze the 3-wave data.

Results: The adjusted R^2 value in the linear regression models were 0.28 and 0.27, whereas in the fixed-effect model, it was .30. Both the linear regression and fixed-effect models yielded a good fit. A positive effect of physicians' effort on the aggregated number of online patients was identified in all models (R^2=0.30 and R^2=0.37 in 2 regression models; R^2=0.23 in fixed effect model; P<.001). The proxies of physicians' reputations indicated different results, with total number of page views of physicians' homepages (R^2=0.43 and R^2=0.46; R^2=0.16; P<.001) and number of votes received (R^2=0.33 and R^2=0.27; R^2=0.43; P<.001) being seen as positive. Virtual gifts were not significant in all models, whereas *thank-you* messages were only significant in the fixed-effect model (R^2=0.11; P=.02). The effort made by physicians online is positively associated with their aggregated number of patients consulted, whereas the effect of a physician's reputation remains uncertain. The control effect of a physician's title and hospital's level was not significant in all linear regressions.

Conclusions: Both the effort and reputation of physicians online contribute to the increased number of online patients' consultation; however, the influence of a physician's reputation varies. This may imply that physicians' online effort and reputation are critical in attracting patients and that strategic manipulation of physician profiles is worthy of study. Practical insights are also discussed.

(J Med Internet Res 2019;21(3):e10170) doi:10.2196/10170

KEYWORDS

physician-rating websites; physician efforts; physician reputations; patient choices; panel data analysis

XSL•FO
RenderX

Introduction

Background

With the increasing popularity of Web 2.0 technologies, people are seeking health information more frequently online [1-6]. The internet has become a popular medium for obtaining medical treatment from physicians, for discussing and receiving medical advice, and for identifying symptoms experienced. It also offers many advantages for patients in comparison with the offline world, such as convenience, time saving, and reduced limitations on space and time. Since the early 2000s, consumers have been able to access the ratings of medical services and choose physicians through websites [2]. Physician-rating websites (PRWs) are an alternative and novel way for patients to obtain information about physicians before receiving consultation. PRWs collect and present information about patients' experiences and whether they were satisfied with their encounters with physicians [7]. They provide patients with an opportunity to discuss their health conditions with physicians and rate his or her quality of service and the care provided [8,9]. The number of PRWs throughout the world is increasing [3], such as in the United States (RateMDs.com and Healthgrades.com), the United Kingdom (iWantGreatCare.org), and Germany (Jameda.de and AOKGesundheitsnavi.de). In China, haodf.com [5] and guahao.com [10] are rising in popularity.

The development of PRWs has made it easier and more flexible for patients to access information and consult with physicians before receiving medical services. From a patient's perspective, most PRWs are perceived to be positive and can help improve the provision of services [11]. They can provide patients with information about physicians, especially in relation to service quality [11,12]. Globally, the number of people engaging and conversing on PRWs is increasing, with many using them to share their health care experiences, seek health care information, and rate the services received from health care practitioners [13,14]. The ratings on PRWs can also affect the patient care received from providers [15]. To make PRWs more reliable for the public, adjusting the content of PRWs to suit the differing information needs of health consumers is important [16].

People obtain information about the quality of physicians through the online view functionality built into PRWs. The information sought relates to the online efforts and reputation of physicians. Physician efforts imply the amount of time and energy that a physician spends online. A physician's reputation is associated with a patient's perception of the physician's characteristics or qualities in general, for example, honesty, capability, and reliability, and it is usually demonstrated in the form of positive evaluation toward the health care service received by the patient. On the basis of the evaluation of these 2 metrics, patients become better informed and make more critical decisions when choosing their physician. An increasing amount of research has tested the effects of online effort and reputation on the sales of products and decision-making by consumers [17-20]. However, little analysis has been conducted into the factors that contribute to the selection of physicians on PRWs. To fill this gap, this paper undertakes an investigation into the choice of physicians by patients on PRWs and aims to explore the factors that affect a patient's choice of a physician. The effects of physicians' effort and reputation online are also tested. Data for this research were collected from the Good Doctor website in China, one of the biggest PRWs in the country. As searching for suitable new physicians or evaluating their current one is deemed important for patients, this study provides insights into understanding how patients make their decisions online. By using the results of this study, marketers and the designers of PRWs can better tailor their online services to the needs of patients. The findings also offer practical insights for online health care providers to encourage patients' utilization of health care websites.

Physician Efforts

The term *effort* was defined by Naylor et al [21] as "the amount of energy 'spent' on an act per unit of time." In the sales and marketing domain, effort represents the amount of time and energy a salesperson devotes to the selling of a product or service, relative to another salesperson [22]. The effect of effort on performance has been considered by many researchers [22-24], with e ort being identified as a direct antecedent of performance in the research of salespeople [25]. As services are intangible, it is difficult to control their quality, compared with, for example, manufactured products. The verbal and nonverbal actions of employees have great influence on customers' perceptions toward value and service quality [26]. The effort of an employee is more important in service settings as customer evaluations of service quality are often linked directly to the performance of the service provider. Social exchange theory (SET) is commonly used to explain the exchange behavior between various parties, with the aim of individual behaviors being to maximize satisfaction and minimize costs [27]. This also applies to the relationship of physicians and patients [28]. From the perspective of social exchange, the physicians' participation in virtual communities is social exchange behavior, while patients can also provide social and economic returns for physicians as exchange returns [29].

An increase in visibility of effort from staff may generally lead to higher perceived quality by customers [26]. An employee's effort, applied to their daily work, can influence consumer perception of the services being received [30]. If the employee is considered to pay extra effort to their work, then he or she may obtain a higher rating from their customers. The employee's efforts will affect the research and purchasing intention of consumers, which is critical to a service organization's overall performance [17]. This positive effect of effort may contribute to the consumer's likelihood to browse and purchase goods in future. As online health care services are part of the service domain, a physician's efforts, applied to the services they deliver, can affect a patient's perception of quality of service and may change their choices and opinions toward physicians. The study by Liu et al [19] indicated that a physician's effort was a positive indicator of online physician popularity, that is, physicians become more popular when they show a greater effort toward their service. When selecting physicians through PRWs, patients can visit the home page of the physicians and obtain additional information, such as their personal blog,

XSL•FO
RenderX

published articles, and previous physician-patient communication. At this stage, patients gain a perception of the efforts made by the physician in the past, which may affect their attitude toward the physician and, thus, the likelihood of them selecting the physician. More effort shown by the physician online toward their service offerings may increase the chances of patients choosing them for consultation. Thus, we hypothesize the following:

H1: Patients prefer to consult with physicians that provide higher amounts of effort online.

Physician Reputation

The definition of reputation varies in different fields of research. In the online marketplace, reputation is understood as a conditional probability that an individual will behave in a certain manner [31]. Early research on reputation focused generally on experimentation [32] as it was difficult to measure reputation in the offline world. With the development of the internet and online user-centered social tools, the ability to measure reputation has developed, with much research being conducted into the development of methods to measure reputation on e-commerce and eHealth websites. As the reputation of a person and/or website can help consumers and vendors make better decisions, communicate more effectively, and improve cooperation, it is now seen to play a major role in online service delivery [33]. Reputation is acknowledged as one of the most influential factors that affect a consumer's behavior and seller's performance in online marketplaces [34]. A high reputation contributes to reducing information asymmetry and the reduction of risk and uncertainty perceived by consumers [35]. In the domain of SET, reputation is taken as an important factor that affects the behavior of online patients [36]. Many researchers have explored the effects of online reputation on sales in fields such as tourism, the retailing of books, and online auctions. Previous research has demonstrated that a correlation exists between reputation and sales, generally with online reputation having a positive effect on sales. For example, Dewan and Hsu [37] found that reputation has a significant effect on the sales of products on customer-to-customer auction websites, using data available on eBay. Similarly, Ye et al [20] indicated that seller reputation has a positive impact on sales volume, following analysis of sales data from Taobao.com, a Chinese e-commerce website.

The online reputation mechanism serves as the basis for online transactions, and it helps consumers obtain more detailed information about products before purchase. As physicians know more about their service quality and patients' health conditions than the patients themselves, information asymmetry can be considered severe in the online health care market [38]. Without PRWs, patients are unable to evaluate accurately the quality of a physician's services before consultation; this may lead to a misunderstanding or misinterpretation of information relating to the physician. Similarly, online consultation experiences can reduce the risks caused by information asymmetry and build trust between the patient and physician [39]. As a result, online reputation mechanisms can also be applied in the delivery of medical services. For example, health services provided online can allow patients to share their experience quickly and objectively [40]. There are few studies relating to the reputation of online health care services. The study by Josang [40] indicated that sound reputation systems can be applied to medical services, whereas word-of-mouth from family and friends of patients is considered important when selecting an appropriate physician [41]. When selecting Web services, reputation plays an important role in a patient's decision-making process [42], with reputation being treated as the most valuable attribute of a physician [43]. Reputation is also a vital quality factor in health care delivery as patients rely heavily on word-of-mouth when deciding which physician to approach [44]. The online reputation of physicians can help patients choose a suitable physician. If a physician is highly regarded, with a high reputation online, patients are more likely to consult with him or her; thus, we present the following hypothesis:

H2: Patients prefer to consult with physicians with a high level of online reputation.

Methods

Data Collection

Data were collected from the Good Doctor website (www.haodf.com) in China, one of the largest online PRWs in the country [5]. Currently, over 7500 hospitals and more than 500,000 physicians are active on the website. According to Good Doctor, online physicians are divided into 28 groups and more than 100 departments. As different medical departments provide varying treatment and patient-physician communication, we selected 1 department for our study to avoid interference from different departments. As China is facing the problem of an aging population [45], patients in the department of heart diseases represent a sizeable group; thus, we selected this department, which specializes in cardiovascular diseases. Using a Java-based program, we collected data from the homepages of 5996 cardiovascular physicians. Data were collected on February 25, April 27, and June 27, 2017, which allowed us to form a longitudinal panel dataset (there is a 60-day interval between these dates). We matched the URL and name information of physicians in 3 different stages. Following this process, it was revealed that 1361 physicians had data with less than 3 stages (among them, 698 physicians had 2 stages missing and 665 had 1 stage missing), 472 physicians had more than 1 null value across 3 stages, and 172 physicians had some abnormal values as the number of proxies decreased over time. The removal of samples with those missing values or abnormal values yielded a final dataset of 4037 patients from 878 specific hospitals. Approximately 46.54% (1879/4037) of the physicians were men, 22.94% (926/4037) were women, and 30.52% (1232/4037) could not be identified. The majority of physicians (3856/4037, 95.52%) are working in tertiary hospitals. Approximately 36.22% (1462/4037) of these physicians possess the title of Director. The details of the characteristics used in our study are listed in Table 1.

XSL•FO
RenderX

Table 1. Demographic characteristics of physicians. N=4037.

Physician characteristics	Statistics, n (%)
Gender	
Male	1879 (46.54)
Female	926 (22.94)
Not reported	1232 (30.52)
Hospital level	
Tertiary hospital	3856 (95.52)
Secondary hospital	174 (4.31)
One	7 (0.17)
Professional title	
Director Physician	1462 (36.22)
Associate Director Physician	1445 (35.79)
Attending Physician	917 (22.71)
Residing Physician	213 (5.28)

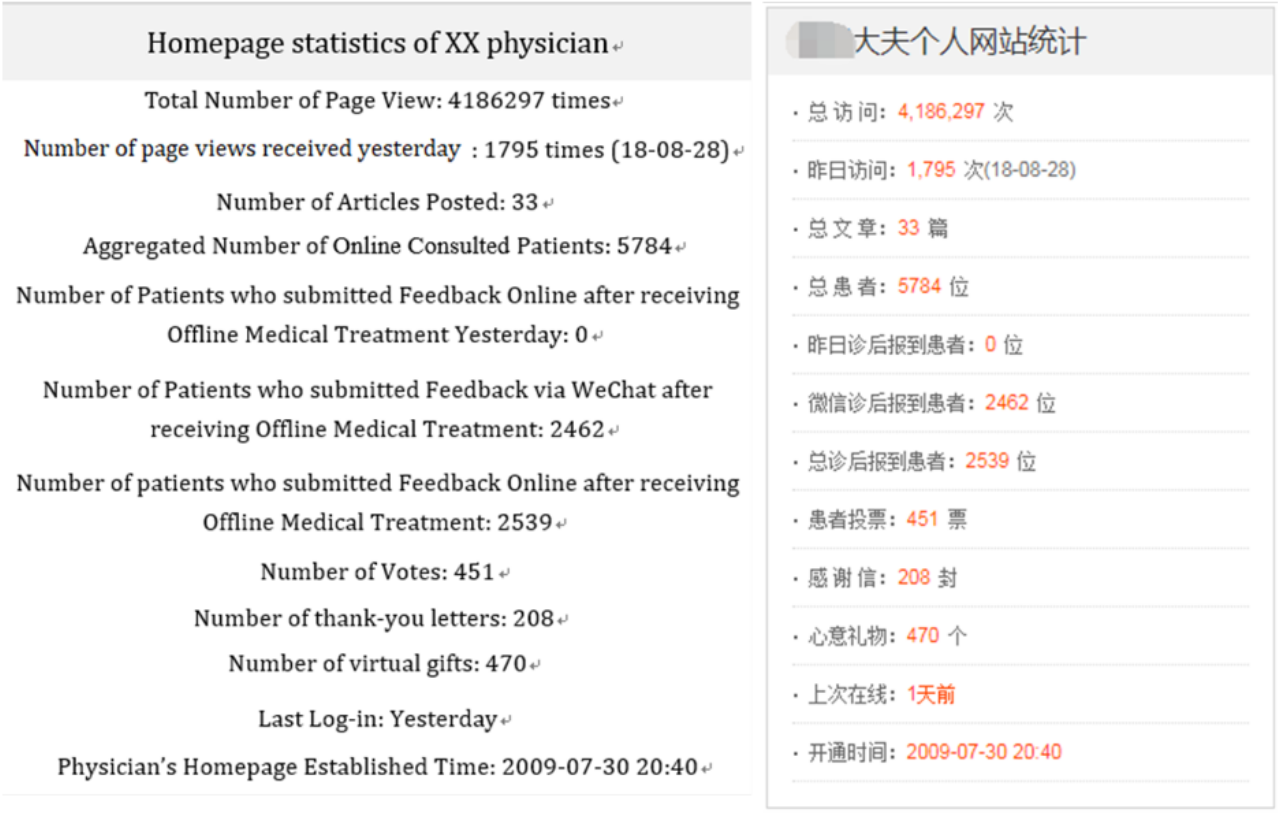
Figure 1. An example of a Physician's Homepage (Accessed on August 29, 2018).



The online homepages of physicians provide many types of information about physicians, including the total number of homepage views, number of votes, and number of thank-you messages and virtual gifts received. A screenshot from a physician's homepage is presented in Figures 1 and 2, which show the different types of information available to the patient.

XSL•FO
RenderX

Figure 2. An example of the physician statistics Web page, calculated by the Good Doctor website (accessed on August 29, 2018).



Dependent Variable

The aggregated number of patients that had received online consultations in the 60-day interval can reflect the patient's choice as interacting with a specific physician was a result of choice.

Independent Variables

The online effort of a physician was estimated by their contribution to the Good Doctor website and patients; the website determines this value to express physician activity and engagement levels, which can help patients choose the most suitable physician. Physicians can affect the value of their contributions by updating their information in a timely manner, publishing articles to educate patients, and answering questions received from previously consulted patients. These behaviors can indicate how much effort the physician puts into online rating websites. As Campbell and Pritchard [46] indicated, as both the duration of time spent working and the intensity of work activities represent important aspects of effort, it is appropriate to choose this contribution score as its proxy.

The physician's reputation possesses more measurement items than the previous. Generally, the persuasiveness of online consumer reviews depends on both their quality and quantity [47]. However, in the Good Doctor website, the star ratings only relate to the quality of service after diagnosis and treatment. The ratings are calculated by the website based on the number of patients diagnosed online and followed up online. A high proportion of patients followed up online will lead to more stars for the physician. If the physician diagnosed more than 100 patients online, he or she will get the star rating displayed on their homepage. As the star rating only refers to the followed-up element and only a small proportion of physicians gained this star rating, we have ignored the star rating in this study and only taken into consideration the quantity. We used the total number of page views shown on the physician's homepage, number of votes, number of thank-you messages, and the number of virtual gifts received as proxies for physician reputation. The total number of homepage views reflects the physician's fame, as online views are the first step and an indicator in learning about the physician. If patients are satisfied with the services provided by the physician, they can vote, write a thank-you message, or send virtual gifts to the physician. All options are free, except for virtual gifts; the website charges the patient for sending virtual gifts, ranging in price from several yuan to several hundred yuan. After deducting a small amount of website operating expenses, the fee will be allocated to the account opened by the physician for their time sacrificed. Table 2 provides sample data on physicians' efforts, observed from the Good Doctor website. Figures 3 and 4 show examples of a virtual gift and thank-you message.

To present major changes of the variables in the 3 waves, the mean value was adopted to show their trend, as shown in Figure 5. For the number of contributions, 1400 was deducted from the value, to fit the size. A log transformation was taken for views as its numerical value was much larger than other values.

XSL•FO
RenderX

Table 2. Measurement items and statistics of variables.

Variable, proxy, and wave	Minimum number of variables	Maximum number of variables	Mean (SD)
Number of consultations			
Consultation number of patients in a 60-day period			
1	0	16,003	143.78 (546.82)
2	0	17,620	154.27 (573.64)
3	0	18,443	160.56 (589.00)
Online effort of physician			
Contribution of physician, calculated by the Good Doctor website			
1	0	210,580	1478.48 (6456.89)
2	0	234,605	1589.70 (6837.36)
3	0	251,440	1687.46 (7112.74)
Online reputation of physician			
Total number of homepage views (log)			
1	3.09	16.23	9.76 (2.00)
2	5.78	16.27	10.00 (1.75)
3	6.32	16.32	10.15 (1.66)
Votes			
1	0	1587	10.23 (38.83)
2	0	1733	10.82 (41.38)
3	0	1860	11.31 (43.52)
Thank-you messages			
1	0	680	3.20 (16.60)
2	0	782	3.58 (18.36)
3	0	863	3.84 (19.71)
Virtual gifts			
1	0	1956	8.23 (49.80)
2	0	2137	8.89 (53.58)
3	0	2424	9.58 (58.74)

Figure 3. An example virtual gift.



XSL•FO
RenderX

Figure 4. An example of the thank-you message.

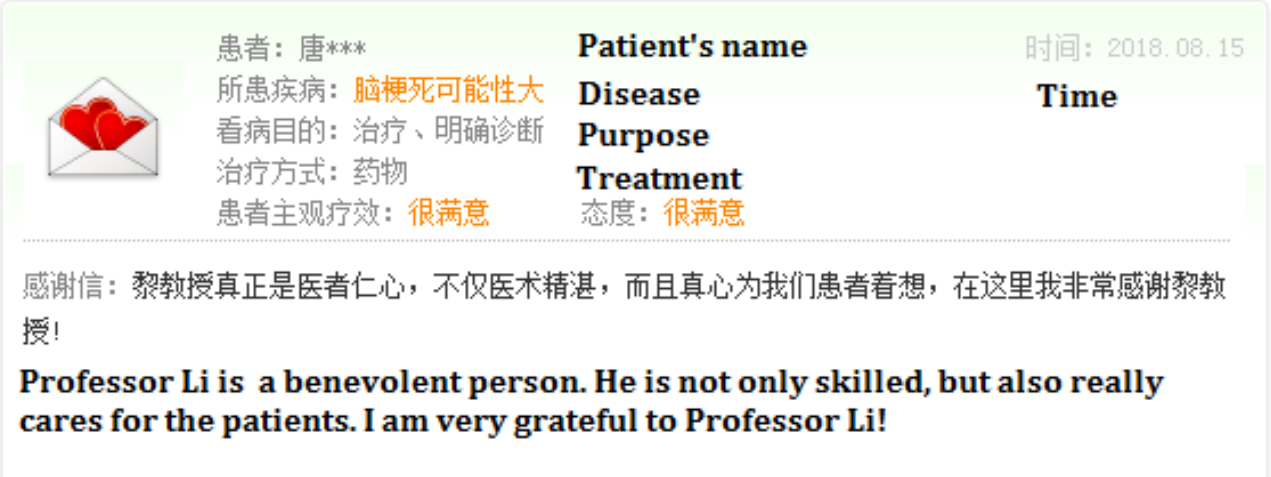
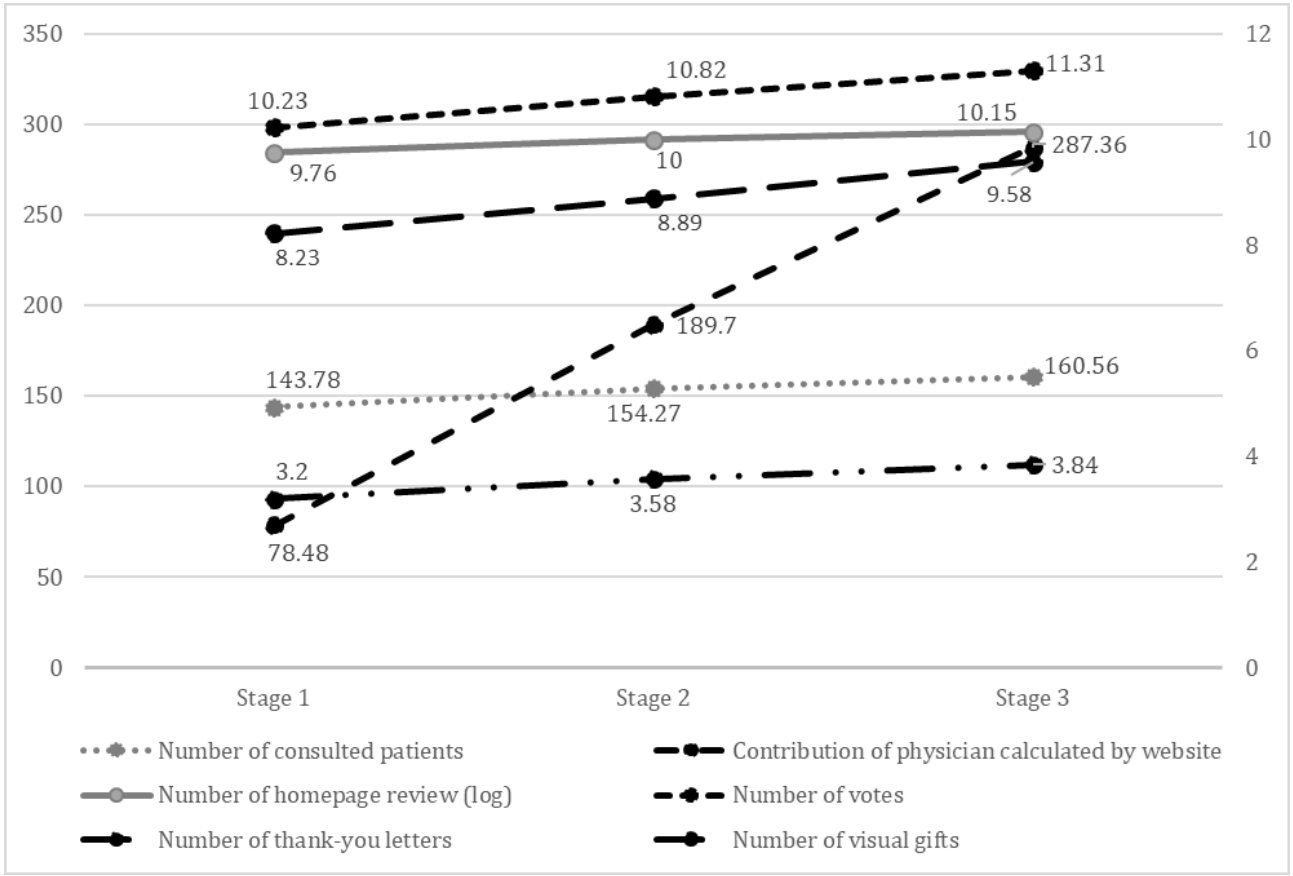


Figure 5. Statistics on the trend of different waves.



Control Variables

The control variables used in this study are (1) the physician's title in the hospital and (2) the level of the hospital. The title of a physician indicates the position and responsibility of the physician in the hospital. It can also reflect the professional expertise of the physician. There are 4 titles used on the Good Doctor website: Director Physician, Associate Director Physician, Attending Physician, and Residing Physician. A higher title suggests greater responsibility on the part of the physician. Two dummy variables (0 or 1) were used to indicate Director Physician, Associate Director Physician, and Attending Physician, respectively. There is also a variable that can represent the standing of the hospital, that is, the ranking of the hospital in China. According to the *standard for the grade management of hospitals* in China, the ranking of hospitals can be divided into 3 levels: level 1 refers to hospitals that are basic and typically provide health care services to communities, level 2 refers to secondary hospitals, and level 3 are tertiary hospitals. Level 3 hospitals employ more staff and own a greater number of beds than the other 2 levels and are often considered to provide a higher quality of service. We also used 2 dummy variables to indicate hospital rankings, ranging from 3 to 2. In our research model, these variables were used to control the effects of a physician's title and the ranking of hospitals on patient choices.

XSL•FO
RenderX

Model Estimation

To test our hypotheses against the effects of a physician's effort and reputation, we formulated a regression equation using the linear regression for each time series. We transformed the dependent variables and continuous independent variables into the log form, as the distribution of these variables may not be normal. The equation is separated into 2 stages, as shown in Figure 6(a).

Log(Consultation$_i$) denotes the number of patients who consulted the physician in the 60-day period. Secondary hospital and tertiary hospital indicate hospital ranking, whereas the title of each physician is also included. Log(Contribution$_i$) represents the effort of the physician online, whereas log(Hompage view$_i$), log(Vote$_i$), log(Thank-you message), and log(Gift$_i$) indicate the number of homepage views, votes, thank-you messages, and virtual gifts received, respectively; these represent the online reputation of physicians. After log transformation, the response variable is approximately distributed (the result of P-P plots is approximate, with a straight diagonal line indicating the data are normally distributed [48]; more details are shown in Multimedia Appendix 1). μ_i is the error term. The physician's title and the rank of the hospital are included as control variables.

Panel data analysis was also conducted, which is a widely used form of longitudinal analysis among social science researchers [49]. The panel data allowed for the control of unobservable variables that change over time and permitted the study of dynamics of change with time series. Meanwhile, the panel data controlled the variables that could not be observed or measured in each group (eg, gender) and the unobservable variables that changed over time but not across entities (eg, physicians with different titles). The panel data allowed for the inclusion of the variables at different levels of analysis.

In this study, gender, the physician's title, and the hospital level were chosen as the level of analysis. The following fixed-effect model was set up to explore the relationship between effort and reputation factors and physician's consultation within each type of physician, as something within the physician groups may impact or bias the predictor or outcome variables; this needed to be controlled in the model. The key insight is that if the unobserved variables do not change over time, then any changes in the outcome variable must be caused by influences other than the fixed characteristics. As such, once the effect of the time-invariant characteristics from the predictor variables is removed, we can assess the predictors' net effect on outcome variables. Generally, 2 approaches to build the fixed-effect model are highlighted, with the binary variable option being chosen as it allows for the separation of the association of the number of consultations and other individual factors. As 3 groups of entities were used (gender, physician's title, and hospital level) to generate binary (dummy) variables, 7 entities from 3 groups were presented in the final model, accordingly (see Figure 7).

Figure 6. Formula of the regression model.

$$\log(Consultation_i)=\alpha_0+\alpha_1\text{Secondary hospital}_i+\alpha_2\text{Tertiary hospital}_i+\alpha_3 Attending\ physician_i+\alpha_4 Associate\ director\ physician_i+\alpha_5\text{Director physician}_i+\alpha_6\log(Contribution_i)+\alpha_7\log(Homepage\ view_i)+\alpha_8\log(Vote_i)\ \alpha_9\log(Thank\text{-}you\ letter_i)+\alpha_{10}\log(Gift_i)+\mu_i \quad \text{(a)}$$

where i=1,2,...,n index all physicians

α_0 to α_{10} are the parameters to be estimated

Figure 7. Formula of the fixed-effect model.

$$Y_{it}=\beta_0+\beta_k X_{it}^k+\gamma_n E_n+\mu_{it}$$

where Y_{it} is the dependent variable (DV) where i = entity (different types of physicians) and t = time,

X_{it}^k $(k = 1,2,\ldots 5)$ represents five independent variables (IV)

β_k (k = 1, 2,... 5) is the coefficient for that IV

E_n(n = 1, 2, 3...7) is the binary entity

γ_n(n = 1, 2, 3...7) is the coefficient for the binary entities

μ_{it} is the error term

XSL•FO
RenderX

Results

Results of the Linear Regression

SPSS 19.0 (IBM) and Stata 12.0 (IBM) were used to analyze the data collected. Table 3 presents the results by ordinary least squares. Equations are presented in hierarchical order. First, the results are shown with only control variables in columns 1a and 1b. Then, the independent variables are added to columns 2a and 2b. The adjusted *R*-square and *F* value both indicate a good fit. The results of the VIF (Variance Inflation Factor) statistics for the variable indicate no multicollinearity (the VIF statistic of every variable is not greater than 10). Not all factors indicated a significant effect. The coefficient of contribution (B1=0.30, B2=0.37; *P*<.001), total number of homepage views (B1=0.43, B2=0.46; *P*<.001), and votes (B1=0.33, B2=0.27; *P*<.001) were all positive and significant. However, as thank-you messages (*P*=.07) and virtual gifts (*P*=.23) were not significant, the hypothesis could not be confirmed.

Results of the Fixed-Effect Model

Table 2 and Figure 5 show a numerical growth of each independent variable. To take the physician's individual factors into consideration, the fixed-effect model of panel data analysis was applied, as shown in Table 4. The effect of thank-you messages is seen to be significant, after the control of individual factors. Hypothesis 1 predicted that patients prefer to choose online physicians with greater effort. The results in Tables 3 and 4 support this hypothesis, as the coefficient of contribution (B1=0.30, B2=0.37; B3=0.19; *P*<.001) was positive and significant. Hypothesis 2 indicated a significant influencing path between the online reputation of the physician and patients' choice on physicians. Hypothesis 2 can be considered partly supported as the coefficient of number of homepage views (B1=0.43, B2=0.46; B3=0.18; *P*<.001) and number of votes (B1=0.33, B2=0.27; B3=0.48; *P*<.001) are significant, whereas thank-you messages are only significant in the fixed-effect model (B3=0.17; *P*=.01). However, virtual gifts (*P*=.11) were not significant in all models. The relationship between the online reputation of the physician and the patient's choice of physician is uncertain and needs further analysis. All control variables are not significant in the linear regression models. In the fixed-effect model, the physician's title effect is confirmed as physicians with the title of Associate Director Physician and Director Physician are more likely to be consulted with than the physician with the title of Residing Physician.

We further predicted a significant relationship between the online effort of physicians' and patients' choices of physicians. Tables 3 and 4 provide support for this hypothesis as the coefficient of contribution in all 3 models was positive and significant. Therefore, we can posit that the online effort of physicians has a positive impact on a patient's choice of physician. We assume that patients would like to consult with a physician who has a good online reputation in hypothesis 2. However, this assumption is only partly supported because of thank-you messages and virtual gifts being not significant.

Table 3. Linear regression results of model 1.

Variables	Number of consultations during the 60-day period[a], coefficient (95% CI)			
	Model 1a	Model 1b	Model 2a	Model 2b
Control variables				
Attending Physician	0.13 (−1.143 to 1.406)	0.41 (−0.689 to 1.504)	-1.94 (−4.941 to 1.068)	−0.79 (−3.386 to 1.809)
Associate Director Physician	0.82 (−0.430 to 2.085)	0.74 (−0.346 to 1.820)	−1.67 (−4.666 to 1.327)	−0.76 (−3.345 to 1.831)
Director Physician	0.81 (−0.076 to 1.701)	0.78 (−0.303 to 1.859)	−1.37 (−4.364 to 1.624)	−0.66 (−3.247 to 1.923)
Secondary hospital	0.76 (−0.313 to 1.830)	0.60 (−0.165 to 1.369)	−0.90 (−4.017 to 2.215)	−0.21 (−2.909 to 2.484)
Tertiary hospital	1.07 (−0.181 to 2.325)	0.70 (−0.069 to 1.459)	0.17 (−2.824 to 3.164)	0.35 (−2.238 to 2.947)
Independent variables				
Contribution (log)	—[b]	0.30[c] (0.214 to 0.375)	—	0.37[c] (0.254 to 0.476)
Total number of homepage views (log)	—	0.43[c] (0.335 to 0.536)	—	0.46[c] (0.358 to 0.564)
Number of votes (log)	—	0.33[c] (0.206 to 0.447)	—	0.27[c] (0.134 to 0.398)
Thank-you messages (log)	—	0.07 (−0.114 to 0.258)	—	0.17 (−0.015 to 0.354)
Virtual gifts (log)	—	0.12 (−0.082 to 0.326)	—	0.14 (−0.073 to 0.347)
Constant	0.84 (−0.597 to 2.282)	2.90[c] (1.462 to 4.345)	3.57 (−0.664 to 7.799)	3.88[d] (0.163 to 7.591)

[a]Model 1a: R^2=0.05, goodness of fit (F_{4037})=9.27. Model 1b: R^2=0.28 and goodness of fit (F_{4037})=33.38. Model 2a: R^2=0.03 and goodness of fit (F_{4037})=4.12. Model 2b: R^2=0.27 and goodness of fit (F_{4037})=27.64.

[b]Not included in the model.

[c]*P*<.01.

[d]*P*<.05.

XSL•FO
RenderX

Table 4. Results of fixed-effect model test (model 2).

Variables[a]	Number of patients consulted during the 60-day period (log)		
	Coefficient	*t* value (degrees of freedom)	*P* value
Contribution (log)	0.19	16.21 (4037)	<.001
Total number of homepage views (log)	0.18	8.04 (4037)	<.001
Number of votes (log)	0.48	40.36 (4037)	<.001
Thank-you messages (log)	0.17	11.82 (4037)	.01
Virtual gifts (log)	−0.06	1.13 (4037)	.17
Male	−0.02	−0.65 (4037)	.52
Female	0.01	0.31 (4037)	.76
Attending Physician	0.17	1.79 (4037)	.07
Associate Director Physician	0.22	2.33 (4037)	.02
Director Physician	0.30	3.15 (4037)	<.001
Secondary hospital	−0.07	−0.22 (4037)	.82
Tertiary hospital	−0.12	−0.37 (4037)	.71
Constant	−1.32	−3.82 (4037)	<.001

[a]Please note that the value of *F* test that all μ_{it} =0 (3643) is 149.15 (*P*<.001) and R^2=0.82.

Discussion

Principal Findings

The main purpose of this study was to examine the relationship between the online effort of physicians (contribution), physician reputation (total number of home page views, number of votes, number of thank-you messages, and virtual gifts received), and patients' choices of physicians on the health care website, Good Doctor. Similar to previous research [19,34,50], the predictions are tested through the data collected from the online health website. However, we then introduced physician effort and reputation and observed the data in 3 waves. To ensure the credibility of the stated results, we conducted 2 analyses and presented their differences in the results.

First, our results indicate that patients are more likely to consult with physicians online who show greater effort. The online effort of a physician was represented by the figure of reputation online, which can be increased by the physicians themselves by revising their personal information regularly, publishing educational articles for patients on their homepage, and answering questions posed by previously consulted patients. These actions can present a positive and hard-working image of a physician on their website. Patients would take these factors into consideration when seeking medical consultation. Liu [19] also suggested that physician efforts would influence patients. From the perspective of SET, researchers have focused on the consumers' effort paid on learning how to use or using online services [51], whereas little attention has been paid to the service providers. We confirm that the physician's effort is also important when providing services. The website's marketers and the physicians themselves should pay more attention to information relating to their effort if they wish to attract more patients for consultation. For example, health care websites can develop an *effort mechanism* whereby physicians can strive to achieve higher scores that are then displayed on their personal homepage.

Second, the effect of physicians' online reputations on the patient's choice is complex. Reputation is taken as an important part from SET. When researchers explore the behavior of online users and patients [36,52], it is a critical benefit for the online users. However, results indicate that not all reputational factors generate similar effects. Some of the factors, such as total number of homepage views and number of votes received, represent a significant result, and the coefficient indicates a positive influence. However, the effect of thank-you messages and virtual gifts is not always significant. These results are not consistent with the study by Yang et al [50]. These differences may be caused by the variances in sample size and time selection, as the study by Yang et al [50] was conducted in 2013, had a comparatively smaller sample size, and was not limited to a specific department. It may also suggest that the physicians in our study have already established their personal reputation through online efforts. As time proceeds, their reputation is less likely to rely on the result of thank-you messages and virtual gifts. In addition, cardiovascular disease patients can be difficult to please because of the nature of their chronic disease, which may also contribute to the insignificance. The coefficient of virtual gifts is not significant as patients must log in to the Good Doctor website and pay additional money for these virtual gifts, and consequently, the gift may not represent the true reputation. As we already understand that reputation is important for offline physicians [53], we are now convinced that it also makes sense in the online world. There are many factors that can reflect reputation; identifying the differences and the characteristics of these factors should be considered critical.

Third, the control effects of the level of hospital and title of physician are not seen as significant in the linear regression model. Level 1 hospitals (0.17%, 7/4037) are much less

XSL•FO
RenderX

significant than those categorized as secondary hospitals or tertiary hospitals. Physicians with the title of Residing Physician are also in a lower proportion (5.28%, 213/4037), which might have led to their result being not significant. However, the result from the fixed-effect model indicates a significant effect for physician's title. The different result with the regression model may be attributed to bias of data from only 1 period. Patients prefer physicians with the title of Associate Director Physician and Director Physician to the title of Residing Physician. The online platform may weaken the impact of gender and hospital and allow patients to focus more on the ability of the physicians.

Finally, the results also suggest a recurring problem, which was identified in the information asymmetry domain. Arrow [54] first proposed the concept of *information inequality* in the research of physicians and patients in medical markets, stating that an information problem exists in the relationship between patients and physicians. Patients try to obtain and analyze information before choosing a suitable physician. As a service provider, this common goal of information requirements by patients should be satisfied and lessen the divide between health supply and demand. Health care providers can offer more channels for both patients and physicians to express themselves, such as tailored functions to make inquiries, release detailed information for physicians, explain and quash possible misunderstandings, and try to maintain a positive image of the physician.

Conclusions

This paper explores the effects of physicians' online efforts and their reputation on patients' decision-making when choosing to consult with a physician. We used linear regression and fixed-effect modeling to test our hypotheses. Some of our assumptions have been proven by the results identified; the physicians' effort is positively associated with patients' decision-making, whereas the effect of a physician's reputation remains uncertain, with 1 or 2 negative items being identified. The differences in the results of linear regression and fixed-effect models indicate the importance of time factor and method selection.

The findings assist us in understanding the effects of such information on a patient's choice of a specific physician and, thus, contribute to the field of online health care research. By exploring the relationship between online physicians' factors and patients' choices of physicians, we contribute to the scarce literature relating to the online choice of physicians. It should be noted that the online information of physicians can affect the choice of patients. This research can help online health care providers and marketers to make strategic decisions on what information to display online to attract and retain a greater number of patients. The findings also extend the research of SET in the field of online health. More attention should be paid to the behavior of physicians.

Our findings also identify some limitations. First, data were collected solely from the Good Doctor website in China, and therefore, the findings may not be generalizable to other health care websites in other countries. Future research could collect and analyze data from multiple websites from different countries. Second, research must be conducted over a much longer period. The data used in this study were collected in 3 phases within a 4-month period only; we will continue the collection of data as part of our further study. Third, although the website confirms that the information about physicians was provided by the physicians themselves, we cannot guarantee that was always the case in reality. Finally, the variables of physician reputation and effort can be illustrated and measured by other items; for example, Wu and Lu [10] indicated that the reputation of one physician's colleagues would affect the quantity of a focal physician's future view.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (Award #: 71671073) and the Fundamental Research Funds for the Central University (HUST#:2018KFYYXJJ062).

Conflicts of Interest

None declared.

Multimedia Appendix 1

P-P plot of variables.

[PDF File (Adobe PDF File), 86KB - jmir_v21i3e10170_app1.pdf]

References

1. Deng Z, Liu S. Understanding consumer health information-seeking behavior from the perspective of the risk perception attitude framework and social support in mobile social media websites. Int J Med Inform 2017 Sep;105:98-109 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.05.014] [Medline: 28750916]
2. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. J Med Internet Res 2012;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]

XSL•FO
RenderX

3. Greaves F, Millett C. Consistently increasing numbers of online ratings of healthcare in England. J Med Internet Res 2012;14(3):e94 [FREE Full text] [doi: 10.2196/jmir.2157] [Medline: 22742977]
4. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. J Med Internet Res 2013 Feb;15(2):e24 [FREE Full text] [doi: 10.2196/jmir.2360] [Medline: 23372115]
5. Hao H. The development of online doctor reviews in China: an analysis of the largest online doctor review website in China. J Med Internet Res 2015 Jun 1;17(6):e134 [FREE Full text] [doi: 10.2196/jmir.4365] [Medline: 26032933]
6. Hao H, Zhang K. The voice of Chinese health consumers: a text mining approach to web-based physician reviews. J Med Internet Res 2016 May 10;18(5):e108 [FREE Full text] [doi: 10.2196/jmir.4430] [Medline: 27165558]
7. Zhang W, Deng Z, Hong Z, Evans R, Ma J, Zhang H. Unhappy patients are not alike: content analysis of the negative comments from China's Good Doctor website. J Med Internet Res 2018 Jan 25;20(1):e35 [FREE Full text] [doi: 10.2196/jmir.8223] [Medline: 29371176]
8. Emmert M, Sander U, Esslinger AS, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. Methods Inf Med 2012 Nov;51(2):112-120. [doi: 10.3414/ME11-01-0045] [Medline: 22101427]
9. Emmert M, Meier F, Heider AK, Dürr C, Sander U. What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. Health Policy 2014 Oct;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]
10. Wu H, Lu N. How your colleagues' reputation impact your patients' odds of posting experiences: evidence from an online health community. Electron Commer Res Appl 2016 Mar;16:7-17. [doi: 10.1016/j.elerap.2016.01.002]
11. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
12. Grabner-Kräuter S, Waiguny MK. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. J Med Internet Res 2015 Apr;17(4):e93 [FREE Full text] [doi: 10.2196/jmir.3991] [Medline: 25862516]
13. Verhoef LM, van de Belt TH, Engelen LJ, Schoonhoven L, Kool RB. Social media and rating sites as tools to understanding quality of care: a scoping review. J Med Internet Res 2014;16(2):e56 [FREE Full text] [doi: 10.2196/jmir.3024] [Medline: 24566844]
14. Okike K, Peter-Bibb TK, Xie KC, Okike ON. Association between physician online rating and quality of care. J Med Internet Res 2016 Dec 13;18(12):e324 [FREE Full text] [doi: 10.2196/jmir.6612] [Medline: 27965191]
15. Emmert M, Meszmer N, Sander U. Do health care providers use online patient ratings to improve the quality of care? Results from an online-based cross-sectional study. J Med Internet Res 2016 Sep 19;18(9):e254. [doi: 10.2196/jmir.5889] [Medline: 27644135]
16. Rothenfluh F, Germeni E, Schulz PJ. Consumer decision-making based on review websites: are there differences between choosing a hotel and choosing a physician? J Med Internet Res 2016 Jun 16;18(6):e129. [doi: 10.2196/jmir.5580] [Medline: 27311623]
17. McClean E, Collins CJ. High-commitment HR practices, employee effort, and firm performance: investigating the effects of HR practices across employee groups within professional services firms. Hum Resour Manage 2011 May 25;50(3):341-363. [doi: 10.1002/hrm.20429]
18. Gregg DG, Scott JE. The role of reputation systems in reducing on-line auction fraud. Int J Electron Comm 2006;10(3):95-120. [doi: 10.2753/JEC1086-4415100304]
19. Liu X, Guo X, Wu H, Vogel D. Doctor's Effort Influence on Online Reputation and Popularity. In: Smart Health. 2014 Presented at: International Conference on Smart Health (ICSH 2014); July 10-11, 2014; Beijing, China p. 111-126 URL: https://link.springer.com/chapter/10.1007/978-3-319-08416-9_11
20. Ye Q, Li Y, Kiang M, Wu W. The impact of seller reputation on the performance of online sales: evidence from TaoBao buy-it-now (BIN) data. ACM SIGMIS Database 2009 Jan 31;40(1):12-19. [doi: 10.1145/1496930.1496935]
21. Naylor JC, Pritchard RD, Ilgen DR. A Theory of Behavior in Organizations. New York: Academic Press; 1980:A.
22. Krishnan BC, Netemeyer RG, Boles JS. Self-efficacy, competitiveness, and effort as antecedents of salesperson performance. JPSSM 2002;22(4):285-295. [doi: 10.1080/08853134.2002.10754315]
23. Brown SP, Peterson RA. The effect of effort on sales performance and job satisfaction. J Mark 1994 Apr;58(2):70-80 [FREE Full text] [doi: 10.2307/1252270]
24. VandeWalle D, Brown SP, Cron WL, Slocum JW. The influence of goal orientation and self-regulation tactics on sales performance: a longitudinal field test. J Appl Psychol 1999;84(2):249-259 [FREE Full text] [doi: 10.1037/0021-9010.84.2.249]
25. Henrique DP, Matos C. The predictors of sales performance: a study with wholesale sellers. RBGN-REV BRAS GEST N 2014;16(52):448-465. [doi: 10.7819/rbgn.v16i52.1686]
26. Testa MR. Organizational Commitment, Job Satisfaction, and Effort in the Service Environment. J Psychol 2010 Apr;135(2):226-236. [doi: 10.1080/00223980109603693] [Medline: 11403344]
27. Emerson RM. Social Exchange Theory. Annu Rev Sociol 1976;2(7):335-362 [FREE Full text]
28. Burgess RL, Huston TL, editors. Social Exchange in Developing Relationships. London: Academic Press; 1979.

XSL•FO
RenderX

29. Guo S, Guo X, Fang Y, Vogel D. How doctors gain social and economic returns in online health-care communities: a professional capital perspective. J Manag Inf Syst 2017 Aug 17;34(2):487-519. [doi: 10.1080/07421222.2017.1334480]
30. Bitner MJ, Booms BH, Mohr LA. Critical service encounters: the employee's viewpoint. J Mark 1994 Oct;58(4):95-106 [FREE Full text] [doi: 10.2307/1251919]
31. Rice SC. Reputation and uncertainty in online markets: an experimental study. Inf Syst Res 2012 Jun;23(2):436-452. [doi: 10.1287/isre.1110.0362]
32. Camerer C, Weigelt K. Experimental tests of a sequential equilibrium reputation model. Econometrica 1988 Jan;56(1):1-36. [doi: 10.2307/1911840]
33. Resnick P, Kuwabara K, Zeckhauser R, Friedman E. Reputation systems. Commun ACM 2000;43(12):45-48. [doi: 10.1145/355112.355122]
34. Liu X, Guo X, Wu H, Wu T. The impact of individual and organizational reputation on physicians' appointments online. Int J Electron Comm 2016 Jun 20;20(4):551-577. [doi: 10.1080/10864415.2016.1171977]
35. Chen Y, Xie J. Online consumer review: word-of-mouth as a new element of marketing communication mix. Manage Sci 2008;54(3):477-491 [FREE Full text] [doi: 10.1287/mnsc.1070.0810]
36. Yan Z, Wang T, Chen Y, Zhang H. Knowledge sharing in online health communities: a social exchange theory perspective. Inf Manag 2016 Jul;53(5):643-653 [FREE Full text] [doi: 10.1016/j.im.2016.02.001]
37. Dewan S, Hsu V. Adverse selection in electronic markets: evidence from online stamp auctions. J Industrial Economics 2004 Dec;52(4):497-516. [doi: 10.1111/j.0022-1821.2004.00237.x]
38. Dulleck U, Kerschbamer R. On doctors, mechanics, and computer specialists: the economics of credence goods. J Econ Lit 2006 Feb;44(1):5-42 [FREE Full text] [doi: 10.1257/002205106776162717]
39. 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]
40. Jøsang A. Online reputation systems for the health sector. Electron J Health Inform 2008;3(1):e8 [FREE Full text]
41. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. J Am Med Assoc 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
42. Wang Y, Vassileva J. A Review on Trust and Reputation for Web Service Selection. In: Proceedings of the 27th International Conference on Distributed Computing Systems Workshops. 2007 Presented at: ICDCSW'07; June 22-29, 2007; Toronto, Canada p. 25. [doi: 10.1109/ICDCSW.2007.16]
43. Romano R, Baum N. Reputation management. J Med Pract Manage 2014;29(6):369-372. [Medline: 25108985]
44. Ramsaran-Fowdar RR. Identifying health care quality attributes. J Health Hum Serv Adm 2005;27(4):428-443 [FREE Full text] [Medline: 16318013]
45. [National Office of the Elderly Working Committee]. 2016. [Aging pressure on China's pension industry to be improved] URL: http://www.cncaprc.gov.cn/contents/16/175826.html [accessed 2019-01-23] [WebCite Cache ID 75ejgisQx]
46. Kanfer R. Motivation theory and industrial/organizational psychology. In: Dunnette M, Hough L, editors. Hand book of Industrial and Organizational Psychology. Chicago: Rand McNally College Publishing; 1990:75-170.
47. Park D, Lee J, Han I. The effect of on-line consumer reviews on consumer purchasing intention: the moderating role of involvement. Int J Electron Comm 2014 Dec 8;11(4):125-148. [doi: 10.2753/jec1086-4415110405]
48. Field A. Discovering Statistics Using IBM SPSS Statistics. London: Sage publications; 2009.
49. Frees EW. Longitudinal and Panel Data: Analysis and Applications in the Social Sciences. New York: Cambridge University Press; 2004.
50. Yang H, Guo X, Wu T, Ju X. Exploring the effects of patient-generated and system-generated information on patients' online search, evaluation and decision. Electron Commer Res Appl 2015 May;14(3):192-203. [doi: 10.1016/j.elerap.2015.04.001]
51. Chou SW, Hsu CS. Understanding online repurchase intention: social exchange theory and shopping habit. Inf Syst E-Bus Manage 2015 Jan 9;14(1):19-45. [doi: 10.1007/s10257-015-0272-9]
52. Shiau WL, Luo MM. Factors affecting online group buying intention and satisfaction: a social exchange theory perspective. Comput Human Behav 2012 Nov;28(6):2431-2444 [FREE Full text] [doi: 10.1016/j.chb.2012.07.030]
53. Torres E, Vasquez-Parraga AZ, Barra C. The path of patient loyalty and the role of doctor reputation. Health Mark Q 2009;26(3):183-197. [doi: 10.1080/07359680903263565] [Medline: 19813122]
54. Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev 1963;53(5):941-971 [FREE Full text]

Abbreviations

PRWs: physician-rating websites
SET: Social Exchange Theory

XSL•FO
RenderX

Edited by G Eysenbach; submitted 24.02.18; peer-reviewed by H Hao, S Bidmon, L Li; comments to author 17.08.18; revised version received 11.10.18; accepted 12.12.18; published 08.03.19.

Please cite as:
Deng Z, Hong Z, Zhang W, Evans R, Chen Y
The Effect of Online Effort and Reputation of Physicians on Patients' Choice: 3-Wave Data Analysis of China's Good Doctor Website
J Med Internet Res 2019;21(3):e10170
URL: https://www.jmir.org/2019/3/e10170/
doi:10.2196/10170
PMID:30848726

©Zhaohua Deng, Ziying Hong, Wei Zhang, Richard Evans, Yanyan Chen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Investigating the Effect of Paid and Free Feedback About Physicians' Telemedicine Services on Patients' and Physicians' Behaviors: Panel Data Analysis

Hualong Yang[1], PhD; Xiaofei Zhang[2], PhD

[1]School of Management, Guangdong University of Technology, Guangzhou, China

[2]Business School, Nankai Univeristy, Tianjin, China

Corresponding Author:
Xiaofei Zhang, PhD
Business School
Nankai Univeristy
94 Weijin Road
Tianjin, 300071
China
Phone: 86 22 23501039
Email: xiaofeizhang@nankai.edu.cn

Abstract

Background: In recent years, paid online patient-physician interaction has been incorporated into the telemedicine markets. With the development of telemedicine and telemedicine services, online feedback has been widely applied, helping other patients to identify quality services. Recently, in China, a new type of service feedback has been applied to the telemedicine markets, namely, paid feedback. Patients who are satisfied with a physician's online service can buy a virtual gift or give a tip to the physicians. This paid feedback can improve the reliability of service feedback and reduce the proportion of false information because it increases the cost for feedback providers. Paid online feedback can benefit the physicians, such as by providing them with monetary incentives; however, research on the impacts and value of such paid feedback from the physician perspective in the telemedicine markets is scant. To fill this research gap, this study was designed to understand the role of paid feedback by developing a research model based on the theories of signaling and self-determination.

Objective: This study aimed to explore the effects of free and paid feedback on patients' choice and physicians' behaviors as well as to investigate the substitute relationship between these 2 types of feedback in the telemedicine markets.

Methods: A JAVA software program was used to collect online patient-doctor interaction data over a 6-month period from a popular telemedicine market in China (Good Physician Online). This study drew on a 2-equation panel model to test the hypotheses. Both fixed and random effect models were used to estimate the combined effects of paid feedback and free feedback on patients' choice and physicians' contribution. Finally, the Hausman test was adopted to investigate which model is better to explain our empirical results.

Results: The results of this study show that paid feedback has a stronger effect on patients' choice (a_5=0.566; t_{2192}=9.160; P<.001) and physicians' contribution (β_4=1.332; t_{2193}=11.067; P<.001) in telemedicine markets than free feedback. Moreover, our research also proves that paid feedback and free feedback have a substitute relationship in determining patients' and physicians' behaviors (a_6=−0.304; t_{2191}=−5.805; P<.001 and β_5=−0.823; t_{2192}=−8.136; P<.001).

Conclusions: Our findings contribute to the extant literature on service feedback in the telemedicine markets and provide insight for relevant stakeholders into how to design an effective feedback mechanism to improve patients' service experience and physicians' engagement.

(J Med Internet Res 2019;21(3):e12156) doi:10.2196/12156

KEYWORDS

telemedicine; feedback; physician rating; health care quality; decision making; physicians' contribution

XSL•FO
RenderX

Introduction

Background

Telemedicine markets have become an important service platform for patients to communicate with physicians to obtain treatment and related medical information [1,2]. Without the limitations of time and space in the offline or conventional patient-physician context, the telemedicine markets enable patients to save medical expense and time [3-6]. Furthermore, physicians in such markets can interact with patients to obtain disease information and promote their online presence and reputation [7,8]. Therefore, telemedicine markets are effective approaches for both patients and physicians in the online context [9,10].

Nevertheless, the development of the telemedicine industry still faces many challenges. On the one hand, choosing the right physician is critical for patients because health care services are related to health and life [11,12]. However, because of information asymmetry and lack of professional health care knowledge, it is difficult for patients to ascertain a physician's competency and service quality based on the limited information and knowledge they obtain [13]. Telemedicine markets should provide efficient and adequate information to help patients make decisions about using such services. On the other hand, physicians' contributions (such as providing consultation services and health care information to patients) are indispensable resources for the development of telemedicine markets [7]. However, participating and contributing in telemedicine markets is burdensome for physicians because of their heavy workload in hospitals or other medical institutions. Thus, both patients and physicians encounter difficulties in participating in telemedicine markets. Understanding how to enable patients to make informed choices and facilitate physicians' contributions has become a managerial agenda for telemedicine practitioners.

Online feedback is regarded as an important information tool to help consumers understand the quality of services and products and interact with sellers, service providers, and other consumers [14-17]. With the development of telemedicine and telemedicine services, online feedback has been widely applied in the telemedicine markets [18]. Patients can draw on the service feedback mechanism to evaluate the quality of physicians' services and reduce information asymmetry in online health consultations [11,19]. Such feedback has become an efficient mechanism based on which patients can make informed decisions. However, the feedback mechanisms have mainly been studied in the electronic commerce context, although the current mechanisms have some disadvantages [20-23]. Generally, online feedback is free to write for the providers, which might result in the growth of fake and exaggerated feedback that can bias consumers' judgments on the service quality. Hence, in the telemedicine markets, a new feedback mechanism on physicians' service quality should be developed.

Recently, a new type of service feedback has been applied to the telemedicine markets, namely, paid feedback. Patients who are satisfied with a physician's online service can buy a virtual gift or give a tip to the physicians. This *paid feedback* can improve the reliability of service feedback and reduce the proportion of false information because it increases the cost for feedback providers. Moreover, this type of feedback can bring both reputational and monetary rewards for physicians, motivate their online contribution, and enhance their service quality. Therefore, paid feedback might have a positive effect on patients' and physicians' behaviors in the telemedicine markets and could substitute for the role of traditional free service feedback to a certain extent. However, little research attention has been paid to this new feedback mechanism, let alone its benefits for the telemedicine markets.

Objectives

Despite the prevalent use of service feedback in the telemedicine markets, empirical studies are still lacking in 2 important areas. First, research that distinguishes the different influences of free and paid feedback in telemedicine markets is scant. Previous research has mainly focused on the role of free feedback [11,19] and has neglected the existence of paid feedback. Second, although there have been extensive studies investigating the effects of online feedback on patient behaviors [18,24], not much research has explored such effects on physician behaviors in the health care context. Physicians provide consultation services, knowledge, and information to help patients understand their disease and obtain treatment as well as to promote the development of telemedicine markets over a long period. Hence, it is important to investigate the role of feedback on physician contribution. To fill these research gaps, the main research questions leading this study are as follows:

1. What is the strength of free and paid feedback on patients' choice and physicians' behaviors?
2. Is there a substitute relationship between free feedback and paid feedback?

Methods

Research Hypotheses

We combined signaling theory and self-determination theory to build our research model and propose the hypotheses. Our study explores the role and strength of free and paid feedback on patients' and physicians' behaviors.

Generally, information asymmetry refers to a situation in which some individuals have more information than others [25]. Signaling theory can help us understand individuals' behaviors and reduce information asymmetry [26]. This theory points out that efficacious signals can help individuals judge the quality of a service or product and influence their decision-making process. To be useful, a signal should possess important characteristics, that is, cost [27]. The cost of a signal refers to the expenditure when a signal was sent out. The cost of a signal relates to the value of the signal and is the core of signaling theory [28].

In telemedicine markets, patients do not have sufficient information to understand physicians' service quality, which generates a condition of information asymmetry [18]. Online feedback is regarded as an effective solution to the issue of information asymmetry. Previous studies have pointed out that the online feedback mechanism can develop trust and

XSL•FO
RenderX

cooperative relationships between users in a virtual context. Online feedback might be a useful signal to help patients judge service quality because such feedback originates from patients who have experienced the service and is more reliable than direct information from the service platforms and physicians [11].

Although both paid and free feedback are effective signals to influence patients' judgment and choice, the strengths of these 2 types of feedback are significantly different. According to signaling theory [26], the strength of signals depends on their cost. Paid feedback has a higher monetary cost than free feedback because patients need to pay related fees for the paid feedback. This means that such paid mechanism increases the providers' cost and decreases fake and exaggerated feedback. Therefore, because the value of paid feedback is higher, the reliability and strength of paid feedback are higher than those of free feedback [28]. When evaluating physicians' service quality, patients will rely more on paid feedback when making their choice. We thus hypothesized the following:

Hypothesis 1: Paid feedback has a stronger effect on patients' choice than free feedback.

According to the signaling theory, a strong signal will decrease the effect of a weak signal on people's behaviors [29]. Strong and weak signals have a substitute relationship in influencing personal decision making [27]. In telemedicine markets, because free feedback is adopted by telemedicine platforms to influence patients' choices, patients might perceive this free service feedback as physicians' marketing strategies and become skeptical. Due to higher monetary costs, patients are likely to think that paid feedback is a stronger signal than free feedback. Paid feedback can alleviate patients' skepticism and substitute for the role of free feedback in relation to their choices. Hence, when paid feedback and free feedback operate in coexistence in the telemedicine markets, these 2 types of feedback can have a substitute relationship in determining patients' choices. Therefore, we further hypothesized the following:

Hypothesis 2: Paid feedback and free feedback have a substitute relationship in affecting patients' choice.

Self-determination theory is useful to explain the motivation behind human behaviors [30]. This theory indicates that individuals are conditioned to behave in a certain way if they can obtain the relevant value to satisfy their basic psychological needs [31]. Although the individuals are not necessarily interested in the specific behavior or activity, they can obtain satisfaction from the extrinsic motivators [32].

In the telemedicine markets, the physicians' contribution is mainly focused on consultation services and health care information sharing. As information and knowledge belong to themselves [31], physicians' contribution is motivated by some motivators. Previous studies have pointed out that an appropriate extrinsic reward (such as reputation or money) can satisfy individuals' inner needs and stimulate their behaviors. Free online feedback from patients who have service experience in telemedicine markets can help physicians establish their reputation and satisfy their reputational needs. However, this type of feedback cannot satisfy their economic needs. Besides reputational rewards, paid feedback can bring monetary rewards that enhance physicians' financial position. Such reputational and monetary rewards will compensate physicians for the effort and time they contributed. According to self-determination theory, the strength of different motivators depends on the extent to which the individual's inner needs are satisfied. Hence, compared with free feedback, paid feedback can better satisfy physicians' need for reputational and monetary reward, rendering a stronger influence on their contribution behaviors. Hence, we hypothesized the following:

Hypothesis 3: Paid feedback has a stronger effect on physicians' contribution than free feedback.

Although both free and paid feedback might positively affect physicians' contribution behaviors, we expected that physicians would value paid feedback more than free feedback. According to self-determination theory [32], 2 types of motivators might have a substitute relationship in affecting individuals' behaviors when 1 type of motivator can better satisfy their inner needs. For instance, when paid feedback and free feedback operate concurrently in the telemedicine markets, physicians are concerned more about the amount of paid feedback they receive because paid feedback brings not only reputational but also monetary rewards. In other words, if physicians already have a certain amount of paid feedback, they will care less about the amount of free feedback they receive. Paid feedback can better satisfy physicians' needs, but free feedback provides only a limited amount of online reputation for physicians. Therefore, this type of feedback can substitute for the effect of free feedback in determining physicians' contribution in the telemedicine markets. On the basis of this, we hypothesized the following:

Hypothesis 4: Paid feedback and free feedback have a substitute relationship in physicians'contribution.

Research Design

To test the research hypotheses in this study, we collected related information and data from a popular telemedicine market in China, namely, "Good Physician Online" [33]. Patients can use this platform to consult physicians about their health conditions and inquire about related suggestions from the physicians. Moreover, this platform provides abundant service feedback to help patients reduce information asymmetry and make informed choices on services or physicians they are interested in or require. This platform provided sufficient data for this study, including physicians' occupational titles, hospital ranking, the amount of free and paid feedback they have received, the number of patients they have, and the level of the physicians' contribution.

The research design of this paper is as follows. First, we built a 2-equation panel model to investigate the effects of free and paid feedback on patients' choice and physicians' contribution, respectively. Second, we adopted a fixed effects model to estimate patients' choice and physicians' contribution—this model mainly controls the fixed differences at the physician level of our research objectives. Third, this study also estimated a random effect model on patients' choice and physicians' contribution—this model mainly controls for the random effects of research objectives in different periods. Fourth, the Hausman

XSL•FO
RenderX

test was conducted to determine which model was better for our research context. Finally, in accordance with a previously used method [34], we compared the coefficient of free and paid feedback to distinguish the strengths of the roles of the 2 aforementioned types of feedback.

Participants and Data

We designed a JAVA program to automatically collect data from the ''Good Physician Online'' website over a 6-month period. The data were organized in a panel and incorporated into our main dataset. After excluding the incomplete records, the online interaction experiences of 418 physicians and their patients were collected.

There were 2 dependent variables in our research model. We used different variables for the behavioral equations of patients and physicians. In the patients' choice equation, the dependent variable was the number of patients who had consulted the physician online. In accordance with a previous study [11], we used this variable as a proxy for patients' choice. In the physicians' contribution equation, the dependent variable was the level of physicians' contribution in telemedicine markets. Our research context provided an index for reflecting physicians' contribution behaviors, that is, *contributionvalue*, which was automatically calculated by the system of the platform. The contribution-values were calculated using 3 aspects of the physicians' behaviors: modifying personal information on the website, publishing health care articles, and responding to patients' online consultations. We drew on this variable to measure the level of physicians' contribution in the telemedicine markets.

In our 2-equation panel model, the independent variables were the number of free feedback and paid feedback items. In our research context, there are 2 types of positive service feedback given by patients: gratitude letters and virtual gifts. After an online consultation, if patients are satisfied with the physician's service, they can write a letter of gratitude or buy a virtual gift for the physician to express appreciation for the service. The gratitude letter is a text service feedback that patients use to express feedback on their experience of a physician's service process and quality. The virtual gift is similar to a type of digital card to express patients' appreciation of a physician's service. Both gratitude letters and virtual gifts can help patients to express feedback on their experience and physicians to develop their online reputation. Gratitude letters and virtual gifts differ in that the former are free, whereas the latter need to be paid for by patients. Physicians can obtain monetary rewards based on the price of the virtual gifts. Both gratitude letters and virtual gifts are associated with positive feedback about a physician's service quality. Hence, this study used these 2 variables as the proxies for free feedback and paid feedback, respectively.

Furthermore, in our research model, we added physicians' occupational title ranking, hospital standing, and the number of patients who have visited the physician's homepage as the control variables. Occupational title ranking refers to a physician's position ranking in an offline hospital. Hospital standing refers to a hospital's ranking as conferred by a professional health organization or the government. For patients, these rankings can reflect physicians' professional status, thereby affecting their judgment and choice decisions. For physicians, their professional status might be associated with their inner needs, thereby affecting their contribution behaviors. Thus, we used these 2 variables as control variables in the 2-equation model. Occupational title rankings, that is, director physician, deputy director physician, chief physician, and resident physician, are expressed, respectively, as 4, 3, 2, and 1. The hospital rankings from highest to lowest are referred to as follows: A class hospital, B class hospital, C class hospital, and D class hospital. In this paper, we have used 4, 3, 2, and 1, respectively, to indicate high-to-low rankings of hospitals. In addition, the promotion of telemedicine markets might increase the number of patients who visit a physician's homepage. When the number of patients who visited a physician's homepage is high, the number of patients who choose that physician will increase. Hence, we used the number of patients visiting the physician's homepage as an additional control variable in the patients' choice equation.

Table 1 presents the description of the variables in our research model. Table 2 presents the descriptive statistics of the variables, and Table 3 presents the correlations of the main variables in the research model, which indicates that there is no significant multicollinearity among the independent variables.

Research Model

A 2-equation panel model was used to test the hypotheses on patients' and physicians' behaviors. The first equation explored the relative effects of free feedback and paid feedback on patients' choices. In the second equation, we investigated the relative impacts of free feedback and paid feedback on the level of physicians' contribution. According to the statistics in Table 2 on the means and variances of the variables, the distributions of the dependent variables and many independent variables were not normal. Thus, this study developed a log-linear regression model. The 2-equation model is as follows:

$$Log(Patient_{it})=\alpha_0+\alpha_1 Log(Visiting_{it})+\alpha_2 Title\ Ranking_i$$

$$+\alpha_3 HospitalRanking_i+\alpha_4 Log(FreeFeedback_{it-1})+\alpha_5 Log(Paid\ Feedback_{it-1})$$

$$+\alpha_6 Log(FreeFeedback_{it-1})Log(PaidFeedback_{it-1})+\mu_i+\varepsilon_{it} \quad (1)$$

$$Log(Contribution_{it})=\beta_0+\beta_1 TitleRanking_i+\beta_2 HospitalRanking_i$$

$$+\beta_3 Log(FreeFeedback_{it-1})$$

$$+\beta_4 Log(PaidFeedback_{it-1})+\beta_5 Log(FreeFeedback_{it-1})$$

$$Log(PaidFeedback_{it-1})$$

$$+\eta_i+\theta_{it} \quad (2)$$

where i=1...N indicate the patient i or physician I and a_0 to a_6 are the parameters to be estimated in the first equation. The signals of μ_i and ε_{it} are error terms in the first equation, where β_0 to β_5 are the parameters to be estimated in the second equation, and η_i and θ_{it} are error terms associated with observation i. The variables $Log(Free_{it-1})\times Log(Paid_{it-1})$ are interaction terms to test the substitutable relationship between free and paid feedback.

XSL•FO
RenderX

Table 1. Description of variables.

Variable type and name	Proxy
Dependent variable	
Patients' choice	The number of patients who have consulted a physician online
Physicians' contribution	Contribution-value
Independent variable	
Free feedback	The number of gratitude letters
Paid feedback	The number of virtual gifts
Control variable	
Title ranking	Title ranking is expressed, respectively, as 4, 3, 2, and 1
Hospital standing	Hospital standing is expressed 4, 3, 2, and 1, respectively
The number of patients visiting	The number of patients who have visited the physician's homepage

Table 2. Descriptive statistics of variables.

Variable	Minimum	Maximum	Mean (SD)
Title ranking	1.000	4.000	2.949 (0.626)
Hospital standing	1.000	4.000	2.990 (0.689)
Number of visiting	2.00	13,520	3035.924 (792.351)
Number of free feedback	0.000	169.000	3.557 (11.971)
Number of paid feedback	0.000	374.000	9.569 (34.337)
Level of contribution	0.000	20,899.000	5192.629 (1864.505)
Number of patients	0.000	14,825.000	401.616 (1208.634)

Table 3. Correlations of variables.

Variable		1	2	3	4	5	6	7
1	Title ranking	1	0.383[a]	0.182[a]	0.203[a]	0.248[a]	0.239[a]	0.204[a]
2	Hospital standing	0.383[a]	1	0.025	0.146[a]	0.121[a]	0.061[a]	0.058[a]
3	Number of visiting	0.182[a]	0.025	1	0.442[a]	0.469[a]	0.463[a]	0.465[a]
4	Number of free feedback	0.203[a]	0.146[a]	0.442[a]	1	0.515[a]	0.497[a]	0.474[a]
5	Number of paid feedback	0.248[a]	0.121[a]	0.469[a]	0.515[a]	1	0.539[a]	0.548[a]
6	Level of contribution	0.239[a]	0.061[a]	0.463[a]	0.497[a]	0.539[a]	1	0.570[a]
7	Number of patients	0.204[a]	0.058[a]	0.465[a]	0.474[a]	0.548[a]	0.570[a]	1

[a]$P<.05$.

Results

Model Estimation

Fixed effects models and random effects models are widely used estimation methods on panel data. Fixed effects models are used to test the fixed effects of research objectives caused by individual differences and heterogeneity, and random effects models are used to test the random effects of research models caused by time variation. Generally, the Hausman test can determine which model is more appropriate for result estimation. We used both fixed effects and random effects models to estimate patients' choices and physicians' contributions, respectively. Then, we selected our main model based on the results of the Hausman test to explain our research results.

First, we used a fixed effects model to estimate the results of patients' choice in the first equation. The first 2 columns in Table 4 represent the results of regression by ordinary least squares (OLS) with fixed effects. Second, we further used a random effects model to estimate the results of the first equation. The latter 2 columns in Table 4 show the results of regressions by generalized least squares (GLS) with random effects. In the first 2 columns in Table 4, the results of the fixed and random effects models are hierarchical. The independent variables are presented in column 1, and the interactions of variables are

XSL•FO
RenderX

presented in column 2. The results of the Hausman test (P<.001; χ^2_5=47.9) indicate the appropriateness of the fixed effects model is over the random effects model. Hence, we consider the results of the fixed effects model appropriate for the first equation model.

Table 4. The results of the first equation model.

Variable	1 Fixed effects	2 Fixed effects	1 Random effects	2 Random effects
Cons	−4.287[a] [−16.140, 2192][b]	−3.833[a] [−13.977, 2191]	−4.022[a] [−19.235, 2192]	−3.989[a] [−19.102, 2191]
Title ranking	0.061 [1.113, 2192]	0.058 [1.085, 2191]	0.120[c] [2.581, 2192]	0.118[c] [2.544, 2191]
Hospital standing	0.114 [1.232, 2192]	0.070 [0.769, 2191]	−0.045 [−0.754, 2192]	−0.055 [−0.910, 2191]
Log (visiting)	0.658[a] [41.482, 2192]	0.648[a] [40.924, 2191]	0.675[a] [47.858, 2192]	0.666[a] [46.981, 2191]
Log (free feedback)	0.212[a] [2.727, 2192]	0.379[a] [4.608, 2191]	0.206[a] [4.604, 2192]	0.378[a] [6.389, 2191]
Log (paid feedback)	0.566[a] [9.160, 2192]	0.756[a] [10.896, 2191]	0.366[a] [8.072, 2192]	0.521[a] [9.175, 2191]
Log (paid feedback) ×Log (free feedback)	—[d]	−0.304[a] [−5.805, 2191]	—[c]	−0.114[a] [−4.452, 2191]
Observations	2198	2198	2198	2198
Number of groups	418	418	418	418
R^2	0.541	0.549	0.538	0.544
F (degree of freedom)	419.023[a] (4, 1776)	361.213[a] (5, 1775)	—[e]	—[e]
χ^2 (degree of freedom)	—[e]	—[e]	3712.2[a](5)	3744.5[a](6)

[a]P<.001.

[b]Values within parenthesis indicate t value and degrees of freedom of the coefficient.

[c]P<.05.

[d]Missing data.

[e]Missing test.

Hypothesis 1 predicted that paid feedback has a stronger effect on patients' choice than free feedback. According to column 1 of Table 4, the coefficient of free feedback (a_4=0.212; t_{2192}=2.727; P<.001) is positive and statistically significant. The coefficient of paid feedback (a_5=0.566; t_{2192}=9.160; P<.001) is also positive and statistically significant at the .01 level. Therefore, both free feedback and paid feedback positively affect patients' choice. Moreover, the coefficient of paid feedback is larger than the coefficient of free feedback. Hence, this is evidence to support hypothesis 1.

Hypothesis 2 postulated that paid feedback and free feedback have a substitute relationship in patients' decisions. Evidence to support this hypothesis lies in the fact that, according to column 2 of Table 3, the coefficient of the interaction term between paid and free feedback is negative and statistically significant (a_6=−0.304; t_{2191}=−5.805; P<.001). This result indicates that the impacts of free feedback and paid feedback on patients' decision behaviors are substitutive.

We further used a fixed effects model and a random effects model to estimate the results of physicians' contribution in the second equation. Table 5 represents the results of regression by OLS with fixed effects and by GLS with random effects. The results of the Hausman test (P<.001; χ^2_4=80.612) indicate the appropriateness of the fixed effects model over the random effects model. Hence, we consider the results of the fixed effects model appropriate for the second equation model.

Hypothesis 3 postulates that paid feedback has a stronger effect on physicians' contribution than free feedback. According to column 1 of Table 5, the coefficient of free feedback (β_3=0.883; t_{2193}=5.879; P<.001) is positive and statistically significant. The coefficient of paid feedback (β_4=1.332; t_{2193}=11.067; P<.001) is also positive and statistically significant. Hence, both free feedback and paid feedback positively affect patients' decisions. Moreover, the coefficient of paid feedback is larger than that of free feedback. Therefore, hypothesis 3 is supported.

Hypothesis 4 postulated that paid feedback and free feedback have a substitute relationship in determining physicians' contributions. Evidence to support this hypothesis lies in the fact that according to column 2 of Table 5, the coefficient of Log(paid feedback)×Log(free feedback) (β_5=−0.823; t_{2192}=−8.136; P<.001) between paid feedback and free feedback is negative and statistically significant, that is, the impacts of free feedback and paid feedback on physicians' contribution are substitutive. Figure 1 shows the results of our 2-equation panel model.

XSL•FO
RenderX

Table 5. The results of the second equation model.

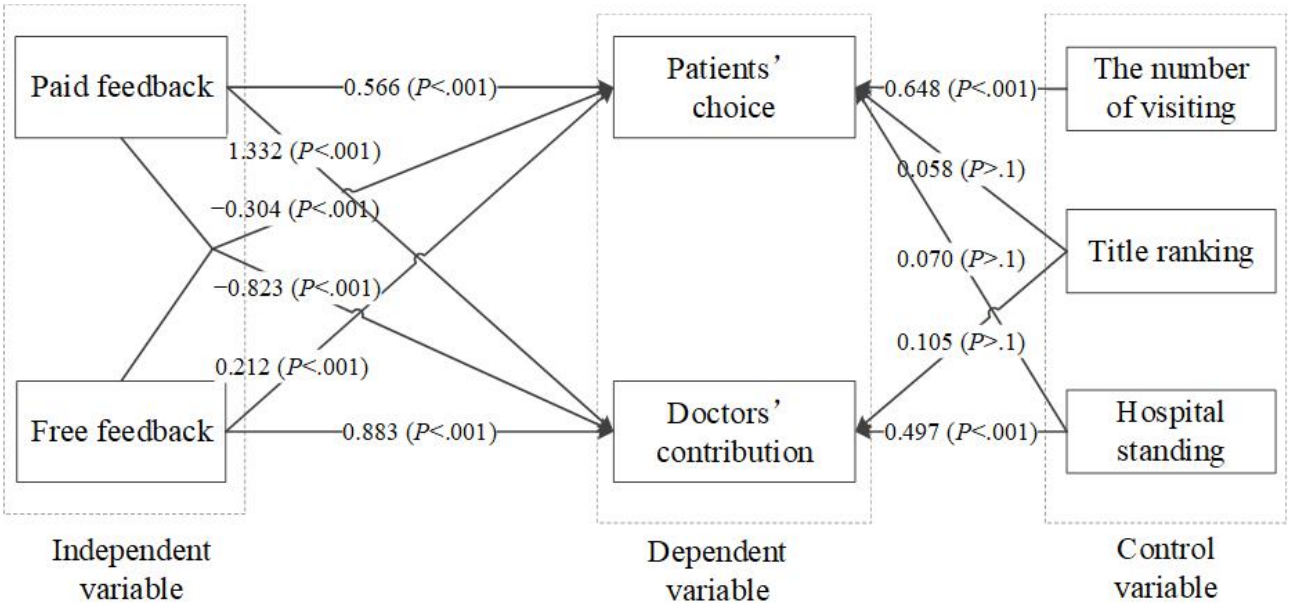
Variable	1 Fixed effects	2 Fixed effects	1 Random effects	2 Random effects
Cons	1.659[a] [3.576, 2193][b]	2.673[a] [5.647, 2192]	2.968[a] [8.245, 2193]	2.581[a] [8.032, 2192]
Title ranking	0.114 [1.061, 2193]	0.105 [0.992, 2192]	0.281[a] [3.004, 2193]	0.268[a] [2.913, 2192]
Hospital standing	0.632[a] [3.493, 2193]	0.497[a] [2.786, 2192]	0.025 [0.213, 2193]	−0.021 [−0.181, 2192]
Log (free feedback)	0.883[a] [5.879, 2193]	1.309[a] [8.351, 2192]	1.126[a] [13.076, 2193]	1.387[a] [12.346, 2192]
Log (paid feedback)	1.332[a] [11.067, 2193]	1.836[a] [8.354, 2192]	0.828[a] [8.995, 2193]	1.739[a] [15.504]
Log (paid feedback) ×Log (free feedback)	—[c]	−0.823[a] [−8.136, 2192]	—[c]	−0.435[a] [−8.382, 2192]
Observations	2198	2198	2198	2198
Number of groups	418	418	418	418
R^2	0.369	0.122	0.436	0.457
F (degree of freedom)	45.932[a] [3, 1777)	51.308[a] [4, 1776)	—[d]	—[d]
χ^2 (degree of freedom)	—[d]	—[d]	490.883[a] (4)	574.732[a] (5)

[a]P<.05.

[b]Values within parenthesis indicate t value and degrees of freedom of the coefficient.

[c]Missing data.

Figure 1. The results of research model.



Robustness Check

To check the robustness of our research model, we used the generalized method of moments to run the 2-equation model again. This method can effectively reduce the endogeneity and heterogeneity of regression models. In this part, we were aimed to investigate the main effect of paid and free feedback on patients' and physicians' behaviors and to prove the causal link between independent and dependent variables. Hence, we selected the generalized method of moments to estimate our model.

Table 6 presents the estimation results of the 2-equation models. Column 1 of Table 6 indicates the results of patients' choice equation and column 2 presents the results of physicians' contribution equation. The coefficients of paid and free feedback in 2 equations are positive and statistically significant. Therefore, the results of robustness checks are consistent with the results of the main model.

XSL•FO
RenderX

Table 6. The results of robustness check.

Variable	Patients' choice equation	Physicians' contribution equation
Logpatient$_{-1}$	−0.022 [−0.256, 2191][a]	0.682[b] [4.908, 2192]
Title ranking	−1.493[b] [−23.623, 2191]	−0.006 [−0.001, 2192]
Hospital standing	0.070[b] [1.927, 2191]	0.276[b] [3.897, 2192]
Log (visiting)	0.720[b] [21.978, 2191]	—[c]
Log (paid feedback)	0.159[b] [2.724, 2191]	0.094[b] [7.098, 2192]
Log (free feedback)	0.379[b] [6.392, 2191]	0.967[b] [9.507, 2192]
Observations	2198	2198
Number of groups	418	418
Wald χ^2 (degree of freedom)	6.28e+06[b] (6)	3.69e+06[b] (5)

[a]The values in parenthesis are the *z* value and degree of freedom of the coefficients.

[b]$P<.001$.

[c]Missing data.

Discussion

Summary of Findings

This study mainly investigated the influence of free and paid feedback on patients' choices and physicians' contributions in the telemedicine markets. On the basis of theories of signaling and self-determination, we developed 4 research hypotheses and established empirical models. The results of our research model support all of our hypotheses. Accordingly, this work provides 3 aspects of key findings. First, we found that paid feedback has a stronger effect on patients' choice than free feedback. Second, we found that paid feedback has a stronger influence on physicians' contribution than free feedback. Third, the empirical results of this study proved that paid feedback and free feedback have a substitute relationship in determining patients' and physicians' behaviors.

Discussion of Research Results

Previous studies have found that online feedback has a positive effect on consumers' decisions. Our study divided online feedback into free and paid feedback and distinguished the strengths of the 2 types of feedback on patients' choices. According to the signaling theory, the strength of signals is dependent on the cost. Paid feedback as a signal that reflects physicians' service quality has a higher cost than free feedback. In the telemedicine markets, free feedback from patients with service experience could be falsified and therefore be misleading for other patients who rely on such feedback. Paid feedback improves the cost of providers and reduces the risk of unreliable information. Thus, the reliability and strength of paid feedback for patients to judge service quality are higher compared with free feedback. Figure 2 shows the strength of free and paid feedback on patients' choice.

Moreover, previous studies on online feedback have mainly focused on the perspective of service receivers. This paper investigated the role and strength of paid and free feedback from the perspective of service providers. According to self-determination theory, although individuals are not necessarily interested in the specific behavior or activity, they can obtain satisfaction from the extrinsic reward. Although free feedback can promote physicians' online prestige, paid feedback not only compensates physicians with reputation but also provides monetary rewards for their time, efforts, and expenses in the telemedicine markets. Hence, the value of paid feedback is higher for physicians, which further stimulates their contribution online. Figure 3 shows the strength of the 2 types of service feedback on physicians' contribution.

In addition, for patients, paid feedback is a stronger signal influencing their judgment about a physician's quality of service and affects their choice because the cost of paid feedback is higher than that of free feedback. According to the signaling theory, strong signals can decrease the role of weak signals in influencing individuals' decision making. There is a substitute relationship between strong and weak signals. Hence, paid feedback can substitute for the role of free feedback on patients' behaviors in telemedicine markets. Furthermore, according to self-determination theory, the strength of motivators on human behavior depends on the individual's needs. Although both free and paid feedback can offer physicians the incentive of a heightened profile and better reputation, paid feedback has the added incentive of monetary compensation. Hence, paid feedback is more valuable for physicians and could weaken the impact of free feedback on their behaviors in telemedicine markets. On the basis of the above discussion, we find that a new feedback mechanism (ie, paid feedback) can substitute for the role of the traditional one (ie, free feedback) in telemedicine markets.

XSL•FO
RenderX

Figure 2. The strength of paid and free feedback on patients.

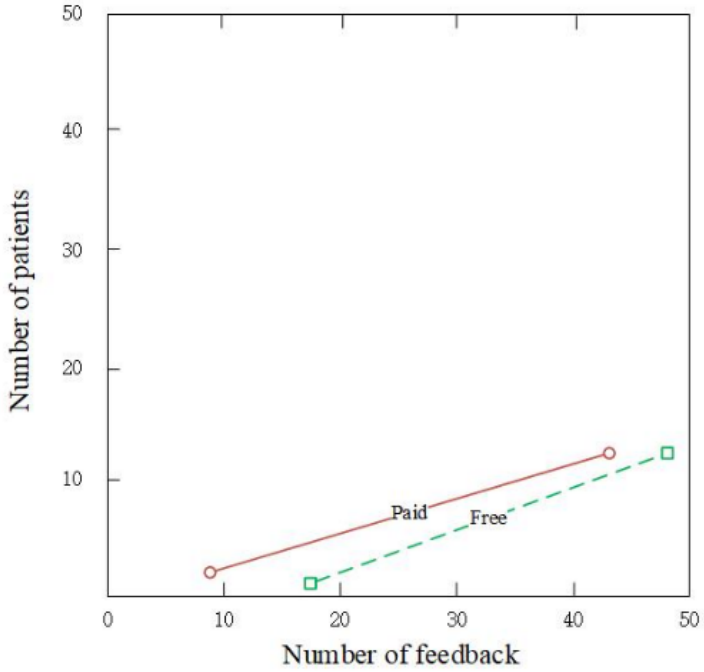
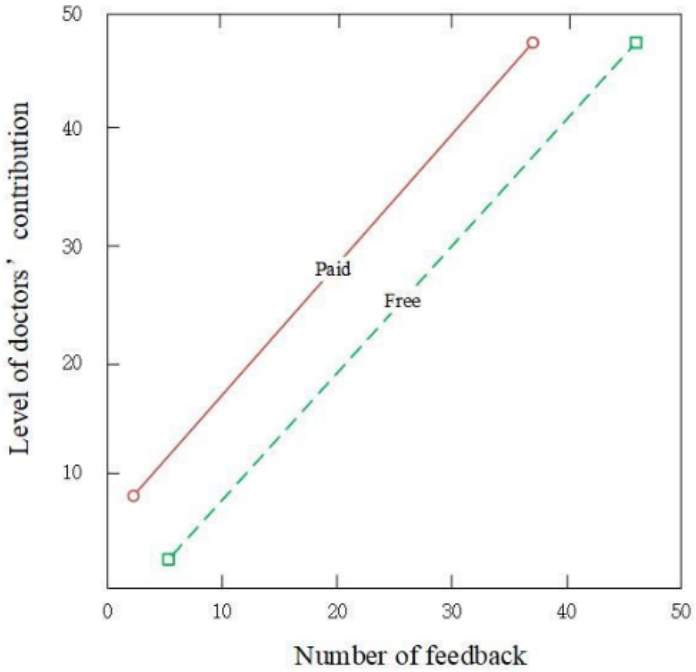


Figure 3. The strength of paid and free feedback on physicians.



Limitations and Future Research

However, this study still has a few limitations. First, although this study investigated the effects of free and paid feedback on patients' and physicians' behavior, we did not figure out the inherent reasons, such as patients' perceived value and trustworthiness of this online feedback. In future research, we plan to develop a questionnaire and use a structural equations model to explore the influence of psychological factors on patients' or physicians' behavior. Second, our research model

XSL•FO
RenderX

did not take into account the moderating effects of patients and physicians' characteristics. The influence of paid and free feedback on different people could vary differently. In the future, we will use more theories and methods to explore the moderating effect of individuals' characteristics on the relationship between online feedback and human behavior. Finally, we focused only on 1 specific telemedicine market to collect our research data. This could limit the generalization of our research results. Our future research aims to collect data from several different platforms to test the role of paid and free feedback.

Contributions

Despite the limitations set out above, this study makes several contributions. First, it has investigated the influences of paid feedback in the telemedicine markets for the first time. Although abundant studies have examined the role of free feedback in online markets, few have explored the influence of paid feedback. In filling this research gap, this study found that both free and paid feedback have positive effects on patients' choice and physicians' contribution in telemedicine markets. Our finding contributes to the literature on online feedback and the current knowledge on telemedicine markets.

Second, it combined the perspectives of both patients and physicians to comprehensively understand the role of online feedback in the telemedicine markets. In the current literature, there are limited studies on the combined perspectives of service providers and receivers. This study drew on signaling theory and self-determination theory to investigate the effects of online feedback on patients' and physicians' behaviors. Our empirical results expand signaling theory and self-determination theory into the research context of telemedicine market.

Third, this study distinguished the relative strengths of the 2 types of service feedback, that is, paid feedback has a stronger effect on patients' choice and physicians' contribution than free feedback. This finding helps us to understand the strength of paid and free feedback in telemedicine markets.

Finally, this study found that paid feedback and free feedback have a substitute relationship in patients' and physicians' behaviors. This finding suggests an extension of related research on online feedback and helps us understand the relationship between paid and free feedback.

Furthermore, this paper also points out important practical strategies for marketers and developers of telemedicine markets. First, how to incentivize physicians to contribute to the telemedicine markets is an important issue for designers because physicians are one of the most important resources for the development of the telemedicine markets. Developers should promote the role of paid feedback and improve reputational and monetary rewards to improve physicians' contributions in the telemedicine markets. Second, although paid feedback could influence patients' opinions on physicians' service quality and support their decision-making process, such feedback will increase the monetary cost of providers and decrease the supply of paid feedback. Hence, developers of telemedicine platforms should design an effective compensation mechanism to reduce the cost for providers. For example, telemedicine platforms can provide some coupons for providers of paid feedback.

Conclusions

Paid feedback is a novel mechanism in online health care services that can help patients judge the quality of service provided by physicians and improve physicians' contribution to telemedicine markets. Although extensive studies have been conducted to investigate the influence of service feedback on patients' behaviors, little research has explored the role and strength of paid and free feedback from the perspectives of patients and physicians. To fill this research gap, we developed a 2-equation panel model based on the theories of signaling and self-determination and used data collected from a real telemedicine market to test our hypotheses. The empirical results of our research indicated that paid feedback has a stronger effect on patients' decisions and physicians' contribution than free feedback. Furthermore, we found that paid feedback can substitute for the role of free feedback in telemedicine markets. Therefore, paid feedback can become a vital tool for both patients and physicians. Our findings may contribute to relevant ongoing or future studies in the telemedicine service sector and provide suggestions for designers and developers of telemedicine markets.

Acknowledgments

This study was partially funded by the China Postdoctoral Science Foundation Grant (2017M622647). The authors greatly appreciate the work of the editors and anonymous reviewers and are grateful for their insightful comments and suggestions. All errors remain ours.

Conflicts of Interest

None declared.

References

1. Liu N, Tong Y, Chan HC. Information seeking in online healthcare communities: the dual influence from social self and personal self. IEEE Trans Eng Manag 2017;99:1-10. [doi: 10.1109/TEM.2017.2703121]
2. Guo S, Guo X, Zhang X, Vogel D. Doctor–patient relationship strength's impact in an online healthcare community. Inf Technol Dev 2017 Mar 09;24(2):279-300. [doi: 10.1080/02681102.2017.1283287]

XSL•FO
RenderX

3. Yang H, Guo X, Wu T. Exploring the influence of the online physician service delivery process on patient satisfaction. Decis Support Syst 2015;78:113-121. [doi: 10.1016/j.dss.2015.05.006]
4. Yan Z, Wang T, Chen Y, Zhang H. Knowledge sharing in online health communities: a social exchange theory perspective. Inf Manag 2016 Jul;53(5):643-653. [doi: 10.1016/j.im.2016.02.001]
5. Jie MG, Gao G, Agarwal R. The creation of social value: can an online health community reduce rural?urban health disparities? MIS Q 2016;40(1):247-263. [doi: 10.25300/MISQ/2016/40.1.11]
6. Zhang X, Guo X, Wu Y, Lai KH, Vogel D. Exploring the inhibitors of online health service use intention: a status quo bias perspective. Inf Manag 2017;54(8):987-997. [doi: 10.1016/j.im.2017.02.001]
7. Guo S, Guo X, Fang Y, Vogel D. How doctors gain social and economic returns in online health-care communities: a professional capital perspective. J Manag Inf Syst 2017 Aug 17;34(2):487-519. [doi: 10.1080/07421222.2017.1334480]
8. Wu H, Lu N. How your colleagues' reputation impact your patients' odds of posting experiences: Evidence from an online health community. Electron Commer Res Appl 2016 Mar;16:7-17. [doi: 10.1016/j.elerap.2016.01.002]
9. Zhang P, Dong L, Chen H, Chai YL, Liu J. The rise and need for mobile apps for maternal and child health care in China: survey based on app markets. JMIR Mhealth Uhealth 2018 Jun 08;6(6):e140 [FREE Full text] [doi: 10.2196/mhealth.9302] [Medline: 29884601]
10. Van Kessel K, Babbage DR, Reay N, Miner-Williams WM, Kersten P. Mobile technology use by people experiencing multiple sclerosis fatigue: survey methodology. JMIR Mhealth Uhealth 2017 Feb 28;5(2):e6 [FREE Full text] [doi: 10.2196/mhealth.6192] [Medline: 28246073]
11. Yang H, Guo X, Wu T, Ju X. Exploring the effects of patient-generated and system-generated information on patients' online search, evaluation and decision. Electron Commer Res Appl 2015 May;14(3):192-203. [doi: 10.1016/j.elerap.2015.04.001]
12. Arrow KJ. Uncertainty and the welfare economics of medical care. 1963. Bull World Health Organ 2004 Feb;82(2):141-149 [FREE Full text] [Medline: 15042238]
13. Yaraghi N, Wang W, Gao G, Agarwal R. How online quality ratings influence patients' choice of medical providers: controlled experimental survey study. J Med Internet Res 2018 Mar 26;20(3):e99 [FREE Full text] [doi: 10.2196/jmir.8986] [Medline: 29581091]
14. Mishra A, Maheswarappa SS, Maity M, Samu S. Adolescent's eWOM intentions: an investigation into the roles of peers, the internet and gender. J Bus Res 2018 May;86:394-405. [doi: 10.1016/j.jbusres.2017.04.005]
15. Wu J. Review popularity and review helpfulness: a model for user review effectiveness. Decis Support Syst 2017 May;97:92-103. [doi: 10.1016/j.dss.2017.03.008]
16. Karimi S, Wang F. Online review helpfulness: impact of reviewer profile image. Decis Support Syst 2017 Apr;96:39-48. [doi: 10.1016/j.dss.2017.02.001]
17. Baek H, Oh S, Yang HD, Ahn JH. Electronic word-of-mouth, box office revenue and social media. Electron Commer Res Appl 2017 Mar;22:13-23. [doi: 10.1016/j.elerap.2017.02.001]
18. Liu X, Guo X, Wu H, Wu T. The impact of individual and organizational reputation on physicians' appointments online. Int J Electron Commer 2016 Jun 20;20(4):551-577. [doi: 10.1080/10864415.2016.1171977]
19. Gao GG, Greenwood BN, Agarwal RM, McCullough JS. Vocal minority and silent majority: how do online ratings reflect population perceptions of quality. MIS Q 2015 Mar 3;39(3):565-589. [doi: 10.25300/MISQ/2015/39.3.03]
20. Anderson ET, Simester DI. Reviews without a purchase: low ratings, loyal customers, and deception. J Mark Res 2014 Jun;51(3):249-269. [doi: 10.1509/jmr.13.0209]
21. Lappas T, Sabnis G, Valkanas G. The impact of fake reviews on online visibility: a vulnerability assessment of the hotel industry. Inf Syst Res 2016 Dec;27(4):940-961. [doi: 10.1287/isre.2016.0674]
22. Luca MZ, Zervas G. Fake it till you make it: reputation, competition, and Yelp review fraud. Manag Sci 2016 Dec;62(12):3412-3427. [doi: 10.1287/mnsc.2015.2304]
23. Balaji MS, Khong KW, Chong AY. Determinants of negative word-of-mouth communication using social networking sites. Inf Manag 2016 Jun;53(4):528-540. [doi: 10.1016/j.im.2015.12.002]
24. Lu N, Wu H. Exploring the impact of word-of-mouth about physicians' service quality on patient choice based on online health communities. BMC Med Inform Decis Mak 2016 Dec 26;16(1):151 [FREE Full text] [doi: 10.1186/s12911-016-0386-0] [Medline: 27888834]
25. Spence M. Job market signaling. Q J Econ 1973;87(3):355-374 [FREE Full text]
26. Connelly BL, Certo ST, Ireland RD, Reutzel CR. Signaling theory: a review and assessment. J Manag 2010 Dec 20;37(1):39-67. [doi: 10.1177/0149206310388419]
27. Cai S, Lin X, Xu DF, Fu X. Judging online peer-to-peer lending behavior: a comparison of first-time and repeated borrowing requests. Inf Manag 2016 Nov;53(7):857-867. [doi: 10.1016/j.im.2016.07.006]
28. Bliege Bird R, Smith E. Signaling theory, strategic interaction, and symbolic capital. Curr Anthropol 2005 Apr;46(2):221-248. [doi: 10.1086/427115]
29. Fischer E, Reuber R. The good, the bad, and the unfamiliar: the challenges of reputation formation facing new firms. Entrep Theory Pract 2007 Jan;31(1):53-75. [doi: 10.1111/j.1540-6520.2007.00163.x]

XSL•FO
RenderX

30. 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. [Medline: 11392867]
31. Sun Y, Fang Y, Lim KH. Understanding sustained participation in transactional virtual communities. Decis Support Syst 2012 Apr;53(1):12-22. [doi: 10.1016/j.dss.2011.10.006]
32. Gagné M, Deci EL. Self-determination theory and work motivation. J Organiz Behav 2005 Apr 14;26(4):331-362. [doi: 10.1002/job.322]
33. Haodf. URL: https://www.haodf.com/ [accessed 2019-02-10] [WebCite Cache ID 7661RvCi3]
34. Lin Z. An empirical investigation of user and system recommendations in e-commerce. Decis Support Syst 2014 Dec;68:111-124. [doi: 10.1016/j.dss.2014.10.003]

Abbreviations

GLS: generalized least squares
OLS: ordinary least squares

Edited by G Eysenbach; submitted 08.09.18; peer-reviewed by Z Deng, F Guo, D Li, P Jiang, Y Li; comments to author 18.10.18; revised version received 26.11.18; accepted 07.01.19; published 22.03.19.

Please cite as:
Yang H, Zhang X
Investigating the Effect of Paid and Free Feedback About Physicians' Telemedicine Services on Patients' and Physicians' Behaviors: Panel Data Analysis
J Med Internet Res 2019;21(3):e12156
URL: http://www.jmir.org/2019/3/e12156/
doi:10.2196/12156
PMID:30900997

©Hualong Yang, Xiaofei Zhang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Reporting of Patient Experience Data on Health Systems' Websites and Commercial Physician-Rating Websites: Mixed-Methods Analysis

Tara Lagu[1*], MPH, MD; Caroline M Norton[2*], BS; Lindsey M Russo[3*], MS; Aruna Priya[1,3*], MA, MSc; Sarah L Goff[1,3*], MD; Peter K Lindenauer[1*], MSc, MD

[1]Baystate Health, Institute for Healthcare Delivery and Population Science and Department of Medicine, University of Massachusetts Medical School-Baystate, Springfield, MA, United States

[2]College of Natural Sciences, University of Massachusetts, Amherst, MA, United States

[3]School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA, United States

[*]all authors contributed equally

Corresponding Author:
Tara Lagu, MPH, MD
Baystate Health
Institute for Healthcare Delivery and Population Science and Department of Medicine
University of Massachusetts Medical School-Baystate
3601 Main St
Springfield, MA, 01199
United States
Phone: 1 4137947688
Fax: 1 4147940300
Email: lagutc@gmail.com

Abstract

Background: Some hospitals' and health systems' websites report physician-level ratings and comments drawn from the Consumer Assessment of Healthcare Providers and Systems surveys.

Objective: The aim was to examine the prevalence and content of health system websites reporting these data and compare narratives from these sites to narratives from commercial physician-rating sites.

Methods: We identified health system websites active between June 1 and 30, 2016, that posted clinician reviews. For 140 randomly selected clinicians, we extracted the number of star ratings and narrative comments. We conducted a qualitative analysis of a random sample of these physicians' narrative reviews and compared these to a random sample of reviews from commercial physician-rating websites. We described composite quantitative scores for sampled physicians and compared the frequency of themes between reviews drawn from health systems' and commercial physician-rating websites.

Results: We identified 42 health systems that published composite star ratings (42/42, 100%) or narratives (33/42, 79%). Most (27/42, 64%) stated that they excluded narratives deemed offensive. Of 140 clinicians, the majority had composite scores listed (star ratings: 122/140, 87.1%; narrative reviews: 114/140, 81.4%), with medians of 110 star ratings (IQR 42-175) and 25.5 (IQR 13-48) narratives. The rating median was 4.8 (IQR 4.7-4.9) out of five stars, and no clinician had a score less than 4.2. Compared to commercial physician-rating websites, we found significantly fewer negative comments on health system websites (35.5%, 76/214 vs 12.8%, 72/561, respectively; P<.001).

Conclusions: The lack of variation in star ratings on health system sites may make it difficult to differentiate between clinicians. Most health systems report that they remove offensive comments, and we notably found fewer negative comments on health system websites compared to commercial physician-rating sites.

(J Med Internet Res 2019;21(3):e12007) doi:10.2196/12007

KEYWORDS

physician reviews; social networking; public reporting

XSL•FO
RenderX

Introduction

Approximately 60% of US consumers report that online reviews are either somewhat or very important when choosing a physician [1]. Of patients who have used reviews to choose a physician, 52% report that they have chosen not to see a given physician because of review content [2]. However, commercial physician-rating websites (designed similarly to websites that review restaurants and hotels) are difficult to use and have few reviews per physician [3,4]. There is an ongoing effort by leaders in the field to systematically collect and publicly report patient narratives about individual physicians, but the potential for widespread implementation of these initiatives remains unclear [5-10].

In response to this gap and as a mechanism to increase market share, some hospitals and health systems across the United States have begun to compile and report physician-level ratings and comments drawn from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys [11]. Typically, a health system engaging in such an effort summarizes CAHPS data through a composite score (hereafter called a "star rating") and posts this score on the physician's biographical webpage within the health system website. Many health systems also post patients' narrative responses to open-ended questions (eg, "What did this clinician do well?" and "What could this clinician do better?"). Although these initiatives have received attention in both the medical and lay press [11,12], there has been only one description of the phenomenon in the medical literature [13]. The number of US health systems that are participating in these efforts is unknown, the content of reviews on health systems' sites has not been described (nor has it been compared to the previously existing narrative content about physicians reviews on commercial rating websites), and the implications for patient experience and quality improvement activities have yet to be explored. Therefore, we aimed to characterize the content of health systems' webpages that report these results, including numbers of star ratings and narrative reviews per clinician, and to compare the content of narrative comments drawn from commercial physician-rating websites.

Methods

Data Sources, Website Identification, Search Characteristics, and Hospital Characteristics

The organizations that first published compiled patient experience data were large systems associated with a hospital or hospitals [12]. Therefore, we obtained data from the Center for Medicare and Medicaid Services' Hospital Compare website and examined the websites of all listed hospitals [14]. For each entity, we verified the name and street address and then determined if reviews were present. To identify participating health systems that were not associated with a hospital, we obtained a published online list of health systems [14] known to post reviews. We examined all sites on this list and supplemented our search using a previously described method for systematically searching Google (eg, "doctor reviews") [4]. We included websites that were functional (ie, had working links) between June 1 and 30, 2016, and had at least one clinician with star ratings or narrative comments. We used American Hospital Association data to generate hospital descriptive statistics (characteristics of health systems without a participating hospital were not captured). Because all included data were publicly available, the Baystate Institutional Review Board deemed that this study did not constitute human subjects research.

Website Structure and Extraction of Physician-Review Data

Using an extraction method described previously, we created an a priori list of website classification criteria to describe included websites [4]. In brief, this included elements such as methods that could be used to search for clinicians (eg, specialty, name, location), description of methods used to remove "offensive" reviews, and review structure (eg, star ratings vs narratives). Three authors (TL, CN, LR) then reviewed the content of the websites and added classification categories as appropriate. Two authors then completed a final review of included websites (CN, LR).

To examine a sample of clinician reviews, we obtained lists of clinicians for 14 of the identified health systems from the National Research Corporation (NRC) [15]. NRC is a for-profit consulting firm focused on improving patient experience and health system brand loyalty. We limited our sample to health systems for which we could obtain lists of clinicians because without a list we were unable to randomly select clinicians. We used random number generation to select 10 clinicians from each of the 14 identified health systems. For each clinician, we then extracted the number of star ratings, number of narrative comments, and total or average star rating. We quantified occurrences of each type of review using descriptive statistics (frequencies and percentages).

Quantitative and Qualitative Analysis of Narrative Reviews

We used qualitative methods to examine narrative reviews for included physicians. We selected the five most recent narratives from each of the 140 (10 clinicians from 14 health systems) randomly selected clinicians' profiles (if there were fewer than five narratives, we took as many as were present). Beginning with themes described in related studies [4,16,17], we created a codebook. We then developed additional codes to capture themes and content that were not in the codebook. We repeated this iteratively until the team felt that the coding categories captured the major substantive content reviewed. After establishing 80% agreement, the researchers each completed coding independently (all reviews were double-coded) and met once more to reach an agreement on all codes, resolving differences by consensus. Applying directed qualitative content analysis methods [18,19], we then organized codes into pertinent major and minor themes. One author (CN) then checked for accuracy of coding and performed second-level coding to synthesize themes and content, which was reviewed with the other authors. We used descriptive statistics to describe the frequency with which major and minor themes occurred.

XSL•FO
RenderX

Comparison of Narrative Reviews From Health System Websites to Those From Commercial Physician-Rating Websites

In a prior study of 600 physicians selected from three geographically diverse US cities, we collected more than 1800 narrative reviews from 28 different commercial physician-rating websites [3]. We conducted a simultaneous qualitative analysis of a randomly sampled set of 214 comments taken from these 28 commercial physician-rating sites and compared the results to those obtained from our qualitative analysis of comments from health systems' websites. Using the codebook created for the analysis of reviews from health systems, the two coders coded commercial website reviews independently and then met to discuss discrepancies in coding. Investigators resolved differences in coding and updated the codebook using an iterative process. We continued this process until no new codes were identified in 10 sequential reviews, resulting in a comprehensive codebook that covered both commercial rating websites and health systems (Multimedia Appendix 1: codebook) and a comprehensive list of themes (Multimedia Appendix 2: themes). Two investigators (CN, LR) then independently coded the remaining reviews. We compared the percentage of reviews for each theme between health system websites and commercial rating websites using the chi-square test and the Fisher exact test. All analyses were performed using SAS version 9.4 (SAS Institute Inc, Cary, NC, USA).

Results

Website Identification, Search Characteristics, and Hospital Characteristics

From 4800 hospitals on Hospital Compare, we identified 161 hospitals (3.4%) that posted star ratings or narrative comments about clinicians. Many of these hospitals were affiliated hospitals within a larger system, so we collapsed the 161 hospitals into 36 health systems. Our search methods identified an additional eight health systems that were not associated with hospitals, which gave us a total of 42 health systems from 26 states. This represented approximately 7% of the 626 health systems in the United States [20]. All identified health systems published star ratings (42/42, 100%) and most published narrative reviews (33/42, 79%).

Website Structure

No sites described their method for calculating star ratings. The majority (27/42, 64%) stated on their main page that they excluded narratives deemed inappropriate or offensive, but none explained how this process was conducted. Most allowed users to search for physicians by name (39/42, 93%), specialty (41/42, 98%), and location (31/42, 74%). Nearly half of included hospitals (79/169, 46.8%) had fewer than 200 beds (Table 1). More than a third (61/169, 36.1%) were located in the western region of the United States. Acute care hospitals made up the majority (145/169, 85.8%) of the sample.

Quantity of Reviews

Of the randomly sampled 140 clinicians from 14 health systems, there were 21,332 quantitative reviews and 4723 narrative reviews. A majority of clinicians had reviews (star ratings: 122/140, 87.1%; narrative reviews: 114/140, 81.4%), with medians of 110 star ratings (interquartile range [IQR] 42-175) and 25.5 narratives (IQR 13-48) per clinician. Only one clinician in the sample did not have any reviews. In general, star ratings were quite high with little variation between physicians: the median rating was 4.8 (IQR 4.7-4.9) out of five stars. Of 140 physicians, none had a score of less than 4.2.

Quantitative and Qualitative Analysis of Narrative Reviews

Using the five (or less if five were not available) most recent reviews from 140 clinicians from 14 health systems, we identified 561 health system narrative reviews for qualitative analysis. As described in Methods, we also analyzed 214 narrative comments previously randomly sampled from 600 physicians across 28 commercial physician-rating websites.[3] Themes that emerged from coding these two sets of data included general positive and negative comments about clinicians, clinician communication and interpersonal skills, technical skills, facility and office experience, patient care experience (independent of these other themes), descriptions of "reasons for seeking care," and "extreme comments" (ie, long descriptions of very positive or negative experiences that did not fit well into other categories). Example quotes from these themes are given in Multimedia Appendix 3 (example quotations for identified themes).

Comparison of Narrative Reviews From Health System Websites to Those From Commercial Physician-Rating Websites

Overall, the vast majority of comments were positive (642/775, 82.8% for all narratives), including 71.0% (152/214) of commercial rating websites comments and 87.3% (490/561) of health systems' websites comments (P<.001). Negative comments were less common, but commercial rating sites had a greater proportion of negative reviews compared to health systems' sites (35.5%, 76/214 vs 12.8%, 72/561, respectively; P<.001) (Table 2). Within subcategories of positive comments, there were some significant differences between the two types of websites. Commercial rating websites had significantly more "clinician communication and personal skills" positive comments (commercial rating sites: 127/214, 59.3%; health systems' sites: 238/561, 42.4%; P<.001), more positive "clinician technical skills" comments (commercial rating sites: 74/214, 34.6%; health systems' sites: 84/561, 15.0%; P<.001), and more "extremely positive" comments (commercial rating sites: 9/214, 4.2%; health systems' sites: 4/561, 0.7%; P=.002) compared to health system websites, whereas health system websites had significantly more positive "patient care experience" comments than commercial websites (commercial rating sites: 12/214, 5.6%; health systems' sites: 106/561, 18.9%; P<.001).

XSL•FO
RenderX

Table 1. Characteristics of 169 hospitals posting physician reviews.

Characteristic	Hospitals, n (%)
Number of beds[a]	
<200	79 (46.8)
200 to 400	43 (25.4)
>400	25 (14.8)
Teaching[a]	20 (11.8)
Region[a]	
Midwest	25 (14.8)
Northeast	27 (16.0)
South	34 (20.1)
West	61 (36.1)
Hospital type[b]	
Acute care	145 (85.8)
Children's	2 (1.2)
Critical access	15 (8.9)

[a]22 of 169 hospitals did not have an American Hospital Association identifier.

[b]7 of 169 hospitals did not have their type/ownership recorded.

XSL•FO
RenderX

Table 2. Comparison of occurrence of themes between health systems' sites and commercial rating sites.

Themes	Total, n (%)	Commercial rating sites, n (%)	Health systems' sites, n (%)	*P* value[a]
Sites	775 (100)	214 (27.6)	561 (72.4)	
Positive themes				
Overall positive comments	642 (82.8)	152 (71.0)	490 (87.3)	<.001
General positive comments about clinicians (great doctor, very good, would recommend, satisfied with care from provider)	345 (44.5)	89 (41.6)	256 (45.6)	.31
Clinician communication and interpersonal skills	365 (47.1)	127 (59.3)	238 (42.4)	<.001
Clinician technical skills	158 (20.4)	74 (34.6)	84 (15.0)	<.001
Facility or office experience and staff characteristics	127 (16.4)	31 (14.5)	96 (17.1)	.38
Patient care experience	118 (15.2)	12 (5.6)	106 (18.9)	<.001
Reason for seeking care	5 (0.6)	2 (0.9)	3 (0.5)	.62[b]
Extremely positive	13 (1.7)	9 (4.2)	4 (0.7)	.002[b]
Negative themes				
Overall negative comments	148 (19.1)	76 (35.5)	72 (12.8)	<.001
General negative comments about clinicians (would not recommend)	12 (1.5)	11 (5.1)	1 (0.2)	<.001[b]
Clinician communication and interpersonal skills	72 (9.3)	40 (18.7)	32 (5.7)	<.001
Technical skills	35 (4.5)	27 (12.6)	8 (1.4)	<.001
Facility or office experience and staff characteristics	66 (8.5)	38 (17.8)	28 (5.0)	<.001
Patient care experience	24 (3.1)	16 (7.5)	8 (1.4)	<.001
Feedback about survey	7 (0.9)	0 (0)	7 (1.2)	—
Extremely negative	13 (1.7)	13 (6.1)	0 (0)	—
Neutral themes				
Neutral patient experience	86 (11.1)	22 (10.3)	64 (11.4)	.65

[a]All chi-square tests, except where noted.

[b]Fisher exact test.

In contrast, commercial websites had a higher percentage of negative comments across nearly all themes. For example, commercial rating websites had more negative "clinician communication and interpersonal skills" comments (commercial rating sites: 40/214, 18.7%; health systems' sites: 32/561, 5.7%; *P*<.001), "technical skills" comments (commercial rating sites: 27/214, 12.6%; health systems' sites: 8/561, 1.4%; *P*<.001), more "facility/office experience and staff characteristics" comments (commercial rating sites: 38/214, 17.8%; health systems' sites: 28/561, 5.0%; *P*<.001), "patient care experience" comments (commercial rating sites: 16/214, 7.5%; health systems' sites: 8/561, 1.4%; *P*<0.001), and more "extremely negative" comments (commercial rating sites: 13/214, 6.1%; health systems' sites: 0/561, 0.0%) compared to health systems' websites.

Discussion

The phenomenon of health systems publishing systematically collected patient experience data about individual clinicians has been hailed as a triumph for transparency and patient-centeredness, but the scope and content of the narratives and reviews on health systems' sites have not been previously described [11,12]. After a comprehensive search, we identified 42 of 626 health systems nationwide (7%) that were early adopters of this practice. Most clinicians' pages had many reviews (both star reviews and narratives), which gives them an important advantage over existing commercial physician-rating websites [3]. However, most clinicians also had near-perfect star ratings, with 75% of physicians having a score of between 4.7 and 4.9 stars out of 5, and a minimum star rating of 4.2. Furthermore, we observed, similar to published literature [4,21], that narratives from both commercial rating websites and health systems' websites were mostly positive (with similar percentages of positive responses across both types of sites), and themes that emerged were similar to themes seen in other studies that have examined review content [4,22,23]. Across a range of subthemes, we observed a statistically significantly smaller number of negative reviews for individual

XSL•FO
RenderX

clinicians on health system sites when compared to commercial physician-rating sites.

There are several important implications to our findings. First, the narrow range of star ratings (the majority were in the 4.7 to 4.9 out of maximum 5) may limit the ability of patients to differentiate between clinicians using only star ratings on health systems' websites. The tightly clustered distribution of scores near 5.0 may be the result of the fact that health systems calculate the composite score from the CAHPS multiquestion survey, and the majority of responses to CAHPS questions are either "usually" or "always" [5]. Less than 5% of respondents choose "never" for any CAHPS category [5]. It has also been previously reported in the literature that most online ratings for physicians are quite high [24]. Notably, we did not find any health systems that explained their methods for calculating the composite metric of "number of stars."

Second, the large number of reviews and narratives on health systems' websites (medians of 110 star ratings and 25 narratives for each physician) may address the gap between patients' desire for reviews [1,2] and the previously reported lack of narrative and star reviews on physician-rating websites.[3,4]. Further, the large numbers of reviews we found may address some of the concerns about lack of an adequate sample (because health systems can also wait to post reviews until a sufficient number are collected or can use older surveys to increase the numbers of reviews and narratives) [12,25]. However, it has been reported that patients use reviews from different sources differently, and this makes it difficult to extrapolate how patients will view star ratings and comments posted by health systems [26]. At least one recent study suggests that patients trust commercial rating sites more than health systems' sites [27].

One potential driver of this lack of trust is the finding that most (64%) health systems we examined stated on their main page that they remove "offensive or inappropriate content"; we also observed a relative lack of negative reviews compared to commercial physician-rating sites [12,25]. The finding that health system sites have more ratings, fewer negative ratings, and fewer negative comments compared to commercial physician-rating websites is consistent with the only published report that compared the two sources [13]. Despite the obvious possibility that negative comments are being removed by system administrators, there are several potential explanations for the observed differences. Removal of some comments is appropriate if the content is genuinely offensive [27], and patients may take a different approach to systematic surveys compared to open online platforms (which include reviews and comments only from respondents who seek out the site).

Our findings suggest that health systems' websites have the potential to provide patients with information about the experience of care with clinicians, but the sites may require improvements on this first iteration. One addition that could improve the narrative content is posting of a protocol for curating patient narratives and calculating star ratings [5-10]. Given the narrow distribution of the star ratings, a posting of the range of all physician scores with an indication where each physician's score is situated would also be helpful (eg, if 4.2 is the lowest score, the patient would know this by seeing where the physician fell within the distribution). Although these improvements would provide patients with a more comprehensive picture of the experience of care with physicians within a given health system, it is also possible that health systems have conflicts of interest (specifically, an interest in increasing market share) that would discourage them from making these changes [11-12]. Another area that warrants further investigation—but about which we have limited information—is the possibility that the process of collecting and publishing the patient experience data has led physicians, practices, and even entire systems to initiate improvement activities based on patient comments [28,29].

This study has several limitations. First, we made extensive efforts to identify all health systems in the United States that are posting reviews of clinicians but may have missed some sites. Second, this is a snapshot of a single point in time, and the number of health systems participating in these efforts has likely changed in the interim. Third, we had a limited sample from which we drew clinician reviews for analysis because of the lack of lists of clinicians for most sites. However, we have no reason to believe that the health systems we sampled were different than the remaining health systems in our study. Finally, we were limited by an inability to assess the impact of these sites on clinicians or prospective patients.

Given the amount of public interest in narrative and quantitative data on individual clinicians, we anticipate that the trend of health systems publishing this information will continue. However, before health systems' websites emerge as the main route by which consumers look for information about prospective clinicians, there may be a need to improve their methods used for curating and posting patient experience data.

Acknowledgments

Dr Lagu is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under award numbers K01HL114745 and R01 HL139985-01A1. Dr Goff is funded by the National Institute for Child Health and Human Development of the National Institutes of Health under K23HD080870.

Authors' Contributions

All authors had access to the data and wrote portions of the manuscript and reviewed the manuscript for content.

XSL•FO
RenderX

Conflicts of Interest

Drs Lagu and Lindenauer have served as consultants for the Yale Center for Outcomes Research and Evaluation, under contract to the Centers for Medicare and Medicaid Services, for which they have provided clinical and methodological expertise and input on the development, reevaluation, and implementation of hospital outcome and efficiency measures. The views expressed in this manuscript do not necessarily reflect those of the Yale Center for Outcomes Research and Evaluation or the Centers for Medicare and Medicaid Services.

Multimedia Appendix 1

Codebook.

[PDF File (Adobe PDF File), 35KB - jmir_v21i3e12007_app1.pdf]

Multimedia Appendix 2

Themes.

[PDF File (Adobe PDF File), 407KB - jmir_v21i3e12007_app2.pdf]

Multimedia Appendix 3

Example Quotations for Identified Themes.

[PDF File (Adobe PDF File), 109KB - jmir_v21i3e12007_app3.pdf]

References

1. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. JAMA 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
2. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. J Med Internet Res 2013;15(8):e187 [FREE Full text] [doi: 10.2196/jmir.2702] [Medline: 23985220]
3. Lagu T, Metayer K, Moran M, Ortiz L, Priya A, Goff SL, et al. Website characteristics and physician reviews on commercial physician-rating websites. JAMA 2017 Feb 21;317(7):766. [doi: 10.1001/jama.2016.18553]
4. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
5. Hays RD, Shaul JA, Williams VS, Lubalin JS, Harris-Kojetin LD, Sweeny SF, et al. Psychometric properties of the CAHPS 1.0 survey measures. Consumer Assessment of Health Plans Study. Med Care 1999 Mar;37(3 Suppl):MS22-MS31. [Medline: 10098556]
6. Giordano LA, Elliott MN, Goldstein E, Lehrman WG, Spencer PA. Development, implementation, and public reporting of the HCAHPS survey. Med Care Res Rev 2010 Feb;67(1):27-37. [doi: 10.1177/1077558709341065] [Medline: 19638641]
7. Grob R, Schlesinger M, Parker AM, Shaller D, Barre LR, Martino SC, et al. Breaking narrative ground: innovative methods for rigorously eliciting and assessing patient narratives. Health Serv Res 2016 Dec;51 Suppl 2:1248-1272 [FREE Full text] [doi: 10.1111/1475-6773.12503] [Medline: 27126144]
8. Schlesinger M, Grob R, Shaller D, Martino SC, Parker AM, Finucane ML, et al. Taking patients' narratives about clinicians from anecdote to science. N Engl J Med 2015 Aug 13;373(7):675-679. [doi: 10.1056/NEJMsb1502361] [Medline: 26267629]
9. Schlesinger M, Kanouse DE, Martino SC, Shaller D, Rybowski L. Complexity, public reporting, and choice of doctors: a look inside the blackest box of consumer behavior. Med Care Res Rev 2014 Oct;71(5 Suppl):38S-64S. [doi: 10.1177/1077558713496321] [Medline: 23999489]
10. Schlesinger M, Kanouse DE, Rybowski L, Martino SC, Shaller D. Consumer response to patient experience measures in complex information environments. Med Care 2012 Nov;50 Suppl:S56-S64. [doi: 10.1097/MLR.0b013e31826c84e1] [Medline: 23064278]
11. Lee TH. Harvard Business Review. 2014 Jun 03. Online reviews could help fix medicine URL: https://hbr.org/2014/06/online-reviews-could-help-fix-medicine [accessed 2019-01-25] [WebCite Cache ID 75h2wEPDE]
12. Lee V. Transparency and trust-online patient reviews of physicians. N Engl J Med 2017 Jan 19;376(3):197-199. [doi: 10.1056/NEJMp1610136] [Medline: 28099823]
13. Ricciardi BF, Waddell BS, Nodzo SR, Lange J, Nocon AA, Amundsen S, et al. Provider-initiated patient satisfaction reporting yields improved physician ratings relative to online rating websites. Orthopedics 2017 Sep 01;40(5):304-310. [doi: 10.3928/01477447-20170810-03] [Medline: 28817163]
14. US Centers for Medicare & Medicaid Services. Medicare.gov. 2010. Hospital Compare URL: https://www.medicare.gov/hospitalcompare/search.html? [accessed 2018-04-19] [WebCite Cache ID 75h7Jr6fW]

XSL•FO
RenderX

15. National Research Council. NRC Health. 2018. URL: https://nrchealth.com/ [accessed 2018-02-12] [WebCite Cache ID 75h3w9OuR]
16. Lagu T, Kaufman EJ, Asch DA, Armstrong K. Content of weblogs written by health professionals. J Gen Intern Med 2008 Oct;23(10):1642-1646 [FREE Full text] [doi: 10.1007/s11606-008-0726-6] [Medline: 18649110]
17. Lagu T, Goff SL, Hannon NS, Shatz A, Lindenauer PK. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. Jt Comm J Qual Patient Saf 2013 Jan;39(1):7-15. [Medline: 23367647]
18. Goff SL, Mazor KM, Gagne SJ, Corey KC, Blake DR. Vaccine counseling: a content analysis of patient-physician discussions regarding human papilloma virus vaccine. Vaccine 2011 Oct 06;29(43):7343-7349. [doi: 10.1016/j.vaccine.2011.07.082] [Medline: 21839136]
19. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
20. Agency for Healthcare Research and Quality. 2017. Compendium of US Health Systems, 2016 URL: https://www.ahrq.gov/chsp/compendium/index.html [accessed 2019-01-24] [WebCite Cache ID 75h3h6FES]
21. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012 Jun;27(6):685-692 [FREE Full text] [doi: 10.1007/s11606-011-1958-4] [Medline: 22215270]
22. Greaves F, Laverty AA, Cano DR, Moilanen K, Pulman S, Darzi A, et al. Tweets about hospital quality: a mixed methods study. BMJ Qual Saf 2014 Oct;23(10):838-846 [FREE Full text] [doi: 10.1136/bmjqs-2014-002875] [Medline: 24748372]
23. Ranard BL, Werner RM, Antanavicius T, Schwartz HA, Smith RJ, Meisel ZF, et al. Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. Health Aff (Millwood) 2016 Apr;35(4):697-705. [doi: 10.1377/hlthaff.2015.1030] [Medline: 27044971]
24. Kadry B, Chu LF, Kadry B, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. J Med Internet Res 2011;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
25. Algorithms for Innovation. Transparency: will it help or harm health care? URL: https://uofuhealth.utah.edu/innovation/algorithms/2015/three.php [accessed 2018-05-21] [WebCite Cache ID 75h3jJdZF]
26. Yaraghi N, Wang W, Gao GG, Agarwal R. How online quality ratings influence patients' choice of medical providers: controlled experimental survey study. J Med Internet Res 2018 Mar 26;20(3):e99 [FREE Full text] [doi: 10.2196/jmir.8986] [Medline: 29581091]
27. Holliday AM, Kachalia A, Meyer GS, Sequist TD. Physician and patient views on public physician rating websites: a cross-sectional study. J Gen Intern Med 2017 Feb 01;32(6):626-631. [doi: 10.1007/s11606-017-3982-5] [Medline: 28150098]
28. Lagu T, Goff SL, Craft B, Calcasola S, Benjamin EM, Priya A, et al. Can social media be used as a hospital quality improvement tool? J Hosp Med 2016 Jan;11(1):52-55 [FREE Full text] [doi: 10.1002/jhm.2486] [Medline: 26390277]
29. Emmert M, Meszmer N, Sander U. Do health care providers use online patient ratings to improve the quality of care? results from an online-based cross-sectional study. J Med Internet Res 2016 Sep 19;18(9):e254 [FREE Full text] [doi: 10.2196/jmir.5889] [Medline: 27644135]

Abbreviations

CAHPS: Consumer Assessment of Healthcare Providers and Systems

Edited by G Eysenbach; submitted 22.08.18; peer-reviewed by J Munro; comments to author 14.10.18; accepted 25.01.19; published 27.03.19.

Please cite as:
Lagu T, Norton CM, Russo LM, Priya A, Goff SL, Lindenauer PK
Reporting of Patient Experience Data on Health Systems' Websites and Commercial Physician-Rating Websites: Mixed-Methods Analysis
J Med Internet Res 2019;21(3):e12007
URL: http://www.jmir.org/2019/3/e12007/
doi:10.2196/12007
PMID:30916654

©Tara Lagu, Caroline M Norton, Lindsey M Russo, Aruna Priya, Sarah L Goff, Peter K Lindenauer. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet

XSL•FO
RenderX

Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis

Amelia Romeo[1], BOccupTher; Sarah Edney[2], BA; Ronald Plotnikoff[3], PhD; Rachel Curtis[2], PhD; Jillian Ryan[2], PhD; Ilea Sanders[2], B Human Movement; Alyson Crozier[2], PhD; Carol Maher[2], PhD

[1]School of Health Sciences, University of South Australia, Adelaide, Australia

[2]Alliance for Research in Exercise, Nutrition and Activity, School of Health Sciences, University of South Australia, Adelaide, Australia

[3]Priority Research Centre for Physical Activity and Nutrition, University of Newcastle, Newcastle, Australia

Corresponding Author:
Carol Maher, PhD
Alliance for Research in Exercise, Nutrition and Activity
School of Health Sciences
University of South Australia
GPO Box 2471
Adelaide, 5001
Australia
Phone: 61 8 830 22315
Email: carol.maher@unisa.edu.au

Abstract

Background: Smartphone apps are a promising tool for delivering accessible and appealing physical activity interventions. Given the large growth of research in this field, there are now enough studies using the "gold standard" of experimental design—the randomized controlled trial design—and employing objective measurements of physical activity, to support a meta-analysis of these scientifically rigorous studies.

Objective: This systematic review and meta-analysis aimed to determine the effectiveness of smartphone apps for increasing objectively measured physical activity in adults.

Methods: A total of 7 electronic databases (EMBASE, EmCare, MEDLINE, Scopus, Sport Discus, The Cochrane Library, and Web of Science) were searched from 2007 to January 2018. Following the Population, Intervention, Comparator, Outcome and Study Design format, studies were eligible if they were randomized controlled trials involving adults, used a smartphone app as the primary or sole component of the physical activity intervention, used a no- or minimal-intervention control condition, and measured objective physical activity either in the form of moderate-to-vigorous physical activity minutes or steps. Study quality was assessed using a 25-item tool based on the Consolidated Standards of Reporting Trials checklist. A meta-analysis of study effects was conducted using a random effects model approach. Sensitivity analyses were conducted to examine whether intervention effectiveness differed on the basis of intervention length, target behavior (physical activity alone vs physical activity in combination with other health behaviors), or target population (general adult population vs specific health populations).

Results: Following removal of duplicates, a total of 6170 studies were identified from the original database searches. Of these, 9 studies, involving a total of 1740 participants, met eligibility criteria. Of these, 6 studies could be included in a meta-analysis of the effects of physical activity apps on steps per day. In comparison with the control conditions, smartphone apps produced a nonsignificant (P=.19) increase in participants' average steps per day, with a mean difference of 476.75 steps per day (95% CI −229.57 to 1183.07) between groups. Sensitivity analyses suggested that physical activity programs with a duration of less than 3 months were more effective than apps evaluated across more than 3 months (P=.01), and that physical activity apps that targeted physical activity in isolation were more effective than apps that targeted physical activity in combination with diet (P=.04). Physical activity app effectiveness did not appear to differ on the basis of target population.

Conclusions: This meta-analysis provides modest evidence supporting the effectiveness of smartphone apps to increase physical activity. To date, apps have been most effective in the short term (eg, up to 3 months). Future research is needed to understand the time course of intervention effects and to investigate strategies to sustain intervention effects over time.

(J Med Internet Res 2019;21(3):e12053) doi:10.2196/12053

XSL•FO
RenderX

KEYWORDS

physical activity; smartphone; mobile phone; app; mobile apps; program; health behavior; systematic review; meta-analysis

Introduction

Background

Physical inactivity is the fourth largest behavioral risk factor contributing to ill health and mortality [1]. The number of adults who are physically inactive is on the rise in many countries, influencing the prevalence of noncommunicable diseases and the general health of the global population [2,3]. Current global recommendations on physical activity for health suggest adults (aged 18 to 64 years) participate in at least 150 min of moderate-intensity physical activity or 75 min of vigorous-intensity physical activity per week [4]. Participation in regular physical activity reduces the risk of noncommunicable diseases by improving muscular and cardiorespiratory fitness, functional health, and mental health [2,3,5]. However, currently, 1 in 4 adults globally do not meet physical activity recommendations [2]. As such, finding effective strategies to increase participation in regular physical activity is an essential public health objective.

Smartphone apps are being recognized as a potential and promising approach to increase adherence to physical activity guidelines. Globally, activated mobile phones outnumber citizens [6] with approximately 63% of the global adult population owning at least 1 smartphone in 2017 [7]. Smartphones are inexpensive and allow users to engage with health information technology in any environment and at any time [8]. They are equipped with advanced technological features, most notably their connection to the internet, global positioning system, and inbuilt accelerometers [9], and offer the capacity to create individualized and interactive apps that collect real-time data [10]. These features, as well as the high usage and convenience of smartphones, make them an attractive tool for researchers to deliver physical activity interventions.

Indeed, studies are increasingly using smartphone apps to try to motivate individuals to be physically active. In 2015, a systematic review and meta-analysis of 12 studies using smartphone apps to promote weight loss and increase physical activity found a nonsignificant difference in physical activity between the control group and smartphone intervention group [11]. In 2016, a systematic review of 15 studies, including both qualitative and quantitative research designs, found that smartphone apps can be effective in increasing physical activity although the effect size was modest [12]. However, these reviews are subject to limitations. For example, they only searched a small number of databases, and they included studies with self-reported physical activity data, which are susceptible to bias [13]. In addition, they included studies where smartphone apps were delivered in concert with other intervention elements (eg, face-to-face appointments and podcasts); thus, it is unclear whether intervention effects were truly due to the smartphone app itself, or rather the other intervention elements. Furthermore, the most recent review [12] included studies published until the end of 2015 only. Given the exponential growth in this field [14], the evidence base has expanded considerably since this time. Thus, it is imperative that smartphone intervention research is updated. Finally, these reviews incorporated studies varying widely in research design. Given the large growth of available research in the field, there are now likely to be enough studies using the "gold standard" of experimental design—the randomized controlled trial design—and employing objective measurement of physical activity to support a meta-analysis of scientifically rigorous studies. Examining only studies that meet this stringent methodological inclusion criteria will heighten the trustworthiness of findings.

Objectives

This systematic review aimed to (1) identify all published randomized controlled trials, which examine the efficacy of physical activity interventions delivered via smartphone apps on increasing objectively measured physical activity and (2) conduct a meta-analysis of these published studies to determine the current state of evidence regarding the effectiveness of smartphone app-based interventions for increasing physical activity in an adult population.

Methods

Information Sources and Search Strategy

This systematic review was completed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [15].

A systematic search was conducted on January 8, 2018, and included 7 electronic databases: EMBASE, EmCare, MEDLINE, Scopus, Sport Discus, The Cochrane Library, and Web of Science. The search strategy was reviewed by an academic librarian before being finalized. Each database was searched individually, and the search strategy for 1 database, EMBASE, is presented in Table 1. In brief, the search strategy combined synonyms for the intervention (mobile phone apps) with synonyms for the outcomes (physical activity and weight loss), and these terms were mapped to MeSH headings where possible. Search results were limited to the English language, humans, and year of publication from 2007 to present (on the basis that the original iPhone was released in 2007). The search strategy for all databases is presented in Multimedia Appendix 1. The reference lists of all eligible studies and of relevant systematic reviews [11,12,16,17] were screened to identify any further studies for inclusion. In addition, 5 prominent researchers in the field were contacted with the list of identified studies and asked to recommend any additional studies that met the inclusion criteria.

XSL•FO
RenderX

Table 1. Search strategy as used in EMBASE on January 8, 2018.

Search category	Search terms
Smartphones	Cell phones[a]/Smartphones[a]/Mobile Applications[a]/("smart phone*" or smartphone* or smart-phone* or "cell*phone*" or "cell-phone*" or "mobile phone*" or "mobile-phone" or "mobile device" or "mobile telephone*" or i*Phone* or android* or iOS or "mobile health" or "mhealth" or "m-health" or app or apps or "mobile application*")
Physical activity	Exercise[a]/Weight Loss[a]/("physical activit*" or exercise* or "active living" or walk* or "active transport*" or "leisure activit*" or fitness or "weight loss" or "weight reduction" or "weight maintenance" or "maintaining weight" or "weight management")
Intervention	(Intervention or program* or trial)
Combined	1 AND 2 AND 3

[a]Denotes MeSH headings.

Eligibility Criteria

Population

Study samples including healthy adults or adults with a specific health condition aged 18 years or over were eligible. Studies with participants characterized by intellectual or marked cognitive impairments or with a severe mobility disorder were excluded from this study as app designs may be specific for limited cognition and mobility.

Intervention

Studies that reported a smartphone app as the primary component to the physical activity intervention were included. The smartphone apps were required to explicitly be stand-alone apps available on a smartphone. Apps within other contexts, for example, Facebook apps, were excluded from the study. In addition, studies with interventions that incorporated secondary components such as ongoing personalized support, either in person or over the phone, were excluded.

Control or Comparator

Studies that included control groups, which received no intervention or minimal intervention (eg, given only a physical activity goal), were included.

Outcomes

For inclusion in the review, studies had to report objectively measured physical activity in the form of either moderate-to-vigorous physical activity (MVPA) min as measured by accelerometry or steps measured either by accelerometry or pedometry at baseline and follow-up.

Study Design

Only randomized controlled trials were included in this review. Conference abstracts were excluded.

Study Selection

Studies were screened for eligibility in duplicate under blinded conditions by 2 independent reviewers (selected from the author list, ie, AR, SE, RC, AC, IS, or JR) as per best practice for systematic reviews [15]. Search results were first screened based on the title and abstract, and any studies that appeared to meet the eligibility criteria, or where eligibility was unclear, progressed to full-text screening. Moreover, 2 independent reviewers then screened the full-text studies to determine eligibility for inclusion in the review. An interrater agreement of 96% (Cohen kappa= 0.53) was reached. Results from each round of screening were compared among reviewers, and conflicts were discussed until consensus was attained.

Data Collection Process

Pairs of reviewers (either AR, SE, RC, AC, or IS) independently extracted data from each included study using a standardized form developed specifically for this review (see Multimedia Appendix 2). Extracted data included country of study, study participants (population and sample size), study design (app features, description of intervention group and control group, and duration of follow-up), outcome measures (measurement tool and timing), and key study results (mean and SD for MVPA or steps at baseline and follow-up). Where the results were not adequately reported within the study paper, authors were contacted to provide additional data. Where there were discrepancies in data extraction, the author team discussed and rechecked the original study until consensus was reached.

Risk of Methodological Bias

Risk of bias was assessed using a 25-item tool developed by Maher [18] and based on the Consolidated Standards of Reporting Trials (CONSORT) checklist [19]. Each study was independently scored by 2 reviewers. Items were scored 1 if the study satisfactorily met the criteria, and 0 if the study did not satisfactorily meet the criteria or if the item was not applicable to the study. Disagreements between reviewers were resolved by checking and discussing the original study until consensus was reached. The most common disagreement related to whether items were not present or not applicable to the study, which did not have a large effect on assessment risk of bias as both were scored 0. Studies were also graded using the 2011 Centre for Evidence Based Medicine Levels of Evidence [20].

Synthesis of Results

The primary outcome measure meta-analyzed in this review was mean change in physical activity either reported in MVPA per day or steps per day. For studies that did not report mean change, it was calculated from the baseline data and follow-up (ie, the end of the intervention) data. Where studies included multiple intervention groups with similar app features, the intervention group with the most app features was included in the meta-analysis.

For the intervention and control group of each study, the mean change in physical activity from baseline to follow-up, SD of

XSL•FO
RenderX

the change, and the number of participants were entered into Review Manager software (Version 5.3, The Nordic Cochrane Centre) and used to calculate the mean difference and standardized mean difference between the change in the intervention group and the change in the control group for each study. Where data were available from 3 or more studies, a meta-analysis calculating the combined effects for all studies was performed.

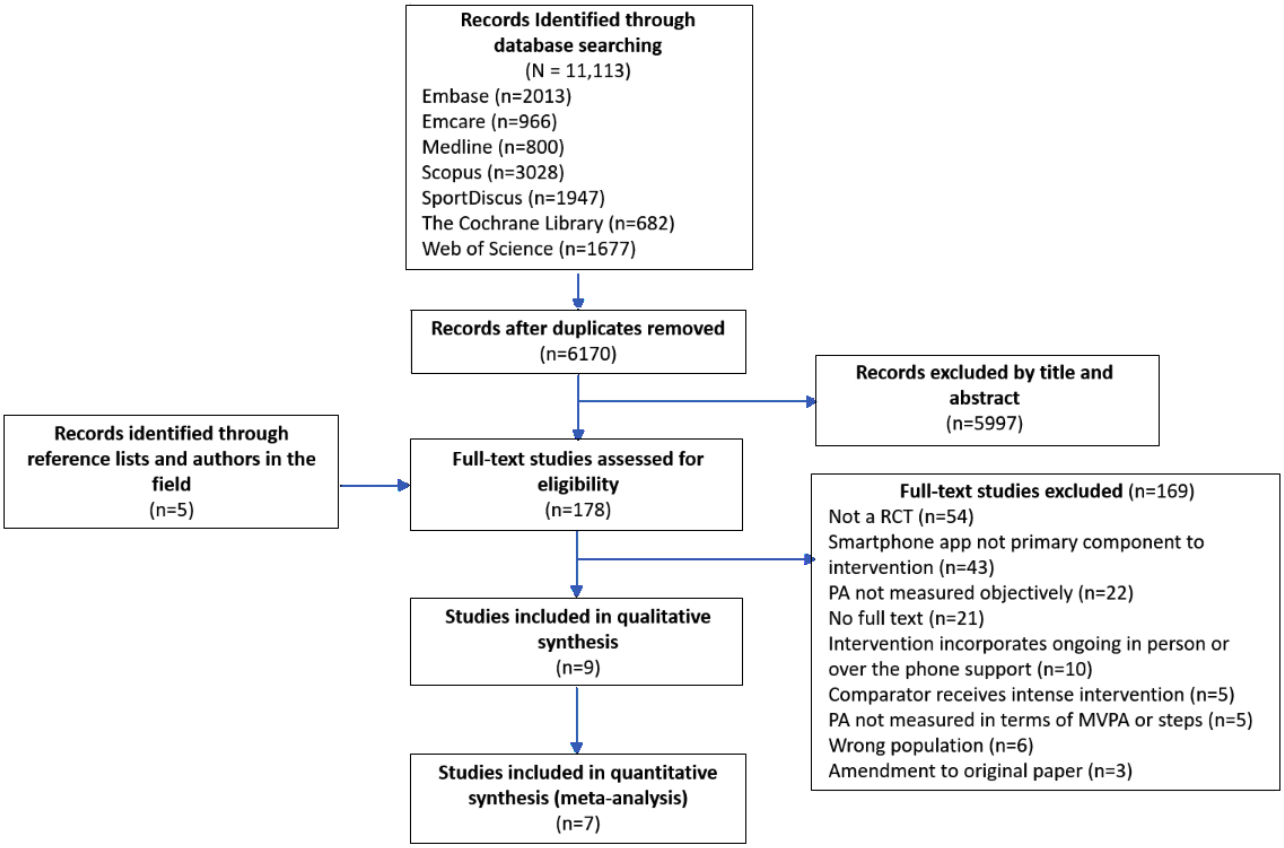
A random effects model approach was used as study heterogeneity was anticipated because of the variance of study populations and intervention designs. The presence of heterogeneity was determined using the I^2 statistic, which describes the percentage of total variation in study estimates that are a result of heterogeneity [21]. The I^2 statistic was selected as the preferred measure of variance as it is robust for small sample sizes [21]. The mean difference and standardized mean difference size were interpreted using the Cohen [22] suggestion that 0.2 is considered a *small* effect size, 0.5 a *medium* effect, and 0.8 a *large* effect, with values smaller than 0.2 being *trivial*. Furthermore, 3 sensitivity analyses were performed to further assess the robustness of our study results by excluding either studies with an intervention duration of 3 months or more (or equivalent in weeks per days) or studies with an app designed to increase physical activity as well as other nonphysical activity behaviors or studies designed for disease population groups.

Results

Study Selection

After the removal of duplicates, a total of 6170 studies were identified from the original database search. An additional 5 studies were suggested by leading authors in the field; however, none of these studies met all inclusion criteria for the systematic review. In total, 9 studies were identified to meet the eligibility criteria for inclusion in this review, and 7 of these studies were included in the meta-analysis. Moreover, 1 study [23] had unusable baseline data due to a technical malfunction in their study; this meant that the mean change in physical activity from baseline to follow-up could not be calculated, and the results were unable to be included in the meta-analysis. In addition, despite best efforts to contact the authors of the study to provide further data, 1 study [24] was excluded from the meta-analysis owing to not reporting baseline or follow-up data for the control group. The study selection process is summarized in Figure 1.

Figure 1. Flowchart for the selection of studies in this meta-analysis. PA: physical activity; MVPA: moderate-to-vigorous physical activity; RCT: randomized controlled trial.



XSL•FO
RenderX

Study Characteristics

All 9 studies reported from this review were randomized controlled trials published in English between 2014 and 2017. The key characteristics of the included studies are presented in Table 2. A total of 3 studies reported being registered with a Clinical Trial Registry [24-26]. Sample sizes ranged from 23 to 833, with a total of 1740 participants across the 9 studies. All studies had a smartphone app designed to increase physical activity as the primary intervention, although the Recio-Rodriguez et al [27] app was also designed to increase both physical activity and adherence to a Mediterranean dietary pattern. Control groups received either no intervention [27-29], a wearable accelerometer only [30,31], or a basic version of a smartphone app [23-26]. Of the 9 studies, 4 reported their app as being based on a recognized behavior-change theory; the social cognitive theory was reported in 3 [24,25,30], and principles of re-enforcement [24], social influencers' perspective [24], and taxonomy of behavior change [28] were reported in 1 study each. Physical activity data were collected using the smartphones' in-built accelerometer in 3 of the studies [23,24,26] and using a separate wearable accelerometer in 6 of the studies [25,27-31]. Intervention length ranged from 6 weeks to 6 months. A total of 6 studies reported physical activity in terms of steps per day [26-31], supporting the meta-analysis of this outcome measure. Only 2 studies reported physical activity data in terms of MVPA in sufficient detail to calculate effect sizes, and meta-analysis for this outcome measure could, therefore, not be performed.

App Features

The intervention apps included a variety of features such as a physical activity performance summary [25,27,29,31], goal setting [24,25,27,28], visual display of goal achievement [24,26,28,29], and motivational prompts [24,30,31]. The only feature common to each of the intervention apps was a visible display of steps or MVPA (see Table 2 for a detailed description of each intervention apps' features).

Risk of Bias in Included Studies

All studies were screened for risk of methodological bias using the CONSORT checklist [19] and scores varied from 15 [23,28,30] to 21 out of 25 [24]. Full details of the bias screening are presented in Multimedia Appendix 3. Each study was graded level 2 as a randomized trial using the 2011 Centre for Evidence Based Medicine Levels of Evidence [20]. All studies fulfilled the CONSORT checklist requirements to provide scientific rationale and clearly describe the intervention. The randomization procedure within most of the studies was considered adequate, with allocation concealment mechanisms detailed in all but 1 study [30]. All studies satisfied the CONSORT checklist criteria for providing primary and secondary outcomes and results for each group except for 2 studies [23,24], which were excluded from the meta-analysis.

Meta-Analysis of Smartphone App–Based Intervention and Steps Per Day

A total of 6 studies, involving a total of 1178 participants, reported change in physical activity in terms of steps per day [26-31]. In comparison with the control (the nature of which varied between studies), the intervention conditions showed a nonsignificant (P=.19) increase in average steps per day, with a mean difference between groups of 476.75 steps per day (95% CI −229.57 to 1183.07; see Figure 2). The standardized mean difference between the control and intervention groups was small in magnitude and favored the intervention group (0.21; 95% CI −0.07 to 0.50; see Figure 3), yet was not statistically significant (P=.14) . The impact of heterogeneity within the studies was significant (I^2=72%), suggesting that 72% of variation across studies was due to heterogeneity rather than chance [32]. Thus, the meta-analysis results should be interpreted with caution.

XSL•FO
RenderX

Table 2. Data extraction characteristics of included studies.

Study	Study population/sample size	Smartphone app features	Intervention description	Outcome measures
Choi et al 2016 [28]	Sample size: Total n=30, intervention n=15, control n=15; Population: Pregnant women between 10 and 20 weeks of gestation with a sedentary lifestyle. Age: Mean 33.7 (SD 2.6); Male (%): 0; Country: United States of America; Attrition rate: 3%	Characteristics: Visual display of steps, distance, flights of stairs climbed, and estimated calories expended. A daily message prompt to support PA[a] was available between 10 am and 7 pm, participants were able to respond to the message and receive feedback. Activity diary available after 7 pm each night. Theory: Social cognitive theory	Focus: Physical activity; Groups: (Intervention) Participants wore a Fitbit and had access to all features within a specifically designed smartphone app. Participants were given the goal of increasing their step count by 10% each week until 8500 steps a day was reached. (Control) Participants wore a Fitbit and were given the goal to increase steps gradually until 8500 steps a day was reached. Additional to app: One face-to-face goal setting session, information provided on healthy diets and recommendations for gestational weight gain, and safety instructions for PA during pregnancy. Duration: 12 weeks; Follow-up post baseline: Weeks 4, 8, and 12.	Primary outcome: Steps per day; Measuring tool: Fitbit Ultra Accelerometer; Secondary outcomes: TV/computer time, self-efficacy, barriers, social support, depressive symptoms, and pregnancy symptoms.
Fanning et al 2017 [30]	Sample size: Total n=116, Group A n=29, Group B n=31, Group C n=26, Group D n=30; Population: Low-active adults; Age: Mean 41.4 (SD 7.6); Male (%): 20; Country: United States of America; Attrition rate: 17%	Characteristics: Tracking of activities, instant feedback on weekly progress, weekly education modules within the app. Guided goal setting module with goals tied within all app features. Points system module with points provided for all in app tasks and accumulated to earn badges. Theory: Social cognitive theory	Focus: Physical activity; Groups: (A) goal setting module and point-based feedback module, (B) goal setting module, (C) point based feedback module, (D) standard app. Additional to app: Text messages reminding participants to goal set and be active and track activities. Support emails at the beginning of each week. Counseling on SMART[c] goal setting (groups A and B). Duration: 12 weeks; Follow-up post baseline: 12 weeks	Primary outcome: Mean daily minutes of MVPA[b]. Measuring tool: Actigraph accelerometers (model GT1 M or newer); Secondary outcomes: self-efficacy, perceived barriers, outcome expectations, goals, use, and usability.
Glynn et al 2014 [31]	Sample size: Total n=90, intervention n=45, control n=45; Population: Existing Android smartphone users; Age: Mean 44.1 (SD 11.5); Male (%): 36; Country: Ireland; Attrition rate: 14%	Characteristics: Automatic feedback and tracking of step count and calories burned, visually appealing display of step count history, and goal achievement. Theory: N/A[d]	Focus: Physical activity; Groups: (Intervention) Access to app, instruction to interact with app and goal of 10,000 steps per day; (control) goal of 30 min activity per day, access to app without visible tracking or display. Additional to app: Physical activity goals, information on the benefits of exercise, and physical activity promotion brochure. Duration: 8 weeks; Follow-up post baseline: Weeks 2-8	Primary outcome: Steps per day; Measuring tool: Accelerometer within smartphone and share data function of the app. Secondary outcomes: Mean systolic blood pressure, mean diastolic blood pressure, mean resting heart rate, weight and body mass index, mental health quality of life, and quality of life.

XSL•FO
RenderX

Study	Study population/sample size	Smartphone app features	Intervention description	Outcome measures
Harries et al 2016 [23]	Sample size: Total n=165, intervention (group 1) n=55, intervention (group 2) n=55, control n=55; Population: Males with an existing mobile phone contract; Age range: 22-40 years; Male (%): 100; Country: United Kingdom; Attrition rate: 8%	Characteristics: Step count, steps taken, miles walked, and calories burned for the day and previous week viewable. Social feedback group were able to view their average step count in comparison with other users' average step counts. Theory: N/A	Focus: Physical activity; Groups: (Intervention group 1) App with feedback on step counts; (intervention group 2) App with feedback on step counts plus social comparison; (control) basic app, no feedback or social features. Additional to app: Standardized text-messages in the first 2 weeks to remind participants to carry their phones in their pockets. Intervention groups received weekly messages to encourage them to walk more. Duration: 6 weeks; Follow-up post baseline: N/A	Primary outcome: Steps per day; Measuring tool: Accelerometer within mobile phone. Secondary outcomes: N/A
King et al 2016 [24]	Sample size: Total n=89, affect group n=22, analytic group n=21, social group n=22, control n=24; Population: Underactive adults aged 45+; Age: Mean 60.0 (SD 9.3); Male (%): 24.7; Country: United States of America; Attrition rate: 6%	Characteristics: (Analytic app) goal-setting, behavioral feedback, tips promoting behavior change, and problem-solving strategies with 2 colorful meters showing progress toward MVPA and sedentary behavior goals; (Social app) social support for behavior change, "just-in-time" social normative feedback, modelling of behaviors by others using avatars on the display, and group-based collaboration and competition "virtual teams"; (affect app) utilized an avatar bird to mirror how active or sedentary the user was throughout the day. The bird avatar changed position, posture, and movement depending on how active/inactive user was. Users received "rewards" as PA levels increased; Theory: (Analytic) social cognitive theory; (social) social influence perspectives; (affect) principles of reinforcement scheduling and attachment, and nurturance motives.	Focus: Physical activity and sedentary behavior; Groups: (analytic) Access to analytic app; (social) access to social app; (affect) access to affect app; (control) access to commercially accessible nonphysical activity dietary app (calorific). Additional to app: Initial 1-hour training on how to use the smartphone app. Duration: 8 weeks; Follow-up post baseline: Weeks 2-8	Primary outcome: Mean daily mins of MVPA; Measuring tool: Accelerometer within smartphone. Secondary outcomes: estimated minutes of sedentary time, self-reported sociological momentary assessment of daily brisk walking and sitting time.

XSL•FO
RenderX

Study	Study population/sample size	Smartphone app features	Intervention description	Outcome measures
Paul et al 2016 [26]	Sample size: Total n=23, intervention n=15, control n=8; Population: Stroke survivors who have had a single unilateral stroke and can walk independently with or without an aid; Age: Mean 55.8 (SD 10.7); Male (%): 48; Country: United Kingdom (Scotland); Attrition rate: 4%	Characteristics: Step count, goal-setting, planning, monitoring, and feedback, as well as rewards and social facilitation. Within the app, participants are represented by a fish within a fish tank. The fish swims and blows bubbles when the participant is active (which other participants can see). Fish fins and tail grow when targets are achieved. Theory: Taxonomy of behavior change	Focus: Physical activity; Groups: (Intervention) Received a smartphone with the STARFISH APP, individual step goals which increased by 5% each week if participants reached their step goal on 5 of 7 days. Individual and group rewards provided when goals achieved; (control) usual care after stroke (no active rehabilitation); Additional to app: (Intervention) face-to-face at week 3 to discuss progress with Clinical Research Facility; Duration: 6 weeks; Follow-up post baseline: 6 Weeks	Primary outcome: Steps per day; Measuring tool: ActivPAL accelerometer; Secondary outcomes: Sedentary time, heart rate, blood pressure, body mass index, fatigue severity scale, instrumental activity of daily living scale, 10-meter walk test, stroke specific quality of life scale, and psychological general well-being index.
Recio-Rodriguez et al 2016 [25]	Sample Size: Total n=833, intervention n=415, control n=418; Population: Selected from the Multicenter Assessment of Experimental Program Promoting Physical Activity; Age: intervention mean 51.4 (SD 12.1); control mean 52.3 (SD 12); Male (%): intervention 40; control 36; Country: Spain; Attrition rate: 16%	Characteristics: Automatic feedback from accelerometer, goal-setting, and self-monitoring/entry of food intake. End of each day the app reported food intake, PA performance summary, and a balance of ingested and spent calories. This information was used by the app to generate a recommended plan for the following day to improve eating habits and increase PA. Theory: N/A	Focus: Physical activity and Mediterranean diet; Groups: (Intervention) Training and access to mobile phone app and initial standardized counseling in PA and the Mediterranean diet; (control) initial standardized counseling in PA and the Mediterranean diet. Additional to app: Initial counseling session on PA and the Mediterranean diet and print out of support materials. Duration: 3 months; Follow-up post baseline: 3 months	Primary outcome: MVPA and Steps per day; Measuring tool: Actigraph GT3X accelerometer; Secondary outcomes: Adherence to the Mediterranean diet, blood pressure, waist circumference, body mass index, and laboratory parameters.
Skrepnik et al 2017 [29]	Sample size: Total n=211, intervention n=107, control n=104; Population: Adults who have had unilateral knee OA[e] and have been suitable for treatment with Hylan G-F 20. Age: mean 62.6 (SD 9.4); Male (%): 49; Country: United States of America; Attrition rate: 2%	Characteristics: The OA GO app provided motivational messages and requested participants enter mood and pain data once a day. The app displayed daily step count, calories burned, and sleep. Daily and monthly cumulative activity trends were available to view. Theory: N/A	Focus: Physical activity; Groups: (Intervention) Jawbone UP activity tracker and access to OA GO mobile app; (control) Jawbone UP activity tracker. Additional to app: All patients received a single 6 ml injection of Hylan G-F 20 and regular follow-ups as per standard of care. Duration: 90 days; Follow-up post baseline: 90 days	Primary outcome: Steps per day; Measuring tool: Jawbone UP 24 activity tracker; Secondary outcomes: Mean percentage change from baseline in the 6-min walk test, patient and physician satisfaction with treatment, percentage change in Patient Activation Measure (PAM)-13 questionnaire score, percentage change in sleep captured by the wearable activity monitor (light, sound, and duration of sleep), and Visual Analog Mood Scale.

XSL•FO
RenderX

Study	Study population/sample size	Smartphone app features	Intervention description	Outcome measures
Vorrink et al 2016 [27]	Sample size: Total n=183, intervention n=102, control n=81; Population: Physiotherapy patients with COPD, GOLD stage 2 or 3 who had completed a pulmonary rehabilitation program of 3 months. Age: intervention mean 62.0 (SD 9.0); control mean 63.0 (SD 8.0); Male (%): 50; Country: United States of America; Attrition rate: 34%	Characteristics: App displayed physical activity in real-time in quantitative and qualitative form. It displays the total number of steps taken each day relative to the daily goal and offers advice on physical activity progress. Theory: N/A	Focus: Physical activity; Groups: (Intervention) smartphone and app with physical activity goals and automated persuasive messages; (control) usual care. Additional to app: Physiotherapists could monitor patients and adjust their goals or send messages through a website. Duration: 12 months; Follow-up post baseline: 12 months	Primary outcome: Steps per day; Measuring tool: SenseWear Pro or MF-SW mini armband accelerometers. Secondary outcomes: Average METS[f], 6-min walking distance, dyspnea, fatigue, emotional function, mastery, and body mass index

[a]PA: physical activity.

[b]MVPA: moderate-to-vigorous physical activity.

[c]SMART: Specific, Measurable, Achievable, Realistic, and Timely.

[d]N/A: not applicable.

[e]OA: osteoarthritis.

[f]MET: metabolic equivalent of task.

Figure 2. Steps per day mean difference. IV: inverse variance.

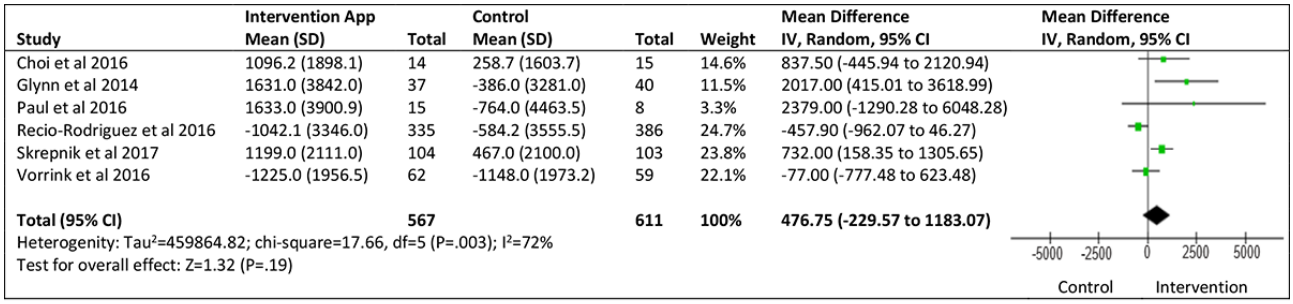
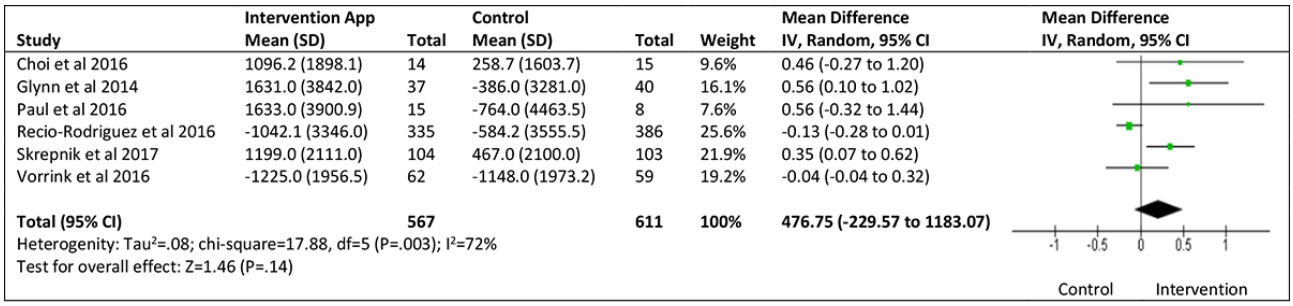


Figure 3. Steps per day standardized mean difference. IV: inverse variance.



Effects of Smartphone App–Based Intervention on Moderate-to-Vigorous Physical Activity Per Day

Time spent in MVPA was reported in 2 studies [25,27] that included a combined total of 732 participants. Both studies reported a nonsignificant trend for daily MVPA minutes to decrease in the intervention group (see Table 3). However, the effect sizes were less than 0.2 (*trivial* [22]) in magnitude as they were in the order of 2 to 3 min per day difference, or an effect size in the order of -0.1 (*trivial* [22]). Due to the small sample (two studies), meta-analysis was not performed.

Sensitivity Analyses

Sensitivity analyses were conducted to determine whether the meta-analysis results were consistent under different conditions. Figures for the results of the sensitivity analyses are available in Multimedia Appendix 4. In the first sensitivity analysis, studies with an intervention length of up to 3 months were included (ie, the studies with an intervention length more than or equal to 3 months were excluded) [27,29-31]. The meta-analysis results suggested physical activity apps significantly increased steps per day by 2074.96 steps per day (95% CI 606.80 to 3543.11, standardized mean difference 0.56, 95% CI 0.16 to 0.97, *P*=.01).

XSL•FO
RenderX

Table 3. Moderate-to-vigorous physical activity effect size for mean difference and standardized mean difference.

Study	Intervention app		Control		Mean difference		Standardized mean difference	
	Mean (SD)	N	Mean (SD)	N	Weight (%)	IV[a], Random, 95% CI	Weight (%)	IV, Random (95% CI)
Fanning et al 2017 [30]	11.9 (25.7)	26	14.1 (24.4)	27	9.0	−2.16 (−15.68 to 11.36)	7.2	−0.08 (−0.06 to 0.45)
Recio-Rodriguez et al 2016 [25]	−7.9 (27.1)	335	−4.3 (29.2)	344	91	−3.16 (−7.85 to 0.63)	92.8	−0.13 (−0.28 to 0.02)

[a]IV: inverse variance.

A second sensitivity analysis examined whether effects for studies that targeted physical activity alone (ie, the one study that intervened on physical activity and diet together was excluded [27]). The meta-analysis results suggested that physical activity apps increased physical activity by 716.86 steps per day (95% CI 38.37 to 1395.36, *P*=.04) or by a standardized mean difference of 0.31 (95% CI 0.07 to 0; *P*=.01).

A final sensitivity analysis examined whether the effects of physical activity apps were consistent for apps targeting a general adult population or populations with specific health conditions (eg, stroke survivors). In both cases, results were consistent with the main meta-analysis; that is, there was a nonsignificant trend for improvement in daily steps in both general adult populations (+649.54 steps per day; 95% CI −822.66 to 2121.74; standardized mean difference 0.24; 95% CI −0.30 to 0.78; *P*=.24) and in populations with specific health conditions (+438.36 steps per day; 95% CI −335.94 to 1212.67; standardized mean difference 0.22; 95% CI −0.09 to 0.53; *P*=.17).

Discussion

Principal Findings

This meta-analysis suggests that app-based physical activity interventions have a nonsignificant, positive influence on objectively measured physical activity. However, sensitivity analyses suggest that effects differ based on study parameters. In particular, there is evidence that smartphone apps have a significant positive effect on physical activity when used over a short-term period (ie, less than 3 months) and where apps target physical activity alone, rather than physical activity in combination with other health behaviors.

To the best of our knowledge, this is the first meta-analysis to establish the effectiveness of app-based physical activity interventions at increasing objectively measured physical activity. Only 1 meta-analysis of app-based physical activity interventions has been published [11], which differs from this study in that it reports on weight loss and subjective physical activity data from both nonrandomized and randomized controlled trials published through to August 2015. Despite these differences, our findings are broadly consistent with both studies finding a nonsignificant increase in physical activity in comparison with control. The other recent systematic review found smartphone apps to have a modest effect on physical activity and noted the limited number of randomized controlled trials that were available at that time to test the efficacy of smartphone apps at increasing physical activity [12]. Our study confirmed that the number of randomized controlled trials evaluating efficacy using objective measurements of physical activity has increased, yet is still limited.

Sensitivity analyses suggested that app-based physical activity interventions were effective when the intervention duration was 3 months or less, compared with longer interventions. This is consistent with findings from 2 other studies [33,34] who similarly reported that physical activity apps appear to be most effective with durations longer than 1 month and 8 weeks, respectively. Taken together, these findings suggest that intervention effects appear to peak within the first couple of months of intervention commencement, and dwindle over time. This raises the possibility that studies with an intervention duration of 3 months or more and those that take their first follow-up assessment at 3 months or later may, in fact, be failing to capture intervention effects, which may have peaked and already started to fade by the time assessments are performed.

That physical activity apps may be most effective in the short term is also consistent with previous studies of engagement with technology-based physical activity interventions, which typically find that engagement declines over time, and this decline corresponds with tapering of intervention effectiveness [18,34-37]. Engagement decline is especially prominent in smartphone-based interventions, as their design precludes human support and supportive accountability [37]. In Recio-Rodriguez et al's [27] trial, just over half of the study population engaged with the app beyond 8 weeks, and the results of the trial favored the control group. Interestingly, the participants in that study who continued to engage with the app for more than 8 weeks actually showed a net increase in MVPA of a mean 44.0 min per week (95% CI 2.1 to 86.0) favoring the intervention group, but this effect was washed out by poor results for the participants who were no longer engaging. Exposure to the intervention is imperative for the intervention to have effect and exposure occurs through participant engagement with the app [36,38]. Thus, these results underscore the notion that ongoing participant engagement with an app is important for intervention effectiveness [36].

Results from the sensitivity analysis suggest that it may be more effective to intervene in physical activity alone rather than in combination with other health behaviors. Note that this interpretation can only be made with caution, given that only 1 study used a multibehavior approach. However, this interpretation is consistent with another recent review that also found apps targeting single health behaviors appear to produce larger improvements than those using a multibehavior approach [34]. In contrast, a recent case study examining the effectiveness

XSL•FO
RenderX

of physical activity apps from the perspective of users suggests apps which combined physical activity and diet components were perceived by users to be more effective than apps with physical activity components alone [33]. User preference for combined physical activity and diet apps is likely owing to the large amount of feedback participants received [33] and the correlation between receiving feedback and motivation to engage with health behaviors [39].

Strengths and Limitations

The stringent inclusion criteria are a key strength of this study, positioning it as the first meta-analysis examining the effectiveness of smartphone app–based interventions on objectively measured physical activity. Maturation of this fast-evolving field allowed us to limit the review to randomized controlled trial methodologies and objective outcomes, heightening the trustworthiness of findings and reducing the likelihood of results being influenced by recall or response bias. In addition, an extensive search of 7 databases was undertaken to reduce the risk of publication bias.

This study is also subject to limitations. Only a relatively small number of studies meeting our strict inclusion criteria were identified and, of these, most had small sample sizes and short intervention lengths. As a result, the confidence intervals for the effect size estimates were quite large, which may have impeded the meta-analysis from determining a significant effect. Despite our strict inclusion criteria, studies were still highly diverse in terms of intervention format, target populations, and study design elements, and heterogeneity scores suggest that the results do not reflect the same pool of data. In particular, some control groups received a minimal intervention [25,26,30,31], which potentially diluted the intervention effect. In addition, although we attempted to focus solely on smartphone apps, some of the included studies included other elements (eg, activity trackers), which in themselves may alter physical activity. This made it impossible to isolate the effects of the mobile phone app component of these interventions. It is acknowledged that although randomized controlled trials are considered the gold-standard experimental design, including only randomized controlled trials in our search criteria excludes studies conducted within ecologically valid designs. As a result, it is possible that our results, based on randomized controlled trials, may differ from those produced by more *real-life* study designs, impeding our ability to make comment on the generalizability of our results to real-world settings. It is also acknowledged that in addition to randomized controlled trials, other study designs with less positivistic assumptions will play an important role in progressing this scientific field [40].

Future Recommendations

This meta-analysis highlights that relatively few high-quality studies have been conducted examining the effectiveness of physical activity smartphone interventions. Future studies should describe their intervention and app features with adequate detail so that results are reproducible, can be learnt from, and advance this field of research.

Future research should be directed toward enhancing understanding of the time course of intervention effects. In particular, increased understanding of the timepoint at which peak effect size is reached, the timepoint at which user engagement decreases, and the factors that underpin these phenomena are required. This may involve future studies with longer follow-up periods and with outcome measurements taken at more regular and frequent timepoints. The relatively short-term nature of positive effects suggests that additional efforts are required to design app features which help sustain user engagement with the app over time, for example, perhaps through modules, unlockable content, and rewards. Sustaining user engagement is particularly important for smartphone-based interventions due to the absence of human support and supportive accountability [37]. Previous research determined the ease of use, function, feedback, tailored information, ability to personalize design, and design-aesthetic as highly ranked engagement strategies [38]. It will be useful for future app designs to incorporate these long-term engagement strategies, as increased exposure to the intervention is suggested to lead to larger, longer lasting effects [35]. Further research is required to confirm or refute our finding that intervening on 1 health behavior could be more effective than interventions targeting multiple health behaviors.

Key recommendations include the following:

1. Research utilizing randomized controlled trial design, in addition to more ecologically valid designs, is required to progress the field.
2. Studies should be designed to improve our understanding of the time course of intervention effects. This could be achieved through more regular assessments throughout the intervention period, rather than the current preponderance for few widely spaced assessments.
3. Strategies to boost ongoing engagement are required to aid sustainable effectiveness.
4. Further research is required to understand whether it is better to target physical activity as a single behavior or in concert with other health behaviors.

Conclusions

This is the first systematic review and meta-analysis of smartphone apps for increasing objectively measured physical activity. Results suggest that such apps lead to a nonsignificant increase in objectively measured physical activity, though effectiveness appears greater in physical activity apps when used in the short term and when the apps target physical activity alone. Overall, the meta-analysis offers modest support for the effectiveness of smartphone physical activity apps.

XSL•FO
RenderX

Acknowledgments

CM is supported by the National Health and Medical Research Council Career Development Fellowship (APP1125913). JR and RC are supported by the NHMRC Project Grant (APP1080186). SE is supported by an Australian Government Research Training Program Scholarship.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy for all databases.

[PDF File (Adobe PDF File), 200KB - jmir_v21i3e12053_app1.pdf]

Multimedia Appendix 2

Blank data extraction form.

[XLSX File (Microsoft Excel File), 12KB - jmir_v21i3e12053_app2.xlsx]

Multimedia Appendix 3

Risk of bias based on Consolidated Standards of Reporting Trials checklist.

[PDF File (Adobe PDF File), 331KB - jmir_v21i3e12053_app3.pdf]

Multimedia Appendix 4

Sensitivity analysis results.

[PDF File (Adobe PDF File), 213KB - jmir_v21i3e12053_app4.pdf]

References

1. Brown WJ, Bauman AE, Bull FC, Burton NW. Australian Government Department of Health. Development of Evidence-based Physical Activity Recommendations for Adults (18-64 years) URL: https://www.health.gov.au/internet/main/publishing.nsf/content/F01F92328EDADA5BCA257BF0001E720D/$File/DEB-PAR-Adults-18-64years.pdf [accessed 2018-06-14] [WebCite Cache ID 71yOyuhi6]
2. World Health Organization. 2018. Physical activity URL: http://www.who.int/mediacentre/factsheets/fs385/en/ [accessed 2019-06-06] [WebCite Cache ID 71qo8nOXd]
3. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
4. World Health Organization. 2010. Global Recommendations on Physical Activity for Health URL: http://apps.who.int/iris/bitstream/handle/10665/44399/9789241599979_eng.pdf;jsessionid=806108ACD914326CF9FB41C060DB22E6?sequence=1 [accessed 2018-05-15] [WebCite Cache ID 71yPFARlS]
5. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. Can Med Assoc J 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
6. GSMA Intelligence. 2018. Global Data URL: https://www.gsmaintelligence.com/ [accessed 2018-08-22] [WebCite Cache ID 71qobDVcK]
7. Molla R. Recode. 2017. Two-thirds of adults worldwide will own a smartphones next year URL: https://www.recode.net/2017/10/16/16482168/two-thirds-of-adults-worldwide-will-own-smartphones-next-year [accessed 2019-06-05] [WebCite Cache ID 71qofnqQe]
8. Sarasohn-Kahn J. California Health Care Foundation. 2010. How Smartphones Are Changing Health Care for Consumers and Providers URL: https://www.chcf.org/wp-content/uploads/2017/12/PDF-HowSmartphonesChangingHealthCare.pdf [accessed 2018-08-27] [WebCite Cache ID 71yPPMpnd]
9. Wu W, Dasgupta S, Ramirez EE, Peterson C, Norman GJ. Classification accuracies of physical activities using smartphone motion sensors. J Med Internet Res 2012 Oct;14(5):e130 [FREE Full text] [doi: 10.2196/jmir.2208] [Medline: 23041431]
10. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]

XSL•FO
RenderX

11. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
12. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Smith S. A review of smartphone applications for promoting physical activity. Jacobs J Community Med 2016;2(1):021 [FREE Full text] [Medline: 27034992]
13. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Res Q Exerc Sport 2000 Jun;71(Suppl 2):1-14. [doi: 10.1080/02701367.2000.11082780] [Medline: 25680007]
14. Müller AM, Maher CA, Vandelanotte C, Hingle M, Middelweerd A, Lopez ML, et al. Physical activity, sedentary behavior, and diet-related eHealth and mHealth research: bibliometric analysis. J Med Internet Res 2018 Apr 18;20(4):e122. [doi: 10.2196/jmir.8954] [Medline: 29669703]
15. 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 healthcare interventions: explanation and elaboration. Br Med J 2009 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
16. Middelweerd A, Mollee JS, van der Wal CN, Brug J, te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
17. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
18. Maher CA, Lewis LK, Ferrar K, Marshall S, de Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
19. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Open Med 2010;4(1):e60-e68 [FREE Full text] [doi: 10.1136/bmj.c332] [Medline: 21686296]
20. Centre for Evidence-Based Medicine. 2016. OCEBM Levels of Evidence URL: https://www.cebm.net/2016/05/ocebm-levels-of-evidence/ [accessed 2018-08-22] [WebCite Cache ID 71qoqam1k]
21. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]
22. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Routledge; 1988.
23. Harries T, Eslambolchilar P, Rettie R, Stride C, Walton S, van Woerden HC. Effectiveness of a smartphone app in increasing physical activity amongst male adults: a randomised controlled trial. BMC Pub Health 2016;16:925. [doi: 10.1186/s12889-016-3593-9]
24. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Effects of three motivationally targeted mobile device applications on initial physical activity and sedentary behavior change in midlife and older adults: a randomized trial. PLoS One 2016;11(6):e0156370 [FREE Full text] [doi: 10.1371/journal.pone.0156370] [Medline: 27352250]
25. Recio-Rodriguez JI, Agudo-Conde C, Martin-Cantera C, González-Viejo MN, Fernandez-Alonso MD, Arietaleanizbeaskoa MS, et al. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the mediterranean diet in primary care: a randomized controlled trial (EVIDENT II study). J Med Internet Res 2016 Dec 19;18(12):e331 [FREE Full text] [doi: 10.2196/jmir.6814] [Medline: 27993759]
26. Paul L, Wyke S, Brewster S, Sattar N, Gill JM, Alexander G, et al. Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: a pilot study. Top Stroke Rehabil 2016 Dec;23(3):170-177. [doi: 10.1080/10749357.2015.1122266] [Medline: 27077973]
27. Vorrink SN, Kort HS, Troosters T, Zanen P, Lammers JW. Efficacy of an mHealth intervention to stimulate physical activity in COPD patients after pulmonary rehabilitation. Eur Respir J 2016 Dec;48(4):1019-1029 [FREE Full text] [doi: 10.1183/13993003.00083-2016] [Medline: 27587557]
28. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101. [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
29. Skrepnik N, Spitzer A, Altman R, Hoekstra J, Stewart J, Toselli R. Assessing the impact of a novel smartphone application compared with standard follow-up on mobility of patients with knee osteoarthritis following treatment with Hylan G-F 20: a randomized controlled trial. JMIR Mhealth Uhealth 2017 May 9;5(5):e64 [FREE Full text] [doi: 10.2196/mhealth.7179] [Medline: 28487266]
30. Fanning J, Roberts S, Hillman CH, Mullen SP, Ritterband L, McAuley E. A smartphone "app"-delivered randomized factorial trial targeting physical activity in adults. J Behav Med 2017 Oct;40(5):712-729. [doi: 10.1007/s10865-017-9838-y] [Medline: 28255750]
31. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014 Jul;64(624):e384-e391. [doi: 10.3399/bjgp14X680461] [Medline: 24982490]
32. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. Br Med J 2011 Feb 10;342:d549. [doi: 10.1136/bmj.d549] [Medline: 21310794]

XSL•FO
RenderX

33. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and physical activity apps: perceived effectiveness by app users. JMIR Mhealth Uhealth 2016;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
34. 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 7;13:127. [doi: 10.1186/s12966-016-0454-y]
35. Gilliland J, Sadler R, Clark A, O'Connor C, Milczarek M, Doherty S. Using a smartphone application to promote healthy dietary behaviours and local food consumption. Biomed Res Int 2015;2015:841368 [FREE Full text] [doi: 10.1155/2015/841368] [Medline: 26380298]
36. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
37. 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 Dec;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
38. Garnett C, Crane D, West R, Brown J, Michie S. Identification of behavior change techniques and engagement strategies to design a smartphone app to reduce alcohol consumption using a formal consensus method. JMIR Mhealth Uhealth 2015;3(2):e73 [FREE Full text] [doi: 10.2196/mhealth.3895] [Medline: 26123578]
39. Bond DS, Thomas JG, Raynor HA, Moon J, Sieling J, Trautvetter J, et al. B-MOBILE--a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. PLoS One 2014;9(6):e100821 [FREE Full text] [doi: 10.1371/journal.pone.0100821] [Medline: 24964010]
40. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. JMIR Mhealth Uhealth 2016 Sep 9;4(3):e107 [FREE Full text] [doi: 10.2196/mhealth.5720] [Medline: 27613084]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
MET: metabolic equivalent of task
MVPA: moderate-to-vigorous physical activity
OA: osteoarthritis
PA: physical activity
SMART: Specific, Measurable, Achievable, Realistic, and Timely

Edited by G Eysenbach; submitted 28.08.18; peer-reviewed by A Mohammadabadi, HL Tong; comments to author 25.10.18; revised version received 10.12.18; accepted 12.12.18; published 19.03.19.

Please cite as:
Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, Crozier A, Maher C
Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis
J Med Internet Res 2019;21(3):e12053
URL: http://www.jmir.org/2019/3/e12053/
doi:10.2196/12053
PMID:30888321

©Amelia Romeo, Sarah Edney, Ronald Plotnikoff, Rachel Curtis, Jillian Ryan, Ilea Sanders, Alyson Crozier, Carol Maher. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Acceptability and Feasibility of Implementing Accelerometry-Based Activity Monitors and a Linked Web Portal in an Exercise Referral Scheme: Feasibility Randomized Controlled Trial

Jemma Hawkins[1], PhD; Joanna M Charles[2], PhD; Michelle Edwards[3], PhD; Britt Hallingberg[1], PhD; Linda McConnon[1], PhD; Rhiannon Tudor Edwards[2], PhD; Russell Jago[4], PhD; Mark Kelson[5], PhD; Kelly Morgan[1], PhD; Simon Murphy[1], PhD; Emily J Oliver[6], PhD; Sharon A Simpson[7], PhD; Graham Moore[1], PhD

[1]Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement, School of Social Sciences, Cardiff University, Cardiff, United Kingdom

[2]Centre for Health Economics and Medicines Evaluation, Bangor University, Bangor, United Kingdom

[3]Division of Population Medicine, Cardiff University, Cardiff, United Kingdom

[4]Centre for Exercise, Nutrition and Health Sciences, University of Bristol, Bristol, United Kingdom

[5]Institute for Data Science and Artificial Intelligence, School of Mathematics, University of Exeter, Exeter, United Kingdom

[6]Department of Sport and Exercise Sciences, Durham University, Durham, United Kingdom

[7]Medical Research Council and Chief Scientist Office Social and Public Health Sciences Unit, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:
Jemma Hawkins, PhD
Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement
School of Social Sciences
Cardiff University
1-3 Museum Place
Cardiff, CF10 3BD
United Kingdom
Phone: 44 02920875184
Email: hawkinsj10@cardiff.ac.uk

Abstract

Background: Exercise referral schemes (ERSs) are recommended for patients with health conditions or risk factors. Evidence points to the initial effectiveness and cost-effectiveness of such schemes for increasing physical activity, but effects often diminish over time. Techniques such as goal setting, self-monitoring, and personalized feedback may support motivation for physical activity and maintenance of effects. Wearable technologies could provide an opportunity to integrate motivational techniques into exercise schemes. However, little is known about acceptability to exercise referral populations or implementation feasibility within exercise referral services.

Objective: To determine the feasibility and acceptability of implementing an activity-monitoring device within the Welsh National ERS to inform a decision on whether and how to proceed to an effectiveness trial.

Methods: We conducted a feasability randomized controlled trial with embedded mixed-methods process evaluation and an exploratory economic analysis. Adults (N=156) were randomized to intervention (plus usual practice; n=88) or usual practice only (n=68). Usual practice was a 16-week structured exercise program. The intervention group additionally received an accelerometry-based activity monitor (MyWellnessKey) and associated Web platform (MyWellnessCloud). The primary outcomes were predefined progression criteria assessing acceptability and feasibility of the intervention and proposed evaluation. Postal questionnaires were completed at baseline (time 0:T0), 16 weeks (T1), and 12 months after T0 (T2). Routine data were accessed at the same time-points. A subsample of intervention participants and scheme staff were interviewed following the initiation of intervention delivery and at T2.

Results: Participants were on average aged 56.6 (SD 16.3) years and mostly female (101/156, 64.7%) and white (150/156, 96.2%). Only 2 of 5 progression criteria were met; recruitment and randomization methods were acceptable to participants, and contamination was low. However, recruitment and retention rates (11.3% and 67.3%, respectively) fell substantially short of target criteria (20% and 80%, respectively), and disproportionally recruited from the least deprived quintile. Only 57.4% of

XSL•FO
RenderX

intervention participants reported receipt of the intervention (below the 80% progression threshold). Less than half reported the intervention to be acceptable at T2. Participant and staff interviews revealed barriers to intervention delivery and engagement related to the device design as well as context-specific technological challenges, all of which made it difficult to integrate the technology into the service. Routinely collected health economic measures had substantial missing data, suggesting that other methods for collecting these should be used in future.

Conclusions: To our knowledge, this is the first study to evaluate short- and long-term feasibility and acceptability of integrating wearable technologies into community-based ERSs. The findings highlight device- and context-specific barriers to doing this in routine practice, with typical exercise referral populations. Key criteria for progression to a full-scale evaluation were not met.

Trial Registration: ISRCTN Registry ISRCTN85785652; http://www.isrctn.com/ISRCTN85785652

(J Med Internet Res 2019;21(3):e12374) doi:10.2196/12374

KEYWORDS

exercise referral; physical activity; feasibility studies; wearable technologies; costs; economic evaluation; fitness trackers; activity trackers; exercise; physical activity; accelerometry

Introduction

Background

Physical inactivity is a major cause of chronic disease [1]. Addressing inactivity at the population level, and among at-risk groups, is a public health priority [2,3]. Interventions for at-risk groups often center around advice and signposting from primary care [4]. Exercise referral schemes (ERSs) are common [5], usually involving health professional referral to community-based structured exercise programs. Although sustained behavior change is consistently associated with internalized, or autonomous, motivation [6-8], patients often enter such schemes motivated by external sources such as general practitioner advice [9-11]. Thus, according to the self-determination theory [12], a key function for ERSs is supporting transition to autonomous motivation through supporting psychological needs for autonomy (volitional and self-endorsed engagement), competence (personal mastery and effectiveness), and relatedness (meaningful interpersonal connections). Although there is evidence of effectiveness of ERSs in the short-term [11,13-15], studies employing multiple follow-ups consistently demonstrate deteriorating effects over time [13,14], perhaps signaling a need for enhanced motivational support to optimize and maintain effects [11].

In Wales, United Kingdom, the National Exercise Referral Scheme (NERS) was established in 2007, which was implemented in 12 local authorities with embedded randomization to test effectiveness before a Wales-wide rollout [16]. After 12-month follow-up, NERS improved physical activity for patients at risk of coronary heart disease. Mediation analyses indicated that increases in autonomous motivation after scheme exit explained almost half of the between-group differences in physical activity 6 months later [17]. Effects on physical activity fell short of significance for the study population as a whole and among patients referred for mental health reasons [18], whereas process evaluation data highlighted a need for postintervention motivational support to maintain changes in the longer term [19,20].

Growing evidence points to potential motivational effects of behavior change techniques (BCTs) such as goal setting, self-monitoring, and personalized feedback on progress toward goals [21-24]. High-quality goal setting and feedback may support autonomous motivation by enhancing patients' sense of competence. The increasing popularity of wearable technologies provides opportunities to enhance goal setting and feedback [25], allowing frequent, automatic feedback on goal progress and tailored updating of goals based on achievement [26]. Incorporation of social components such as remote contact with intervention providers and interaction with other service users may support motivation through promoting relatedness to others. Research on wearable technologies in exercise interventions is growing [27] and suggests that use of wearable technologies may increase physical activity levels [28-32]. Evidence, to date, suggests that the combination of wearable activity monitors (eg, pedometers and accelerometers) and accompanying Web components (eg, websites, social media, and cloud technology) can support exercise motivation [25,32-34].

Some research indicates that existing technologies may lack important BCTs, which are known to play a part in increasing physical activity, such as action planning and problem solving [35]. Thus, it is useful to explore the utility of such technologies as additions to physical activity interventions such as ERSs where they may align with or add to BCTs already in use. Furthermore, little is known about the acceptability of wearable technologies to ERS populations, who, due in part to the typically older age of ERS patients [13], may have less technology experience than, or use technology differently to, younger users [36]. Although several studies have examined the perception and use of activity monitors in older populations, this has not been explored specifically within the context of ERSs [37,38]. The role of ERS staff in supporting setup and use of technological interventions, and the feasibility of randomized trial methods to evaluate the supplementation of ERSs with technological interventions, remains to be established [39]. Hence, before a trial of effectiveness, which may fail to deliver definitive answers at great cost should the intervention or evaluation design prove infeasible, feasibility testing is required to investigate the suitability for technological intervention within an ERS context [40-42].

XSL•FO
RenderX

Table 1. Summary of progression criteria.

Progression criteria[a]	Measures used	Assessment of whether criteria have been met
PC1. Feasibility to recruit a sufficient proportion of new NERS patients to participate in the trial, with appropriate retention to 12-month follow-up (T2)	Percentage of eligible patients recruited; Percentage of participants retained at T2; Regression models used to identify predictors of loss to follow-up	If >20% of new patients recruited=proceed (*green*); if <5%=full-scale trial unlikely to be feasible (*red*). If 5%-20% (*amber*) of the trial steering committee (TSC) will consider the feasibility of proceeding to a full-scale trial bearing in mind the data presented, representativeness of the sample, and possible steps to increase recruitment; If >80% retained at T2=proceed (*green*), if <60%=full-scale trial unlikely to be feasible (*red*). If 60%-80% (*amber*) of the TSC will consider the feasibility of proceeding based on available data and possible steps to increase retention
PC2a. Trial methodology delivered as intended PC2b. Intervention delivered as intended	Summary statistics for intervention fidelity measures overall and by area; Compliance with study invite processes; Compliance with randomization processes	The TSC will consider the data presented and make a judgement about whether the intervention and trial methodology were delivered as intended
PC3. At least 1 of the 2 intervention components is acceptable to participants	Percentages of participants reporting acceptability of intervention components on self-report questions; Issues regarding acceptability of the intervention components explored in qualitative interviews	The TSC will consider the quantitative and qualitative data and make an overall judgement on whether the intervention is acceptable
PC4. Recruitment and randomization processes acceptable to >50% of recruited participants	Percentages of participants reporting acceptability of recruitment and randomization processes on patient questionnaires; Exploration of understanding and acceptability of recruitment and randomization processes in qualitative interviews	>50% of recruited participants report *agree* or *strongly agree* about the acceptability of recruitment and randomization processes; The TSC will apply discretion in judging whether this criterion has been met or could be addressed to improve acceptability in a full-scale trial
PC5. <20% of control group exposed to the intervention components	Percentage of participants in intervention and control groups who report that they were provided with an MWK[b] device or accessed the MWC Web platform	<20% of control participants report they have used an MWK device during the study period; <20% of control participants report that they have accessed MWC during study period

[a]PC: progression criteria.

[b]MWK: MyWellnessKey.

In this study, we have described the results of a feasibility randomized controlled trial (RCT) [43] of the implementation of an activity monitor (MyWellnessKey [MWK], Technogym, Italy; [44]) and linked Web portal (MyWellnessCloud [MWC], Technogym, Italy) within the NERS in Wales.

Objectives

Our primary aim was to assess the feasibility and acceptability of implementing and evaluating the use of MWK activity monitors within the Welsh NERS, to inform decisions on whether to, or how to, proceed to a full trial (see Table 1 for details about the progression criteria). The main objectives were to investigate the following:

1. The feasibility of recruitment and retention.
2. The extent of contamination between arms.
3. The fidelity of intervention and trial methodology.
4. The acceptability of the intervention.
5. The acceptability of randomization.
6. The direction of effect of the intervention on the primary outcome (physical activity) and main hypothesized change mechanism (autonomous motivation).
7. The feasibility of collecting the primary and secondary outcomes, process outcome measures, and economic evaluation methods.

Methods

Design

This study was a feasibility RCT, with process evaluation and exploratory economic analysis. Full details of the methodology, including the intervention and measures, are provided in an open-access peer-reviewed study protocol [43].

Recruitment

Recruitment occurred from January to August 2016 from 8 local authorities in Wales, United Kingdom, purposively selected to provide variation in area characteristics (eg, deprivation, population size, and rurality). Participants were eligible if they (1) were referred to the NERS generic pathway (see Textbox 1) and (2) had the capacity to use the activity monitor (ie, computer access or literacy and an email address). Participants were initially recruited using opportunistic invites from NERS staff [43]. Initial recruitment rates were slower than anticipated. Hence, from week 16 until recruitment closed at week 28, local area co-ordinators forwarded invitation packs containing an information sheet and an expression of interest form (to return to the research team) to all new generic pathway referrals before their initial consultation. On receipt of expression of interest forms, the research team posted recruitment packs to formally recruit interested clients. Participants who returned signed

XSL•FO
RenderX

consent forms and completed baseline questionnaires were randomly assigned 1:1 to receive either the intervention (NERS plus MWK) or the control treatment (usual NERS practice) via a computer-generated random allocation sequence created by the South East Wales Trials Unit. During the third month of recruitment, a chance imbalance in allocation of participants to intervention and control groups was noticed (26 control/46 intervention). The randomization algorithm was investigated and was not found to be erroneous; nonetheless, after consultation with the trial steering committee (TSC), it was agreed to amend the randomization to a 2:1 allocation to balance the groups sufficiently to investigate feasibility parameters. The proposed sample size for the study was 286 participants [43]; however, because of slow initial response rates previously mentioned, it was agreed by the TSC and study funder that the study could proceed with a reduced sample of 156 participants. This smaller sample allowed the estimation of feasibility proportions of adherence and retention to within at least 11.5 percentage points either side using 95% CIs (conservatively assuming proportions of 0.5). Owing to the delays in recruitment and study funding constraints, the 16-month follow-up acceleromtery assessment could only be carried out with participants who reached the 16-month point before 31 August 2017; as such only 63.5% of the total sample (99/156) were eligible to complete this final measure. A subsample of intervention participants were recruited to participate in qualitative interviews following randomization. From the individuals who expressed an interest in taking part in the interviews, participants were purposively recruited to provide variation in local authority area, age, and sex.

Procedure

Questionnaire data were collected at baseline (time 0: T0), at the end of the 16-week NERS program (T1) and 12-months postbaseline (T2) via a postal survey. The data collected routinely within NERS were obtained from each of these time points. Semistructured telephone interviews were conducted with a subsample of intervention participants shortly after intervention receipt (n=19) and again at T2 (n=18) and with a sample of NERS exercise professionals (n=11) at the same time points. Participants received full information about the study procedures and the intervention before providing consent, including which intervention was the *intervention of interest*. This study was given favorable ethical opinion for conduct in the National Health Service on 1 December 2015 by the South East Scotland Research Ethics Committee 02 (REF: 189587) and registered with the International Standard Randomized Control Trial Number Register before recruitment. Figure 1 shows the flow of participants through the trial using a CONSORT flow diagram.

Textbox 1. The National Exercise Referral Scheme (NERS) generic pathway referral criteria.

For referral into the NERS generic pathway, patients must:

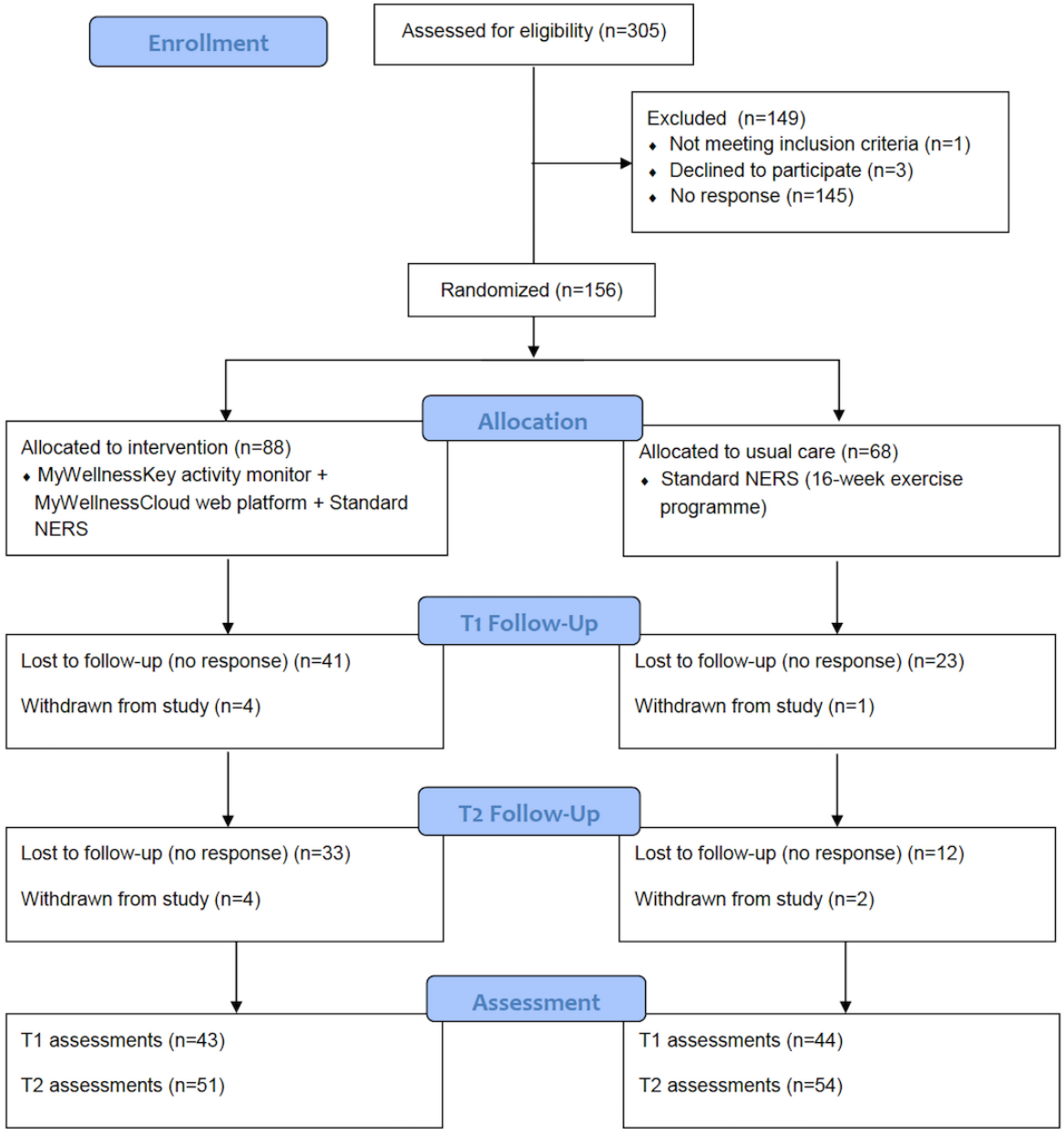
- be aged 16 years or above;
- be sedentary (defined as not moderately active for 3 times per week or deconditioned through age or inactivity);

have at least 1 of the following:

- Raised blood pressure 140/90
- Body mass index >28
- Cholesterol >5.0
- Controlled diabetes or impaired glucose intolerance
- Family history of heart disease or diabetes
- At risk of osteoporosis and/or musculoskeletal pain
- Mild arthritis or poor mobility
- Mild-moderate chronic obstructive pulmonary disorder, asthma, bronchitis, and emphysema
- Mild anxiety, depression, or stress
- Multiple sclerosis

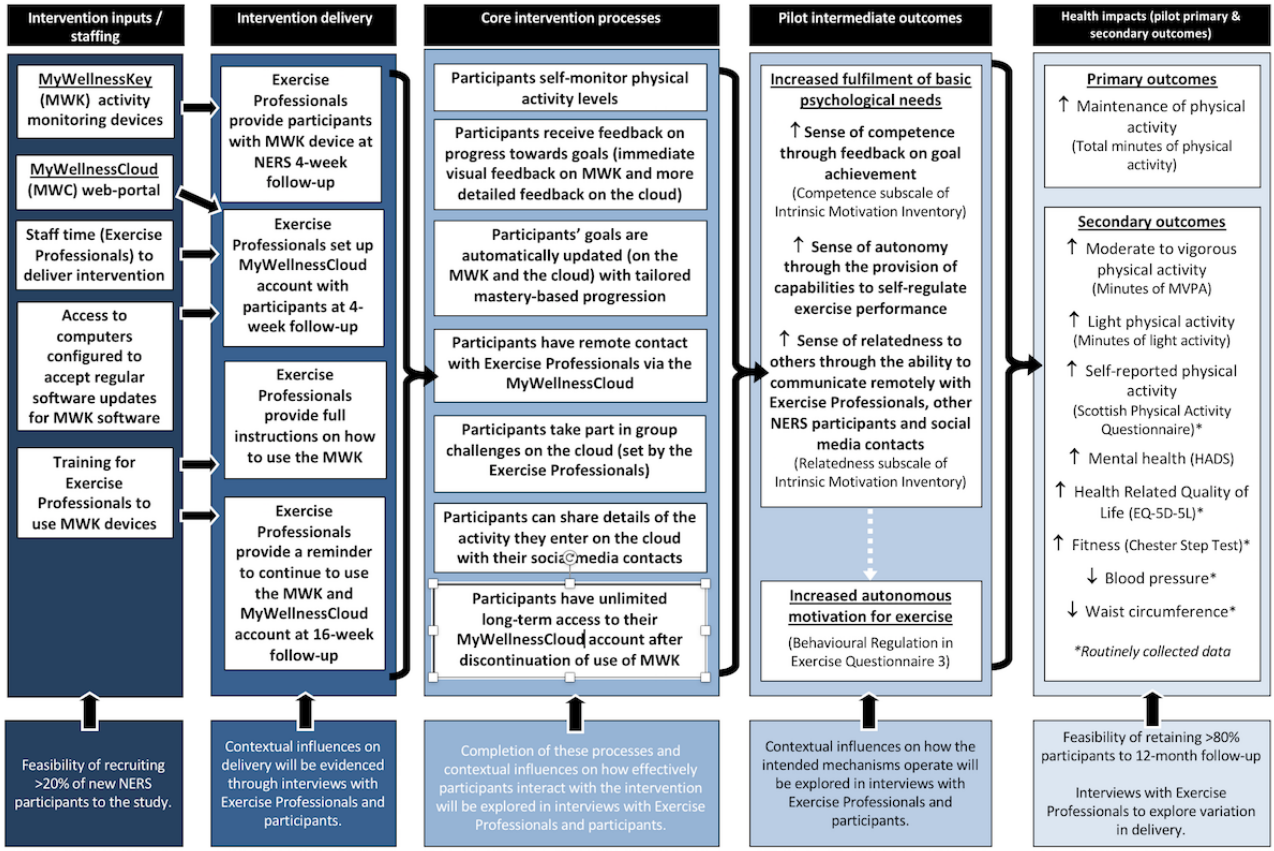
XSL•FO
RenderX

Figure 1. Study flow diagram.



XSL•FO
RenderX

Figure 2. Intervention logic model with study progression criteria.



Intervention

The intervention was an enhanced exercise referral program, which includes usual care (NERS standard practice; [16]) and an accelorometry-based activity monitor (MWK) and associated Web platform (MWC). The MWK is a uniaxial accelerometer, worn at the hip, with a small screen that provides real-time visual feedback. The MWC allows for provision of more detailed feedback and facilitation of support for behavior change following connection of the MWK to a computer via a Universal Serial Bus (USB). For more information about the MWK and MWC, see the study protocol [43], the intervention logic model (Figure 2), and Multimedia Appendix 1. The MWK devices were chosen for use within the NERS in part because of existing use of other Technogym exercise equipment within the centers in which the scheme is delivered. The NERS exercise professionals followed a protocol to provide intervention participants with an MWK and set up their account on the MWC during their 4-week consultation. Participants could use the MWK and MWC up until their 12-month consultation. Control participants received usual NERS care (a 16-week structured exercise program supported with consultations with an exercise professional at the start, 4 weeks, scheme exit (16 weeks) and 12-month follow-up [16]).

Measures

A process evaluation was conducted to examine the acceptability and feasibility of intervention and evaluation methods, including intervention delivery and fidelity, potential contamination, and contextual influences. In total, 5 prespecified progression criteria were agreed among the research team and refined after discussion with the TSC. Various quantitative measures, supported by qualitative interview data, were used to assess whether these criteria (see Table 1) were met. This included a traffic light system for certain criteria (red=stop; amber=discuss with TSC whether there is evidence that sufficient improvements can be made to proceed to full trial without another feasibility assessment; and green=proceed). Questions related to acceptability of the intervention were based on key concepts in technology user acceptance such as ease of use and outcome expectations [45,46]. The data sources used to assess the progression criteria are summarized in Table 1. The feasibility of collecting the primary outcome measure for an effectiveness trial (objectively measured physical activity using accelerometry) was examined in a subsample of intervention and control participants using a separate research grade accelerometer (GT3X ActiGraph). Various secondary outcome measures were collected to inform a future trial including self-reported autonomous motivation (Behavioural Regulations in Exercise Questionnaire 3 [BREQ-3]; [47]), psychological need support, anxiety and depression symptoms, physical activity and routinely collected physiological health measures such as blood pressure, body mass index, and fitness (for more information see the protocol paper by Hawkins et al [43]). Measures for the feasibility of an economic evaluation included an adapted Client Service Receipt Inventory (CSRI) to capture client health and social care service use and health-related quality of life measured by the EQ-5D-5L [48]. Copies of the surveys used to collect the self-report measures can be obtained by contacting the corresponding author.

XSL•FO
RenderX

Data Analysis

Quantitative Analysis

The main quantitative analysis involved descriptive summary statistics for each of the study progression criteria (as outlined in Table 1) as follows:

- PC1: *Recruitment:* Percentage of new NERS generic pathway patients recruited to the trial (excluding the first 8 weeks on the assumption that recruitment rates would stabilize over time); and *Retention:* Percentage of participants retained to 12-month follow-up (returning a completed T2 questionnaire).
- PC2: *Trial methods fidelity:* A summary score of adherence to trial recruitment procedure within audio recordings of initial consultations. Recordings were scored according to whether 3 key pieces of essential information about the study were provided, with a total possible score of 3. *Intervention delivery fidelity:* Percentage of intervention participants reporting receipt of the intervention in the T2 questionnaire.
- PC3: Percentages of participants reporting acceptability and use of the MWK and MWC in the T2 questionnaire.
- PC4: Percentages of intervention and control participants reporting understanding and acceptability of the randomization process in the T1 questionnaire.
- PC5: Percentages of intervention and control participants reporting exposure to the intervention (MWK and MWC) during the study in the T2 questionnaire.

Regression models were used to estimate direction of intervention effects on accelerometer-measured physical activity (16 months) and autonomous motivation (16 weeks and 12 months) as measured with the BREQ-3 [47]. Accelerometer data were processed using standard procedures; periods of ≥60 min of zero counts were recorded as *nonwear time* and removed. Participants were included in the analysis if they provided ≥3 valid days of 500 min of data between 6 am and 11 pm; this value, which is at the lower end of thresholds typically used in the literature [49], was selected to maximize representativeness of the sample within the sedentary population under study. Threshold values for mean minutes of different intensity activity were based on Troiano et al [50]. Sedentary time was estimated based on a cut-point of less than 100 counts per minute, and mean sedentary minutes per day were derived. Linear regression models were fitted for each physical activity outcome controlling for age, gender, baseline self-reported physical activity, and allocation arm. Owing to skewness, mean minutes of moderate to vigorous activity were transformed using a square root transformation. For autonomous motivation, models were fitted for the Relative Autonomy Index (RAI; [47]) score of the BREQ-3 controlling for baseline RAI, age, gender, and number of referral reasons.

Health Economics Analysis

The economic analysis was conducted from a public sector multiagency perspective. Completeness and availability of data using descriptive statistics was used to examine the feasibility of calculating cost-effectiveness alongside a future RCT. Costs of the intervention were calculated by revisiting and revising the costing methodology used in previous economic analysis of the NERS [51]. Quality adjusted life years (QALYs) were calculated from the EQ-5D-5L using the area under the curve method [52]. To address uncertainty in outcomes (QALYs) and costs (service use), bootstrapping (5000 replications) was used to produce 95% CIs around differences in costs and outcomes. Further details can be found in the study protocol [43].

Qualitative Analysis

Qualitative data were transcribed verbatim and organized and coded into a thematic framework using NVivo 11 software (QSR International). The analytic approach incorporated a deductive and inductive approach [53] with data coded using an *a priori* coding scheme of categories aligning with the progression criteria as a means of organizing the data for subsequent interpretation. An element of flexibility was maintained to account for emergence of any new and unexpected themes.

Triangulation

Quantitative and qualitative data were analyzed in isolation with individuals responsible for each analysis blind to the other (eg, statistical analysis conducted by members of the team who were not present for management group meetings where qualitative findings were discussed). On completion of all analyses, the data were then brought together; qualitative data were used to provide further detail and highlight possible explanations for the quantitative findings. Data are organized thematically, drawing on both quantitative and qualitative data sets to provide insights into quantitative feasibility metrics and qualitative insights into barriers and facilitators from multiple perspectives, before an overall picture of progression criteria and decision making on proceeding is presented.

Results

Baseline Characteristics

Participants (N=156) were aged 56.6 (SD 16.3) years and mostly female (101/156, 64.7%) and white (Table 2). There was substantial socioeconomic bias in uptake, with more than half of recruited participants residing in the least deprived quintile of Wales.

XSL•FO
RenderX

Table 2. Baseline (T0) participant characteristics.

Characteristics	Intervention (N=88)	Control (N=68)	Total (N=156)
Age (years), mean (SD)	55.1 (17.6)	58.5 (14.4)	56.6 (16.3)
Female, n (%)	51 (60)	50 (74)	101 (64.7)
White, n (%)	84 (96)	66 (97)	150 (96.2)
Welsh Index of Multiple Deprivation, n[a] (%)			
1–most deprived	2 (2)	0 (0)	2 (1.3)
2	3 (3)	5 (8)	8 (5.2)
3	8 (9)	2 (3)	10 (6.5)
4	22 (25)	24 (36)	46 (29.9)
5–least deprived	52 (60)	36 (54)	88 (57.1)
Income, n[b] (%)			
Less than £5000/year	4 (5)	4 (7)	8 (5.5)
£5000-£9999	7 (8)	11 (18)	18 (12.4)
£10,000-£15,499	22 (27)	15 (24)	37 (25.5)
£15,500-£20,999	18 (22)	10 (16)	28 (19.3)
£21,000-£30,999	12 (15)	10 (16)	22 (15.2)
£31,000-£50,999	16 (19)	7 (11)	23 (15.9)
£51,000 and more	4 (5)	5 (8)	9 (6.2)

[a]A total of 2 participants did not complete this measure, 1 from intervention and 1 from control.

[b]A total of 11 participants did not complete this measure, 5 from intervention and 6 from control.

Recruitment and Retention to the Trial, Contamination, and Acceptability of Randomization (Progression Criteria 1, 2, 4, and 5)

Recruitment

Recruitment fell substantially short of the target of 20% (11.28% [156/1382] of new NERS patients were recruited). After excluding the first 8 weeks, this figure remained similar at 10.99% (111/1010), with 9.1% (31/339) recruitment achieved in the final 8 weeks. Only 6 of 26 (23%) staff provided the audio recordings (N=12) required for assessment of fidelity to the recruitment process. In total, 5 recordings scored 0, with the highest score achieved being 1.75 (out of a total of 3); key information was frequently omitted, which participants might require to make a decision about being contacted by the research team, such as the intervention involving an activity monitoring device or that using it required access to a computer. Qualitative interviews with staff provided explanations for limited adherence to recruitment procedures, including that it was easy to forget to mention the study because it was not part of usual practice, with parts of the procedure often omitted (confirmed by recordings):

> *It's quite difficult, 'cause sometimes even during the consultations, you're kind of talking through it, and 'cause we're on auto pilot, when it comes to asking [about their interest in joining the study], I didn't always remember to do it.* [EP22-T2]

> *If there is someone who is referred and they can hardly move and they're old and they don't have a computer I don't see the point even to talk about it.* [EP71-T1]

Retention and Attrition

A retention rate below the target of 80% was achieved at T2 (105/156; 67.3%). Univariate logistic regression (Table 3) explored potential predictors of follow-up. Retention in the control group of 75% (51/68) was achieved versus 61% (54/88) in the intervention group (odds ratio (OR) 0.53 (95% CI 0.26 to 1.06).

In follow-up phone calls with the 21 participants who did not respond to the T2 questionnaire, 9 reported that disengagement from the NERS was the reason for not completing study questionnaires and 5 cited issues with the MWK as their reason. Staff perceptions of barriers to recruitment and retention also focused on technological problems with the MWK such as lack of internet access or use of another activity monitor and typical disengagement with the NERS:

> *Yeah. there were a couple of older clients who weren't computer literate, and there was one or two who said they didn't have access to any sort of computing.* [EP51-T2]

> *We have three attempts to get back in touch with [non-engaging] clients, like three phone calls and a letter, and if they don't respond, I can't harass them.* [EP82-T2]

XSL•FO
RenderX

Table 3. Predictors of loss to follow-up.

Variable	Odds ratio (95% CI)
Intervention group (N=156)	0.53 (0.26-1.06)
Female (N=156)	1.29(0.65-2.58
Most affluent (N=154)	0.62 (0.31-1.25)
Baseline motivation (N=129)	1.01 (0.99-1.02)
Baseline physical activity (N=134)	0.97 (0.61-1.54)
Multiple referral reasons (N=134)	0.73 (0.32-1.71)

Contamination and Fidelity to Random Allocation

At T2, 10% of control participants (5/51) reported exposure to 1 of the 2 intervention components, whereas 22% of responding intervention participants (12/54) reported that they had not been given an MWK during the study. The proportion of participants who reported using non-MWK activity monitors in the last 12 months was similar in both the control group (12/51, 24%) and intervention group (13/54, 24%). One individual from the control group reported that their decision to use another activity monitor was influenced by participation in the trial. In total, 2 intervention participants reported that they had used another device because of problems they had with the MWK. Staff interviews confirmed the occurrence of contamination, with 1 member reporting giving an MWK to a control participant and 3 reporting advising control participants on how they could access an MWK elsewhere.

Acceptability of Randomization

At T1, 93% (79/85) of participants reported understanding the use of a control group, whereas 84% (72/86) either agreed or strongly agreed that it was acceptable to only give the MWK to half of the participants and 96% (82/86) either agreed or strongly agreed that it was acceptable that the MWK was given to half of the participants at random. Despite high acceptability of randomization in quantitative data, the staff reported that some clients were disappointed by control group allocation. Although the use of other devices was similar across arms, interview data from intervention participants suggested that some might have bought a different activity monitor if they had been allocated to the control group.

Feasibility and Fidelity of Intervention Delivery (Progression Criterion 2)

At T2, 57% of intervention participants (31/54) reported that they had received an MWK during the study, which was below the criterion threshold of 80%. Of those who received the intervention and participated in the T1 questionnaire (n=40), 94% (34/36) stated that their exercise professional provided them with information on how to use the MWK and MWC. Of these, 35% (12/34) reported they received sufficient information on the MWK only, whereas 62% (21/34) received enough information about both the MWK and MWC. Qualitative interviews with staff highlighted a number of issues with Information Technology (IT) and time constraints, which were perceived to have hampered the setup process:

> *I'm aware that some had issues with our MWKs. I know we had issues with setting up the MWKs and with our IT...And also for me, as an instructor, it took a bit of time to set them up.* [EP42-T2]

Most of these issues were linked to either the MWK device or the delivery context (eg, issues with USB devices and IT system security), with fewer being staff-specific (eg, having not attended training or low IT literacy):

> *Because our laptops are encrypted there sometimes can be a bit of an issue with trying to open up the MWC. Also, we couldn't actually download the software to assign MWKs to people [because of firewalls] so the IT department had to over-ride it for us.* [EP61-T1]

In some areas, the staff made attempts to overcome IT issues by using their own laptops or helping participants to set themselves up with the MWK at home:

> *Well I charge the MWK and I give it to them and I give them the instructions to do it at home [...] I ask them beforehand if they're computer literate and would they be happy to do it themselves.* [EP71-T1]

Intervention Acceptability (Progression Criterion 3)

Use of both intervention components was reported by approximately half of the participants, though in both cases this diminished over time. At T2, 57% of intervention participants (31/54) reported using the MWK at some point during the study. However, only 8% (4/49) had used it in the past month. Just under half reported using the MWC at some point during the study, with only 6% (3/47) having used the MWC within the past month. See Multimedia Appendix 2 for a summary of mean scores for acceptability and usage questions. Patient interviews suggested that some engaged with the device initially but stopped owing to a combination of device malfunctions and the novelty factor wearing off:

> *Over a few days, I did quite a lot of exercise and nothing was registered on there. So to be honest with you, I lost a lot of confidence in it. I explained it to my instructor, and he said just carry on with the exercise anyway. So I haven't really used it because nothing was registering.* [BL130-T2]

Other factors influencing engagement included lack of access to a computer and/or internet, poor IT literacy, and technical issues with charging and syncing the device, sometimes

XSL•FO
RenderX

highlighting reliance on a relative or the instructor to support continued use:

> *I needed technological help to explain what had to be done really, and I wasn't altogether the most brilliant at this technology on the system, so I had help from my instructor about that.* [PE088-T2]

The proportion of patients rating the components as easy to use was 49% (21/43) for the MWK and 33% (14/42) for the MWC. Qualitative data highlighted challenges in understanding how the device worked, wearability issues, and not understanding how to use the MWC:

> *I don't think it's the best design to be perfectly honest, it's difficult to attach to your clothing, I think perhaps for a chap it's a little bit easier because they generally wear something with a waistband but women, especially in the summer time often don't, and I think I'm going to struggle in the summer when I'm wearing dresses to find somewhere to put it where it's horizontal.* [AN080-T1]

The proportion of patients reporting that they would use either device in future if they could was 37% (17/46) for the MWK and only 15% (7/46) for the MWC. In the qualitative interviews, participants suggested that they would be more likely to use the intervention in the future if it was easier to understand, technical issues were addressed, and it had better wearability:

> *If it was easier to charge, 'cause the battery kept going, and if it was easier to wear. Being a girl...if I had a dress on for instance, there was nowhere to put it...if I didn't have a pocket or anything like that, then there was nowhere to actually wear it. So if it'd been like on a wristband or something similar, then I probably would have worn it more, I would have just left it on the top with my watch and put it on every day and I'd probably still be using it.* [MO106-T2]

A small majority (26/46, 57%) reported that the device met their expectations in terms of motivating them to be physically active; qualitative data suggest that the reasons it did not meet expectations were linked to the issues reported above:

> *It was beyond what I was hoping for, I've got to be honest. I enjoyed that you could manually enter [on the MWC] if you were doing individual weights and weight machines...or if you were in the garden, and these sorts of things, so I wasn't expecting that.* [PE154-T2]

> *I was hoping it'd be more like a Fitbit, 'cause Fitbits are generally quite easy. But it seemed to be a little bit more complicated than that, I thought, or needed more attention than the Fitbit.* [BR148-T2]

Direction of Effect on Physical Activity and Hypothesized Change Processes

Of the 99 participants (53 control and 46 intervention participants) eligible to provide accelerometer data, 54% (53/99) consented to do so; and 89% of consenting participants provided valid useable data (26/30 control and 21/23 intervention). Of the 6 people who did not provide valid data, 3 did not record sufficient data to meet validity thresholds and 3 did not return the accelerometer. As displayed in Table 4, trends were in the direction of a positive outcome only for sedentary behavior, though with wide CIs for all outcomes. For autonomous motivation, trends were in the direction of a negative outcome at both 16 weeks and 52 weeks (Table 4).

Feasibility of Conducting an Economic Evaluation

Response Rates and Level of Completion

Overall, 156 participants completed baseline economic measures, 85 participants at T1 and 105 participants at T2. Missing data ranged from 0% to 22% (see Multimedia Appendix 3). The EQ-5D (5L) [54], which was obtained from the NERS database, had the largest proportion of missing data of the 2 economic measures. The limited missing data from the CSRI within the study questionnaire show that it is feasible to collect health and social care service use from patients in a future trial. There were limited missing data for measures of productivity losses, ranging from 0% to 18% (see Multimedia Appendix 3).

As shown in Table 5, a total of 25 cases were available for between group comparison of QALYs, and there were 105 cases available for between group comparison of total service use. Service use was lower in the control group, with a significant difference between groups of £386 (including cost of intervention), whereas there was a nonsignificant difference in QALYs between groups of 0.07 QALYs in the opposite direction, equating to 26 days.

Table 4. Direction of intervention effects on physical activity and autonomous motivation.

Variable	Coefficient (95% CI)
Moderate to vigorous physical activity (N=45)	−0.23 (−1.54 to 1.09)
Volume of physical activity (N=45)	−1.20 (−82.42 to 80.0)
Sedentary behavior (N=45)	−18.5 (−81.99 to 44.91)
Autonomous motivation	
16 weeks (N=74)	−3.63 (−14.24 to 6.97)
52 weeks (N=95)	−4.14 (−13.47 to 5.19)

XSL•FO
RenderX

Table 5. Mean quality adjusted life years (QALYs) at 52-week follow-up (T2) by group (mean QALYs at follow-up and 5000 bootstrapped 95% CIs all rounded to 2 decimal places). Mean total service use costs at 52-week follow-up (T2) including the cost of the intervention (mean total service use costs at follow-up and 5000 bootstrapped 95% CIs all rounded to 2 decimal places).

Variable	Intervention group		Control group		Difference between groups (5000 bootstrapped 95% CI)
	n	Mean (SD)	n	Mean (SD)	
QALYs over one year (T2)	11	0.71 (.09)	14	0.78 (.14)	0.07 (0.016 to 0.02)
Total service use costs at T2 including cost of intervention	54	£870 (1332.66)	51	£484 (1230.27)	£386 (35.80 to 452.53)

Table 6. Costs of delivering the National Exercise Referral Scheme (NERS) with MyWellnessKey (MWK) as part of the feasibility trial.

Annual NERS operational costs 2016-2017	Total (£)[a]
National costs paid by the Welsh government	
Consultant	2384
Physical activity specialist (Grade 8a)	10,684
Administrative support	2530
Health improvement coordinator	1392
Meeting costs	300
Exercise professionals (91.5 Whole Time Equivalent [WTE])	2,631,385
Coordination and office costs (eg, printing and stationary) for all 22 local authorities	71,848
Training	64,495
Travel	80,547
Joint national and local costs	
Co-ordinator salary (23 WTE) funding is split between local authorities (£368,438) and the Welsh Government (£478,319)	846,757
Local authority costs	
Staff management	75,000
Promotional material	22,000
Room hire (no charge as covered by session costs)	0
Attending conferences	2200
Total NERS annual operating costs (without MWK)	3,811,522
Participants in NERS[b]	15,626
Cost per participant	244
Additional costs related to MWK	
Cost of MWK activity monitor device (based on 88 units purchased for the trial intervention group)	3960 (£45 per monitor×88)
Cost of MWC annual license fee (professional Web cloud) including Value-Added Tax	3360
Total MWK operating costs	7320
Participants in receipt of MWK as part of the trial	88
Cost per participant for MWK	83
Total cost per participants for NERS with MWK[c]	327

[a]Costs rounded to the nearest pound (£).

[b]Participants in the NERS based on 15,470 individuals who took up the NERS program from September 2016 to August 2017 including the 156 participants taking part in the trial (intervention n=88, control n=68).

[c]Calculation—total annual operational cost per participant and total cost per participant for MWK.

XSL•FO
RenderX

Table 7. Summary of results of progression criteria assessment.

Progression criteria	Results	Criteria met or not
PC1. Feasibility to recruit a sufficient proportion of NERS patients, with appropriate retention rates to T2	11.3% of new NERS patients recruited; 67.3% of study participants retained at T2; No significant predictors of loss to follow-up identified	Not met
PC2a. Trial methodology delivered as intended; PC2b. Intervention delivered as intended	57.4% of intervention participants reported having received the intervention; 35.3% of intervention participants received sufficient information on how to use the MWK; 61.8% received sufficient information for both the MWK and MWC	Not met
PC3. At least 1 of the 2 intervention components is acceptable to participants	49% (MWK) and 33% (MWC) of participants reported the intervention components as easy to use; 37% (MWK) and 15% (MWC) of participants reported that they would use the intervention components in the future; Interview data highlighted challenges in IT and device literacy, technical issues, wearability, and computer access	Not met
PC4. Recruitment and randomization processes acceptable to >50% of recruited participants	92.9% of participants reported understanding the use of a control group; 83.7% of participants agreed that it was acceptable to only give the intervention to half of participants and 95.4% agreed that random allocation was acceptable	Met
PC5. <20% of control group exposed to the intervention components	9.3% of control group participants reported exposure to one of the 2 intervention components	Met

Costs of the Intervention

The costs of NERS are presented for the cost year 2016-2017. Under a delivery framework in which the intervention was absorbed into existing staff roles, the only additional cost of the intervention was the cost of the MWK devices and the annual licence fee for the MWC, which combined with the usual NERS delivery totals £3,818,842 equating to £327 per person based on the 88 intervention participants in this study (Table 6).

As part of a sensitivity analysis, the costs of the NERS were varied using the retail price of the MWK device of £90, rather than the lower price of £45 that they were purchased at. Multimedia Appendix 4 shows the cost of the NERS with MWK when varying the price of the MWK device was £3,822,802; this equates to a cost per person of £352, based on the 88 intervention participants in this study.

Willingness to Pay for Device

In the T2 questionnaire, participants (n=54) responded that they were willing to pay a mean of £29 to use the device during the NERS, reducing to £23 to keep the device afterward. Participants reported willingness to pay as much as £110 (n=2) for the device; however, the minimum amount participants were willing to pay was £0 (n=11).

Progression Criteria for a Full-Scale Trial

Only 2 of the 5 criteria for progressing to a full-scale evaluation were met (see Table 7). Recruitment and retention rates were within the *amber* progression zone, indicating the need for TSC discussion about the feasibility of proceeding, taking into account the data and feedback presented (Table 1). Although the qualitative data provided some insight into barriers to recruitment and retention and possible steps to improve this, it was felt that these issues could not be addressed sufficiently to justify progression without further feasibility work. In addition, the intervention acceptability data did not meet the criterion threshold and there were issues with fidelity of intervention delivery and compliance with randomization processes. In discussion with the TSC, it was agreed that an effectiveness trial would not be feasible given the issues faced with recruitment and retention and the feasibility and acceptability issues related to implementation of the intervention in practice.

Discussion

Principal Findings and Comparisons With Prior Work

This study identified a range of challenges in integrating accelerometer-based wearable technologies into an existing community-based exercise referral program and evaluating this using RCT methodology. There were a number of issues with recruitment and retention of participants and intervention implementation. High attrition, particularly in intervention groups, is common in technological and Web-based health intervention research [55], reflecting challenges in maintaining interest in utilizing such technologies; nonusage attrition often increases steadily over time, with disengagement over time typically observed [55,56]. In this study, less than 10% of intervention participants were still engaging with the device in the final month of the study.

Acceptability of the activity monitors to the ERS population was mixed, with various barriers to use identified. This included wearability and technical problems (eg, difficulty connecting the device to computers and accuracy problems with activity tracking). Comfort and practicality of device wearing has been commonly raised [57,58], particularly among females and within older populations (demographics who make up a large proportion of ERS participants). How a device looks and how secure it is when it is attached are key issues associated with device usage [59], with comfort and wearability closely tied to a device's perceived ease of use and individuals' decision to use [60]. Technical problems are also commonly reported as barriers to engagement [61,62], with individuals reporting issues including device malfunction [63], problems with accuracy of

XSL•FO
RenderX

devices for tracking activities [57], and difficulties accessing feedback on the activity tracked [58]. In this study, some participants cited limited IT literacy as a barrier to engagement, with some noting their age as a factor in their unfamiliarity with such technology, consistent with previous evaluations in older populations. Possibly linked to this, engagement and acceptability ratings for the MWC Web platform were lower than those for the device itself.

Although some studies suggest that similar technologies are both acceptable and feasible to use with adults aged up to 75 years, some have cited difficulties with software installation and the use of associated websites [63]. New and emerging technologies have been perceived as outside the comfort zone of older populations, with an added need to learn the *language* of a device being a barrier to use [59]. Although research on wearable technologies has commonly focused on younger populations, evaluations in populations similar to that of the NERS have identified a need for more extensive support for participants in setup and troubleshooting devices [64], with some implementing high levels of training and troubleshooting within initial months of intervention delivery [38]. Although this study focused on a low-cost delivery model, which integrated this support into roles of existing scheme staff, it may be that such a model is appropriate only where interventions occur in a context of relatively high IT literacy populations. Challenges were perhaps exacerbated by some professionals' lack of buy-in, with professionals describing being distracted by more pressing concerns during time with patients, which limited their ability or willingness to recruit to the study.

Long-term physical activity assessments revealed challenges in use of accelerometers as an outcome measure in this community-based intervention, including low response rates. This study was not intended to assess effectiveness, given its size and limited power. For measures of motivation and physical activity (although not sedentary behavior), directions of effect pointed toward negative impacts, although with wide CIs either side of 0. It is common practice to provide between group comparisons for primary outcomes within feasibility trials to demonstrate that a planned analysis approach is likely to be feasible. However, as feasibility studies are small and underpowered, interventions commonly continue to be refined after a feasibility study and as samples are likely to be unrepresentative of those recruited to a larger trial, such estimates are unlikely to provide meaningful estimates of the likely effect of an intervention. Hence, such data ought to be interpreted with extreme caution. There were substantial missing data from the health economic measures that were routinely collected within the NERS, providing difficulty with conducting an economic evaluation using these methods of data collection. Future work within this population should collect economic data through other self-report methods.

Strengths and Limitations

This study has a number of strengths and limitations. To our knowledge, it is the first to evaluate the use of wearables in an exercise referral population and to explore issues associated with embedding and evaluating technologies within an established community-based intervention. It employed a robust study design including a mixed-methods process evaluation at multiple time points to measure and understand engagement, acceptability, and usability of the intervention alongside piloting measures for an effectiveness study. However, although study sites were purposively sampled to provide a range in levels of area deprivation, the recruited sample was skewed toward a more affluent population, perhaps owing to the more affluent study sites having larger populations or reflecting differences in engagement with activity monitor interventions between socioeconomic groups. The study sample size was originally planned to provide power to detect an effect on the hypothesized mediator, autonomous motivation. However, owing to lower than expected recruitment, sample size targets were revised and hence analyses lack power. Finally, the study evaluated a commercially available device, which is no longer being manufactured. Although there are newer technologies available that overcome some of the issues identified with the MWK (lack of Bluetooth connectivity and issue of wearing at the hip), it is not clear whether these offer the same range of opportunities for behavior change support—particularly from exercise intervention providers.

Nevertheless, these findings offer a number of important insights for future studies. First, although the potential efficiency gains of integrating support and troubleshooting roles into those of existing staff may be appealing, where working with populations with more limited IT skills, additional investment in external support may be required. This is perhaps particularly the case where interventions operate in uncontrolled real-world settings, where professionals serve a large number of clients, as in the NERS, and hence cannot commit much time to supporting engagement with the intervention. Clearly, an introduction of additional technical support components would drive up intervention costs, meaning that effects would perhaps need to be relatively large to justify this investment. Second, recruiting participants to an RCT via routine consultations held by exercise professionals proved challenging; a future full-scale evaluation of similar interventions would require feasible alternative recruitment mechanisms to be established. Finally, as many issues were raised by participants related to the specifics of the MWK devices, the extent to which findings are generalizable to other wearable technology interventions is not always clear. The rapidly evolving nature of wearables and similar technologies presents challenges for efficient and timely evaluation, and RCT methods have been suggested as too slow an approach compared with other more efficient methodologies when evaluating technologies which become out of date during the study period [27].

Conclusions

This study provided an examination of the short- and long-term feasibility and acceptability of integrating wearable technologies into existing community-based ERSs, highlighting some of the possible device- and context-specific barriers. Key criteria for progression to a full-scale evaluation were not met owing to difficulties integrating the technology into routine practice, facilitating uptake by patients, and in methodological challenges relating to the collection of long-term follow-up data. This study demonstrated the importance of investing small amounts of research funding in feasibility assessment before conducting

XSL•FO
RenderX

expensive full-scale effectiveness evaluation, which may fail to be fully executed because of problems with implementing the intervention or evaluation methodology.

Acknowledgments

The authors acknowledge the contributions of research placement students Zoe Hurell and Rachel Morris and research assistant Jordan Van Godwin to data collection during the study. The study was funded by the Welsh Government through Health and Care Research Wales (Health Research Award REF: HRA1019). The work was also undertaken with the support of The Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence. Joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government, and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. SS was supported by an MRC Strategic Award (MC-PC-13027, MC_UU_12017_14, SPHSU14). RJ is partly funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care West (CLAHRC West) at University Hospitals Bristol NHS Foundation Trust.

Authors' Contributions

JH was the principal investigator and responsible for overall management of the study. GM was the principal co-investigator and supported JH with management of the study. JH and GM led the writing of the manuscript. ME and LMcC assisted with implementing the protocol, collecting the data, and analyzing the qualitative data. BH performed statistical analyses and assisted with writing the manuscript. JMC performed the health economic analyses and assisted with writing the manuscript. RJ, MK, SM, EJO, SAS, RTE, and KM were all members of the Trial Management Group, involved in the development of the protocol, provided support for the management of the study, and assisted with writing the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the MyWellnessCloud Web platform.

[PDF File (Adobe PDF File), 356KB - jmir_v21i3e12374_app1.pdf]

Multimedia Appendix 2

Mean scores for usage and acceptability of intervention components.

[DOCX File, 19KB - jmir_v21i3e12374_app2.docx]

Multimedia Appendix 3

Frequencies and percentage of missing data for whole sample at each time point.

[DOCX File, 22KB - jmir_v21i3e12374_app3.docx]

Multimedia Appendix 4

Costs of delivering the National Exercise Referral Scheme (NERS) with the MyWellnessKey (MWK) as part of the feasibility trial varying unit cost of the MWK device.

[DOCX File, 14KB - jmir_v21i3e12374_app4.docx]

References

1. British Heart Foundation. The economic costs of physical inactivity. British Heart Foundation National Centre: Loughborough University; 2013. URL: http://www.ncsem-em.org.uk/wp-content/uploads/2018/12/59541-Economic-Costs-of-Physical-Activity.pdf [accessed 2019-02-25] [WebCite Cache ID 76STTdfrB]
2. Department of Health. London: Her Majesty's Stationery Office; 2011. Start Active, Stay Active: A report on physical activity from the four home countries URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216370/dh_128210.pdf [accessed 2019-02-25] [WebCite Cache ID 76STnB9ts]

XSL•FO
RenderX

3. National Institute for Health and Clinical Excellence. London; 2014. Physical activity: exercise referral schemes URL: https://www.nice.org.uk/guidance/ph54/resources/physical-activity-exercise-referral-schemes-pdf-1996418406085 [accessed 2019-02-25] [WebCite Cache ID 76SVgDfqK]
4. Din NU, Moore GF, Murphy S, Wilkinson C, Williams NH. Health professionals' perspectives on exercise referral and physical activity promotion in primary care: findings from a process evaluation of the National Exercise Referral Scheme in Wales. Health Educ J 2015 Nov;74(6):743-757 [FREE Full text] [doi: 10.1177/0017896914559785] [Medline: 26527835]
5. Sowden SL, Raine R. Running along parallel lines: how political reality impedes the evaluation of public health interventions. A case study of exercise referral schemes in England. J Epidemiol Community Health 2008 Sep;62(9):835-841. [doi: 10.1136/jech.2007.069781] [Medline: 18701737]
6. Teixeira PJ, Silva MN, Mata J, Palmeira AL, Markland D. Motivation, self-determination, and long-term weight control. Int J Behav Nutr Phys Act 2012;9:22 [FREE Full text] [doi: 10.1186/1479-5868-9-22] [Medline: 22385818]
7. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
8. Fortier MS, Wiseman E, Sweet SN, O'Sullivan TL, Blanchard CM, Sigal RJ, et al. A moderated mediation of motivation on physical activity in the context of the Physical Activity Counseling randomized control trial. Psychol Sport Exerc 2011 Mar;12(2):71-78. [doi: 10.1016/j.psychsport.2010.08.001]
9. Rouse PC, Ntoumanis N, Duda JL, Jolly K, Williams GC. In the beginning: role of autonomy support on the motivation, mental health and intentions of participants entering an exercise referral scheme. Psychol Health 2011 Jun;26(6):729-749. [doi: 10.1080/08870446.2010.492454] [Medline: 21827332]
10. Markland D, Tobin VJ. Need support and behavioural regulations for exercise among exercise referral scheme clients: the mediating role of psychological need satisfaction. Psychol Sport Exerc 2010 Mar;11(2):91-99. [doi: 10.1016/j.psychsport.2009.07.001]
11. Rowley N, Mann S, Steele J, Horton E, Jimenez A. The effects of exercise referral schemes in the United Kingdom in those with cardiovascular, mental health, and musculoskeletal disorders: a preliminary systematic review. BMC Public Health 2018 Aug 02;18(1):949 [FREE Full text] [doi: 10.1186/s12889-018-5868-9] [Medline: 30068338]
12. Deci EL, Ryan RM. The "What" and "Why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/S15327965PLI1104_01]
13. Pavey TG, Taylor AH, Fox KR, Hillsdon M, Anokye N, Campbell JL, et al. Effect of exercise referral schemes in primary care on physical activity and improving health outcomes: systematic review and meta-analysis. Br Med J 2011 Nov 04;343:d6462 [FREE Full text] [Medline: 22058134]
14. Williams NH, Hendry M, France B, Lewis R, Wilkinson C. Effectiveness of exercise-referral schemes to promote physical activity in adults: systematic review. Br J Gen Pract 2007 Dec;57(545):979-986 [FREE Full text] [doi: 10.3399/096016407782604866] [Medline: 18252074]
15. Campbell F, Holmes M, Everson-Hock E, Davis S, Buckley WH, Anokye N, et al. A systematic review and economic evaluation of exercise referral schemes in primary care: a short report. Health Technol Assess 2015 Jul;19(60):1-110 [FREE Full text] [doi: 10.3310/hta19600] [Medline: 26222987]
16. Murphy S, Raisanen L, Moore G, Edwards RT, Linck P, Williams N, et al. A pragmatic randomised controlled trial of the Welsh National Exercise Referral Scheme: protocol for trial and integrated economic and process evaluation. BMC Public Health 2010 Jun 18;10:352 [FREE Full text] [doi: 10.1186/1471-2458-10-352] [Medline: 20565846]
17. Littlecott HJ, Moore GF, Moore L, Murphy S. Psychosocial mediators of change in physical activity in the Welsh national exercise referral scheme: secondary analysis of a randomised controlled trial. Int J Behav Nutr Phys Act 2014 Aug 27;11:109 [FREE Full text] [doi: 10.1186/s12966-014-0109-9] [Medline: 25209188]
18. Murphy SM, Edwards RT, Williams N, Raisanen L, Moore G, Linck P, et al. An evaluation of the effectiveness and cost effectiveness of the National Exercise Referral Scheme in Wales, UK: a randomised controlled trial of a public health policy initiative. J Epidemiol Community Health 2012 Aug;66(8):745-753 [FREE Full text] [doi: 10.1136/jech-2011-200689] [Medline: 22577180]
19. Moore GF, Moore L, Murphy S. Integration of motivational interviewing into practice in the national exercise referral scheme in Wales: a mixed methods study. Behav Cogn Psychother 2012 May;40(3):313-330. [doi: 10.1017/S1352465811000531] [Medline: 22008506]
20. Moore GF, Raisanen L, Moore L, Din NU, Murphy S. Mixed-method process evaluation of the Welsh national exercise referral scheme. Health Educ 2013 Oct 14;113(6):476-501. [doi: 10.1108/HE-08-2012-0046]
21. O'Brien N, McDonald S, Araújo-Soares V, Lara J, Errington L, Godfrey A, et al. The features of interventions associated with long-term effectiveness of physical activity interventions in adults aged 55-70 years: a systematic review and meta-analysis. Health Psychol Rev 2015;9(4):417-433. [doi: 10.1080/17437199.2015.1012177] [Medline: 25689096]
22. Shilts MK, Horowitz M, Townsend MS. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. Am J Health Promot 2004;19(2):81-93. [doi: 10.4278/0890-1171-19.2.81] [Medline: 15559708]

XSL•FO
RenderX

23. National Institute for Health and Clinical Excellence. London: National Institute for Health and Clinical Excellence; 2014. Behaviour change: individual approaches URL: https://www.nice.org.uk/guidance/ph49/resources/behaviour-change-individual-approaches-pdf-1996366337989 [accessed 2019-02-25] [WebCite Cache ID 76SVd73WM]
24. Franklin NC, Lavie CJ, Arena RA. Personal health technology: a new era in cardiovascular disease prevention. Postgrad Med 2015 Mar;127(2):150-158 [FREE Full text] [doi: 10.1080/00325481.2015.1015396] [Medline: 25690685]
25. Goode AP, Hall KS, Batch BC, Huffman KM, Hastings SN, Allen KD, et al. The impact of interventions that integrate accelerometers on physical activity and weight loss: a systematic review. Ann Behav Med 2017 Feb;51(1):79-93 [FREE Full text] [doi: 10.1007/s12160-016-9829-1] [Medline: 27565168]
26. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
27. McCallum C, Rooksby J, Gray CM. Evaluating the impact of physical activity apps and wearables: interdisciplinary review. JMIR Mhealth Uhealth 2018 Mar 23;6(3):e58 [FREE Full text] [doi: 10.2196/mhealth.9054] [Medline: 29572200]
28. Hunt K, McCann C, Gray CM, Mutrie N, Wyke S. "You've got to walk before you run": positive evaluations of a walking program as part of a gender-sensitized, weight-management program delivered to men through professional football clubs. Health Psychol 2013 Jan;32(1):57-65. [doi: 10.1037/a0029537] [Medline: 23316853]
29. Annette Dean MC. Using pedometer step-count goals to promote physical activity in cardiac rehabilitation: a feasibility study of a controlled trial. Int J Phys Med Rehabil 2013;04(07):157-162. [doi: 10.4172/2329-9096.1000157]
30. Vaes AW, Cheung A, Atakhorrami M, Groenen MT, Amft O, Franssen FM, et al. Effect of 'activity monitor-based' counseling on physical activity and health-related outcomes in patients with chronic diseases: A systematic review and meta-analysis. Ann Med 2013 Sep;45(5-6):397-412. [doi: 10.3109/07853890.2013.810891] [Medline: 23952917]
31. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
32. Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J. Using an electronic activity monitor system as an intervention modality: a systematic review. BMC Public Health 2015;15:585 [FREE Full text] [doi: 10.1186/s12889-015-1947-3] [Medline: 26104189]
33. Butryn ML, Arigo D, Raggio GA, Colasanti M, Forman EM. Enhancing physical activity promotion in midlife women with technology-based self-monitoring and social connectivity: a pilot study. J Health Psychol 2014 Dec 8;21(8):1548-1555. [doi: 10.1177/1359105314558895] [Medline: 25488937]
34. Muntaner A, Vidal-Conti J, Palou P. Increasing physical activity through mobile device interventions: a systematic review. Health Informatics J 2016 Dec;22(3):451-469. [doi: 10.1177/1460458214567004] [Medline: 25649783]
35. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
36. Olson KE, O'Brien MA, Rogers WA, Charness N. Diffusion of technology: frequency of use for younger and older adults. Ageing Int 2011 Mar;36(1):123-145 [FREE Full text] [doi: 10.1007/s12126-010-9077-9] [Medline: 22685360]
37. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. JMIR Res Protoc 2016 Nov 30;5(4):e237 [FREE Full text] [doi: 10.2196/resprot.6534] [Medline: 27903490]
38. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR Mhealth Uhealth 2016;4(2):e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
39. Hawkins JL, Oliver EJ, Wyatt-Williams J, Scale E, van Woerden H. Assessing the feasibility of using uniaxial accelerometers with an online support platform in the delivery of a community-based exercise referral scheme. J Prim Care Community Health 2014 Oct;5(4):288-290. [doi: 10.1177/2150131914536053] [Medline: 24861974]
40. Hallingberg B, Turley R, Segrott J, Wight D, Craig P, Moore L, et al. Exploratory studies to decide whether and how to proceed with full-scale evaluations of public health interventions: a systematic review of guidance. Pilot Feasibility Stud 2018;4:104 [FREE Full text] [doi: 10.1186/s40814-018-0290-8] [Medline: 29854417]
41. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. Br Med J 2008;337:a1655 [FREE Full text] [Medline: 18824488]
42. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2:64 [FREE Full text] [doi: 10.1186/s40814-016-0105-8] [Medline: 27965879]
43. Hawkins J, Edwards M, Charles J, Jago R, Kelson M, Morgan K, et al. Protocol for a feasibility randomised controlled trial of the use of Physical ACtivity monitors in an Exercise Referral Setting: the PACERS study. Pilot Feasibility Stud 2017;3:51 [FREE Full text] [doi: 10.1186/s40814-017-0194-z] [Medline: 29238607]
44. Sieverdes JC, Wickel EE, Hand GA, Bergamin M, Moran RR, Blair SN. Reliability and validity of the Mywellness Key physical activity monitor. Clin Epidemiol 2013;5:13-20 [FREE Full text] [doi: 10.2147/CLEP.S38370] [Medline: 23378783]

XSL•FO
RenderX

45. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decision Sci 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
46. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
47. Markland D, Tobin V. A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. Journal of Sport and Exercise Psychology 2004 Jun;26(2):191-196. [doi: 10.1123/jsep.26.2.191]
48. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
49. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017 Sep;47(9):1821-1845 [FREE Full text] [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]
50. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
51. Edwards RT, Linck P, Hounsome N, Raisanen L, Williams N, Moore L, et al. Cost-effectiveness of a national exercise referral programme for primary care patients in Wales: results of a randomised controlled trial. BMC Public Health 2013 Oct 29;13(1):1021 [FREE Full text] [doi: 10.1186/1471-2458-13-1021] [Medline: 24164697]
52. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials. Oxford: Oxford University Press; 2014.
53. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Meth 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
54. 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]
55. Watson S, Woodside JV, Ware LJ, Hunter SJ, McGrath A, Cardwell CR, et al. Effect of a web-based behavior change program on weight loss and cardiovascular risk factors in overweight and obese adults at high risk of developing cardiovascular disease: randomized controlled trial. J Med Internet Res 2015;17(7):e177 [FREE Full text] [doi: 10.2196/jmir.3828] [Medline: 26183659]
56. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006 Jun 23;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
57. Shih PC, Han K, Poole ES, Rosson MB, Carroll JM. Use and adoption challenges of wearable activity trackers. In: iConference 2015 Proceedings.: iConference Proceedings; 2015 Presented at: iConference; 2015; Newport Beach, CA.
58. AARP (American Association of Retired Persons). Building a Better Tracker: Older consumers weigh in on activity and sleep monitoring devices URL: https://www.aarp.org/content/dam/aarp/home-and-family/personal-technology/2015-07/innovation-50-project-catalyst-tracker-study-AARP.pdf [accessed 2019-02-25] [WebCite Cache ID 76SXLxegW]
59. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
60. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
61. Blake H, Batt ME. Employee perceptions of a pedometer walking intervention in a hospital workplace. Int J Health Promot Educ 2015 Mar 02;53(5):257-270. [doi: 10.1080/14635240.2015.1016621]
62. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. Process evaluation of physical activity counselling with and without the use of mobile technology: A mixed methods study. Int J Nurs Stud 2016 Jan;53:3-16. [doi: 10.1016/j.ijnurstu.2015.10.008] [Medline: 26518108]
63. Meyer J, Hein A. Live long and prosper: potentials of low-cost consumer devices for the prevention of cardiovascular diseases. Med 2 0 2013;2(2):e7 [FREE Full text] [doi: 10.2196/med20.2667] [Medline: 25075242]
64. Ehn M, Eriksson LC, Åkerberg N, Johansson A. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):e34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]

Abbreviations

BCTs: behavior change techniques
BREQ: Behavioural Regulations in Exercise Questionnaire
CSRI: Client Service Receipt Inventory
ERS: exercise referral scheme

XSL•FO
RenderX

MWC: MyWellnessCloud
MWK: MyWellnessKey
NERS: National Exercise Referral Scheme
PC: progression criteria
QALYs: quality adjusted life years
RAI: Relative Autonomy Index
RCT: randomized controlled trial
TSC: trial steering committee

Edited by G Eysenbach; submitted 05.10.18; peer-reviewed by C McCallum, B Kim, J Lee; comments to author 12.12.18; revised version received 20.12.18; accepted 26.12.18; published 29.03.19.

Please cite as:
Hawkins J, Charles JM, Edwards M, Hallingberg B, McConnon L, Edwards RT, Jago R, Kelson M, Morgan K, Murphy S, Oliver EJ, Simpson SA, Moore G
Acceptability and Feasibility of Implementing Accelorometry-Based Activity Monitors and a Linked Web Portal in an Exercise Referral Scheme: Feasibility Randomized Controlled Trial
J Med Internet Res 2019;21(3):e12374
URL: http://www.jmir.org/2019/3/e12374/
doi:10.2196/12374
PMID:30924791

©Jemma Hawkins, Joanna M Charles, Michelle Edwards, Britt Hallingberg, Linda McConnon, Rhiannon Tudor Edwards, Russell Jago, Mark Kelson, Kelly Morgan, Simon Murphy, Emily J. Oliver, Sharon A. Simpson, Graham Moore. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Detecting Hypoglycemia Incidents Reported in Patients' Secure Messages: Using Cost-Sensitive Learning and Oversampling to Reduce Data Imbalance

Jinying Chen[1,2], PhD; John Lalor[2,3], MS; Weisong Liu[2,4], PhD; Emily Druhl[2], MPH; Edgard Granillo[1,2], MD; Varsha G Vimalananda[2,5], MPH, MD; Hong Yu[2,3,4,6], PhD, FACMI

[1]Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States
[2]Bedford Veterans Affairs Medical Center, Center for Healthcare Organization and Implementation Research, Bedford, MA, United States
[3]College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA, United States
[4]Department of Computer Science, University of Massachusetts Lowell, Lowell, MA, United States
[5]School of Medicine, Boston University, Boston, MA, United States
[6]Department of Medicine, University of Massachusetts Medical School, Worcester, MA, United States

Corresponding Author:
Jinying Chen, PhD
Department of Population and Quantitative Health Sciences
University of Massachusetts Medical School
Albert Sherman Center, 9th Floor
368 Plantation Street
Worcester, MA, 01605
United States
Phone: 1 508 856 6063
Fax: 1 508 856 8993
Email: jinying.chen@umassmed.edu

Abstract

Background: Improper dosing of medications such as insulin can cause hypoglycemic episodes, which may lead to severe morbidity or even death. Although secure messaging was designed for exchanging nonurgent messages, patients sometimes report hypoglycemia events through secure messaging. Detecting these patient-reported adverse events may help alert clinical teams and enable early corrective actions to improve patient safety.

Objective: We aimed to develop a natural language processing system, called HypoDetect (Hypoglycemia Detector), to automatically identify hypoglycemia incidents reported in patients' secure messages.

Methods: An expert in public health annotated 3000 secure message threads between patients with diabetes and US Department of Veterans Affairs clinical teams as containing patient-reported hypoglycemia incidents or not. A physician independently annotated 100 threads randomly selected from this dataset to determine interannotator agreement. We used this dataset to develop and evaluate HypoDetect. HypoDetect incorporates 3 machine learning algorithms widely used for text classification: linear support vector machines, random forest, and logistic regression. We explored different learning features, including new knowledge-driven features. Because only 114 (3.80%) messages were annotated as positive, we investigated cost-sensitive learning and oversampling methods to mitigate the challenge of imbalanced data.

Results: The interannotator agreement was Cohen kappa=.976. Using cross-validation, logistic regression with cost-sensitive learning achieved the best performance (area under the receiver operating characteristic curve=0.954, sensitivity=0.693, specificity 0.974, F1 score=0.590). Cost-sensitive learning and the ensembled synthetic minority oversampling technique improved the sensitivity of the baseline systems substantially (by 0.123 to 0.728 absolute gains). Our results show that a variety of features contributed to the best performance of HypoDetect.

Conclusions: Despite the challenge of data imbalance, HypoDetect achieved promising results for the task of detecting hypoglycemia incidents from secure messages. The system has a great potential to facilitate early detection and treatment of hypoglycemia.

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e11990) doi:10.2196/11990

KEYWORDS

secure messaging; natural language processing; hypoglycemia; supervised machine learning; imbalanced data; adverse event detection; drug-related side effects and adverse reactions

Introduction

Significance and Background

Diabetes mellitus is a highly prevalent disease estimated to affect 425 million people worldwide with a cost of US $673 billion in 2015 [1]. Glycemic control is important for preventing long-term complications of diabetes. However, many factors, including improper dosing of antidiabetic medications such as insulin and sulfonylureas, can lead to hypoglycemic episodes, increasing the risk of severe morbidity or even death [2-5]. It is important to report these hypoglycemia incidents to clinical teams quickly, so that early corrective actions can be taken to improve patient safety.

Secure messaging is a popular functionality of patient portals [6-9] that has been increasingly used in recent years [10,11]. This technology allows for secure communication between patients and health care providers between episodic in-person clinic visits. The use of secure messaging has been associated with improved health care quality and outcomes [12-18]. Although secure messaging was designed for exchanging nonurgent messages (eg, clinical teams in the US Veterans Affairs' [VA] health care system are allowed up to 3 days to respond to patients' messages), patients sometimes use it to report urgent issues that need immediate attention, including adverse events [19-21]. It is important for health care providers to attend to these urgent issues early.

However, the volume of patient-provider secure messages can be huge. For example, the number of secure messages exchanged in the first quarter of 2018 at VA's My Health*e*Vet [8,22] reached about 3.7 million [23]. It is beneficial and necessary to use automated methods to facilitate inspection of these data.

Objective

This study aimed to develop HypoDetect (Hypoglycemia Detector)—to our knowledge, the first natural language processing (NLP) system to automatically identify hypoglycemia incidents from patients' secure messages to facilitate timely responses. This task is challenging because, like other adverse events, hypoglycemia incidents are rare. In addition, patients report hypoglycemia in diversified, informal ways (detailed in the Results and Discussion sections). We addressed this challenge by using supervised learning methods with strategies to handle data imbalance and empirically evaluated our approach.

Related Work

Natural Language Processing for Secure Message Classification

Previously, secure messages were typically analyzed by human experts [20,24], which is difficult to scale up.

Recent studies applied NLP methods to analyze secure messages to identify patients' information needs [25-27]. Cronin and colleagues applied machine learning methods to classify patients' information needs into 5 categories: clinical information, medical, logical, social, and other [25,26]. They found that random forest and logistic regression models and term frequency features were most effective for this task. Sulieman and colleagues extended that work by investigating new semantic and contextual features and deep learning models such as convolutional neural networks [27]. They found that convolutional neural networks with paragraph embeddings outperformed other models.

Learning From Imbalanced Data

Imbalanced data refers to datasets in which some classes have much fewer instances than others. Without treating data imbalance, automated systems often have poor recall (sensitivity) for the minority class [28], which will be a severe problem when the minority class is the target to predict.

Previous work in learning imbalanced clinical data focused on cancer screening and diabetes diagnosis [29-32]. For example, Zahirnia and colleagues adopted a hybrid cost-sensitive learning approach to predict diabetes [29]. Ramezankhani and colleagues showed that the synthetic minority oversampling technique (SMOTE) improved classifiers' sensitivity, but not precision and F1 score, when predicting diabetes [31].

Blagus and Lusa empirically studied two methods for combining cross-validation and sampling techniques [32]. The first one divides the dataset into multiple folds and then samples the training set in each fold independently. The second method first samples the whole dataset and then divides the sampled dataset into multiple folds. Their results indicated that the first method is technically correct, especially for oversampling techniques including SMOTE. This is because oversampling on the whole dataset is likely to add similar or identical instances into both the training and test sets, causing an overestimation of classification performance.

Methods

Secure Messages

We collected 3000 secure message threads between patients with diabetes and VA clinical teams for this study. A secure message thread (abbreviated as *a thread* for convenience) refers to a single, entire thread of messages exchanged between a patient and his or her VA clinical providers. A message refers to a single instance of a communication in a thread. A thread includes 1 or more messages.

We conducted our data sampling process in 2 stages. In the first stage, we obtained a list of patient identification numbers from the VA data service for patients who had a diabetes outpatient visit (*International Classification of Diseases, Ninth Revision*

XSL•FO
RenderX

[ICD-9] codes: 249.x and 250.x; *ICD Tenth Revision* [ICD-10] codes: E08-E13) between 2009 and 2017 and used these identification numbers to retrieve 2.3 million secure message threads from the VA's Corporate Data Warehouse (Office of Information Technology, Department of Veterans Affairs, Washington, DC, USA).

In the second stage, we sampled our evaluation set from the 2.3 million threads obtained in stage 1. We first randomly sampled 1000 threads. An expert in public health annotated those threads and found that only 1 contained a hypoglycemia incident. We therefore used an enrichment approach to sample 2 additional sets of threads. Specifically, we used 2 rule-based methods to improve the recall of positive examples. Both methods constrained their sampling to outpatient visits that had diagnosis codes related to hypoglycemia (ICD-9 codes: 251.0, 251.1, 251.2; ICD-10 codes: E16.1, E16.2) and to the secure message threads sent within 30 days before or after those outpatient visits. The first method randomly sampled 1000 threads that contained at least one of the following keywords: blur, confused, dizzy, headache, hungry, pale, shake, sleepy, sweat, weak, dose, drop, and down. We selected these keywords based on information from the "Hypoglycemia" webpage posted on the US National Institute of Diabetes and Digestive and Kidney Diseases' website [33]. The second method randomly sampled 1000 threads that contained the keyword sugar and at least one of the keywords used in the first method. We searched keywords by using a fuzzy match algorithm written in Transact-SQL that could retrieve inflectional variants of a keyword. We ensured that the 3 sets of threads had no duplicates and combined them into a single set (3000 threads in total) for this study.

Annotation

An expert in public health who has worked in the civilian and military health care fields for 18 years annotated each thread as containing a hypoglycemia incident (positive) or not (negative).

We created a simple annotation guideline based on the American Diabetes Association's standard [34] and Miller et al [2] (see Multimedia Appendix 1). We deemed a message to be positive if it (1) mentioned a blood glucose level <70 mg/dL (<3.9 mmol/L) [34], or (2) described typical hypoglycemia symptoms [2] that could not be contextually attributed to other possible causes (eg, high blood sugar and low blood pressure). We annotated a message thread as positive if it contained at least one positive patient message.

Multimedia Appendix 1 shows excerpts from several deidentified positive and negative messages in our dataset. It is worth noting that we judged as positive any messages that did not report blood glucose levels or reported borderline values but included typical symptoms in a context where hypoglycemia was likely to have occurred (eg, skipping a meal, taking diabetes medication, and then feeling lightheaded and sweaty). Examples 2 and 3 in the first table in Multimedia Appendix 1 fall into this category. We judged as negative the messages that were too vague (ie, lacking clear context) to determine whether an incident of hypoglycemia had occurred. Examples 2 and 4 in the second table in Multimedia Appendix 1 fall into this category.

We further asked a physician who specializes in family medicine to annotate 100 threads from these data independently. The 100 threads combined 2 sets of data. The first set contained 50 threads randomly selected from the 3000 threads, with 5 positive and 45 negative threads. The second set contained 25 randomly selected positive threads and 25 randomly selected negative threads.

Training and Evaluation Data

We trained and evaluated HypoDetect by 10-fold cross-validation (detailed in Experimental Settings). We found that a secure message thread frequently contained 2 or more secure messages from patients and providers, with the first message being from the patient to raise questions or report problems. Because our goal was to develop a system to facilitate timely response to patient-reported hypoglycemia, we expected the system to make a prediction right after seeing the first patient message. In addition, we found that patients almost always reported hypoglycemia incidents in the first message of a thread. Therefore, we used the first message from each thread for our experiments.

This treatment also helped us to regularize the data. For example, it reduced the length variation of training and test examples, a factor that may affect the effectiveness of frequency-based features such as term frequency-inverse document frequency (TF-IDF). In addition, it helped the system focus on text regions where patients reported problems (including hypoglycemia events) and reduced distractive signals elsewhere.

The HypoDetect System

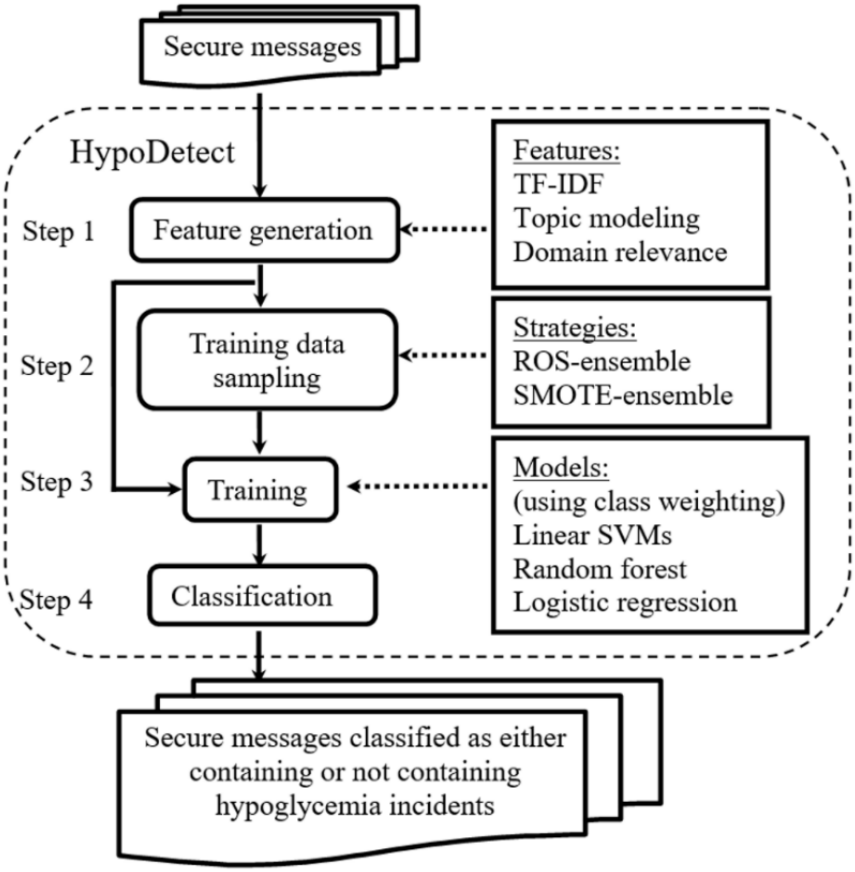
System Overview

HypoDetect processed the data in 4 steps: feature generation, training data sampling, training, and classification (Figure 1). We investigated 2 oversampling strategies in mitigating data imbalance (Figure 1, step 2) and 3 machine learning methods for text classification (Figure 1, step 3, trained using class weighting).

Cost-sensitive learning and data sampling are 2 strategies that have been widely used to address the problem of data imbalance [28,35,36], including problems in the clinical domain [29-32]. Cost-sensitive learning addresses data imbalance at the algorithm level by associating high costs with misclassifying minority examples (also called class weighting) when training machine learning models. Sampling methods modify the training data to make them balanced and thus suitable for standard learning algorithms. Details are as follows.

XSL•FO
RenderX

Figure 1. Workflow of HypoDetect. ROS: random oversampling; SMOTE: synthetic minority oversampling technique; SVM: support vector machine; TF-IDF: term frequency-inverse document frequency.



Data Sampling to Reduce Data Imbalance

Random Oversampling

This method randomly samples minority examples (positive examples in our case) with replacement to increase the number of positive examples. Previous work [31] found that sampling the training set to be completely balanced (ie, having equal numbers of positive and negative instances) was an effective strategy. We therefore adopted this strategy for random oversampling and SMOTE (described below) in our study.

Synthetic Minority Oversampling Technique

Instead of randomly oversampling minority (positive) examples, SMOTE [37] creates "synthetic" positive examples. Specifically, for each positive example x_i, SMOTE generates a new example by using this example and its k positive-class nearest neighbors in the feature space, as defined in equation (1) (Figure 2), where x'_i is the new example synthesized from the positive example x_i and the example x_j that is randomly selected from x_i's k nearest neighbors, and λ is a random value ranged in [0,1]. We set k=5 by following previous work [37].

By its definition in equation (1), SMOTE usually will not remove a word feature (ie, set the feature value to 0) from a synthesized message if the word occurred in the positive message used to generate the synthesized message. Another property of SMOTE is that it can enrich the representation of a message by using additional words that occurred in messages similar to this message. This treatment may alleviate the data sparsity problem that often occurs when using word features.

SMOTE is widely used for learning from imbalanced data due to the simplicity of its sampling procedure and its robustness when applied to different types of problems [38], including clinical classification problems [31,32]. By comparing SMOTE with its 3 variations (borderline SMOTE, support vector machine [SVM] SMOTE, and adaptive synthetic sampling approach) in our preliminary experiments using the first fold of our data, we found that SMOTE worked consistently better with the 3 machine learning algorithms used by HypoDetect. We therefore chose SMOTE for this study.

Ensembled Oversampling Methods

We extended each oversampling method to an ensemble version to improve model robustness. Specifically, during the training phase, we ran an oversampling method on the training set 10 times to train 10 models. We then classified the test examples by voting from the 10 models.

XSL•FO
RenderX

Figure 2. Equations for (1) synthetic minority oversampling technique, (2) inverse document frequency (IDF), (3) term frequency-inverse document frequency (TF-IDF), and (4) F1 measure.

$$x'_i = x_i + \lambda(x_j - x_i) \quad (1)$$

$$IDF(t) = log\frac{N+1}{DF(t)+1} + 1 \quad (2)$$

$$TF\text{-}IDF\ (t) = TF(t) \times IDF(t) \quad (3)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

Machine Learning Models and Class Weighting

While deep learning has shown success in text classification [39], it mostly worked well when the training dataset is large. Due to the knowledge bottleneck challenge, the clinical training dataset is frequently much smaller. Recent work showed that feature-based supervised machine learning approaches outperformed deep learning approaches in certain clinical classification tasks [40]. Therefore, we experimented with 3 feature-based machine learning algorithms (details in Multimedia Appendix 2) that have been widely used and are the state-of-the-art for text classification: linear support vector machines [41], random forest [42], and logistic regression.

We used balanced class weighting when training the cost-sensitive machine learning models. Specifically, we weighted each class by the reciprocal of the number of training instances belonging to that class.

Features for Machine Learning

Term Frequency-Inverse Document Frequency

TF is the number of occurrences of a word in each individual secure message. IDF and TF-IDF are calculated by equations (2) and (3) (Figure 2), where *t* is a word, *DF(t)* is the number of secure messages in a data collection that contained *t* (also called document frequency of *t*), and *N* is the total number of secure messages in the data collection. We computed IDF on the 3000 secure messages used in this study. We removed words that occurred in the stop-word list from scikit-learn [43] or occurred less than 3 times in our secure message corpus. In total, we generated 5910 normalized, real-valued TF-IDF features.

Topic Features

Topic features, as represented by $P(topic_i|\ d)$ (i=1, 2,… K), are real-valued features in [0,1] to indicate the probability of the i th *topic* given a document *d* (ie, a secure message in our case). K is the number of topics used in topic modeling.

We first used the latent Dirichlet allocation algorithm [44] implemented by the Machine Learning for Language Toolkit (MALLET) [45] to train a topic model on 10,000 secure messages that were randomly selected from the same data pool we used to select the evaluation data. We then applied the topic model to the 3000 secure messages to obtain the topic features, that is, the topic distribution over each secure message. We set the topic number K to 100 after manually assessing the informativeness and granularity of topics generated by using different Ks (K=20, 30, 50, 100, 200). We set other hyperparameters of topic modeling to default values. For example, we set the concentration parameters Alpha (for distribution of topics per document) and Beta (for distribution of words per topic) to 50/K and 0.01. The first table in Multimedia Appendix 3 shows examples of topics generated by topic modeling.

Domain-Relevance Features

We defined domain-relevance features by word embeddings and predefined domain-specific keywords.

Word embeddings are distributed vector representations of words. Word embeddings have emerged as a powerful technique for word representation and proved beneficial in a variety of biomedical and clinical NLP tasks. We used word2vec software to create the word embeddings [46,47] and trained word2vec using 4.86 million electronic health record notes, including progress reports, discharge summaries, history and physical reports, and consult notes, from UMass Memorial Medical Center, Worcester, MA, USA. We used 200-dimension vectors by following Jagannatha et al [48] and Pyysalo et al [49]. We used the continuous bags of words model with a window set of 8, negative sampling (25 negative samples for each positive sample), and a subsampling threshold of 0.0001 for training.

To generate the domain-relevance features, we manually created 8 topics by keywords describing hypoglycemia symptoms or incidents (see the second table in Multimedia Appendix 3). We then used these topics to create 28 features in the following steps.

XSL•FO
RenderX

We first included 8 binary-valued features indicating whether the message contained a word belonging to a topic. Then, we included 4 binary features indicating whether the message contained a domain-specific topic word, a number, a number lower than 70, and keywords such as hypoglycemia, low sugar, and low blood sugar.

To generate real-valued features, we represented a topic by the average word embeddings of its topic words [48,50]. For each secure message and each topic, we computed the cosine similarities between this topic and the words in this message and chose the maximum similarity score as the feature value for this topic. This way, we obtained 8 real-valued features associated with 8 topics. We then normalized the 8 features to obtain another 8 real-valued features.

Baseline Systems

To examine the effectiveness of using cost-sensitive learning and oversampling to mitigate the challenge of data imbalance, we compared the HypoDetect systems that use these strategies with 3 types of baselines: (1) a rule-based classifier, (2) the systems that did not treat data imbalance, and (3) the systems that used ensembled undersampling.

The rule-based method classified a message as positive if it satisfied either of the following criteria: (1) it contained the keyword low blood pressure, low sugar, hypoglycemia, or hypoglycemic; or (2) it contained the keyword sugar or glucose, and contained at least two typical symptoms related to hypoglycemia (details in Multimedia Appendix 4).

Ensembled undersampling extends undersampling that randomly selects a subset of examples from the majority class to balance the training data. It has been shown to be effective or even better than oversampling for some classification tasks [51]. However, this method may not work well when the number of positive instances is small and the feature space is large.

Experimental Settings

We used 2-layer 10-fold cross-validation to develop and evaluate supervised learning systems. Specifically, we divided the 3000 messages into 10 folds using stratified random sampling. Each fold had 300 messages, with 10 to 12 positive messages. For each fold, we used the remaining 9 folds to train the system and evaluated the system on this fold. By repeating this procedure for each fold, we obtained the evaluation results on the full dataset. When training the system, we used 10-fold cross-validation on the training data to find the optimal hyperparameters. This approach allows for all the data to be used as the evaluation set while ensuring that the training, validation, and evaluation data are separated.

When applying the data sampling techniques, we only sampled the training data by following Blagus and Lusa's work [32]. After oversampling, the training set for each fold contains 5196 (2598 positive and 2598 negative) or 5194 (2597 positive and 2597 negative) examples. After undersampling, the training set for each fold contains 204, 206, or 208 (with equal numbers of positive and negative examples) examples.

We used the open source software scikit-learn [43], version 0.19.1, to build the systems and develop the ensembled sampling techniques.

We report our evaluation results at the corpus level (ie, first merging the system outputs from the 10 folds and then calculating the evaluation metrics) in the Results section and the fold-level results in Multimedia Appendix 5. In addition, in the third table in Multimedia Appendix 5, we provide statistics (mean, standard deviation, minimum, and maximum) of the performance scores of individual classifiers used by the ensembled oversampling models.

Evaluation Metrics

Sensitivity (Recall), Specificity, Precision, and F1 Score

Sensitivity, or recall, is the number of true positives (ie, secure messages that contained hypoglycemia events and were correctly predicted by the model) divided by the total number of positive instances (ie, total number of secure messages that contained hypoglycemia events).

Specificity is the number of true negatives divided by the total number of negative instances.

Precision is the number of true positives divided by the total number of instances that were predicted to be positive by the model.

The F1 score is the weighted average of precision and recall, as defined by equation (4) (Figure 2). The F1 score takes both false positives (measured by precision) and false negatives (measured by recall) into account. This measure is often used to assess a classifier's performance on handling uneven class distribution, that is, imbalanced data [28,31,51].

Area Under the Receiver Operating Characteristic Curve

This computes the area under the receiver operating characteristic curve (AUC-ROC), which plots the true-positive rate (y-coordinate) against the false-positive rate (x-coordinate) at various threshold settings.

For each ensemble model, we used the mean of the probabilities output by its 10 single models to compute the AUC-ROC. To calculate the AUC-ROC for the rule-based method, we assigned 0 to messages that did not contain hypoglycemia symptoms, 1 to messages that contained hypoglycemia symptoms but did not satisfy the criteria used by the rule-based method to select positive messages (see the Baseline Systems subsection above for the criteria), 2 to messages that satisfied the second criterion used by the rule-based method, and 3 to messages that satisfied the first criterion.

Accuracy

Accuracy is the number of correctly classified instances divided by the total number of instances. Although traditionally accuracy is the most common measure for classification, it is less effective and sometimes even improper when measuring performance on imbalanced classes [52,53]. In this paper, we provide accuracy for readers' interest, but we compared system performance based on other measures (eg, sensitivity and F1 score) [52,53].

XSL•FO
RenderX

Feature Analysis

We conducted feature ablation experiments to examine the effects of features. Specifically, we first selected the 3 best variants of HypoDetect that used different machine learning algorithms and different strategies to address data imbalance. We then compared these systems with their counterparts that dropped each single type of feature respectively.

To gain some understanding of the effects of individual features, we used a hybrid method to identify indicative features. Our method was motivated by the fact that our system, like typical NLP systems, uses a large number of features that are potentially redundant and may have dependencies among each other. As a result, the feature weights from the full model that used all the features may not accurately reflect a feature's impact. To address this problem, our method took into account the feature's adjusted (when used in the full model) and unadjusted (when used alone) effects. Specifically, we first used the best full model to identify 100 features with the largest positive feature weights. We then evaluated 100 single-feature models (which used the same machine learning method as the full model and used only 1 feature) through cross-validation and ranked the 100 features based on the corresponding F1 scores.

Error Analysis

To identify sources of errors, we analyzed hundreds of false-positive and false-negative instances that were predicted with high confidence by the 3 best variants of HypoDetect.

Results

Secure Messages and Patient-Reported Hypoglycemia Events

The interannotator agreement between the 2 annotators on the 100 secure message threads was Cohen kappa=.976.

Our dataset contained 3000 secure messages. The distribution of the number of words contained in these messages was right skewed (the first figure in Multimedia Appendix 5), with a median length of 92 (interquartile range 49-168) words. A total of 2850 (95.00%) of the 3000 messages had fewer than 435 words, and 114 (3.80%) messages were annotated as positive, indicating that the data were highly imbalanced.

Diabetic patients reported both mild and severe hypoglycemia incidents through secure messaging (see Textbox 1 and the first table in Multimedia Appendix 1). As Textbox 1 shows, patients wrote messages in diversified, informal ways (eg, "eating low carb" in example 1 and "blood sugar #" in example 2) and with typos (eg, "Gllipizide," "stablize," and "to much" in example 2). In addition, patients often elaborated on symptoms rather than directly reporting blood glucose levels.

Performance of Different HypoDetect Systems on the Evaluation Set

Corpus-level evaluation (Table 1) showed that logistic regression with class weighting achieved the best AUC-ROC (0.954) and F1 score (0.590). This classifier had a high specificity (0.974) and balanced sensitivity (0.693) and precision (0.513).

The 3 baseline machine learning systems (without treating data imbalance) consistently had very high specificity and very low sensitivity because they classified most examples as negative. Class weighting and oversampling (ROS-ensemble and SMOTE-ensemble) improved the baselines' sensitivity substantially (0.123-0.728 absolute gains) and their overall performance (as measured by the F1 score and AUC-ROC). Class weighting worked best for linear SVMs and logistic regression, whereas SMOTE-ensemble worked best for random forest.

Undersampling (RUS-ensemble in Table 1) boosted the baselines' sensitivity even higher but dropped their specificity and precision substantially. The rule-based method had higher sensitivity than the baseline machine learning systems but had lower performance than systems using class weighting or oversampling for all the metrics.

The fold-level evaluation showed similar results (see the first and second tables and the second figure in Multimedia Appendix 5). The individual classifiers used by an ensembled oversampling model had similar performance (the third table in Multimedia Appendix 5).

Effects of Features

We tested the effects of features on the 3 best variants of HypoDetect, namely linear SVMs with class weighting, random forest with SMOTE-ensemble, and logistic regression with class weighting. The results (Table 2) showed that dropping TF-IDF or domain-relevance features decreased the comprehensive metrics (AUC-ROC, F1 score, and accuracy) of all 3 systems and also decreased most single metrics (especially precision and specificity). Dropping topic features had mixed results. It decreased most metrics for logistic regression with class weighting and random forest with SMOTE-ensemble but increased most metrics for linear SVMs with class weighting.

Textbox 1. Excerpts from 2 secure messages reporting incidents of hypoglycemia.

Example 1: "Can you tell me what glucose level is too low? The last couple of nights, I've woken up in the middle of the night sweating profusely and shaky. I got up to check my blood sugar and it's been 63 both nights. Is that too low? After testing, I eat a snack and test again, and it goes up, to 73 Wednesday night/Thursday morning and to 70 Thursday night/Friday morning. I'm dieting, I've lost 7 pounds since the first of the month, but I'm not really eating low carb. This evening before I go to bed, I'm going to test my blood and if it's low, eat something before going to bed."

Example 2: "I took Gllipizide in the am before breakfast and one before dinner. Last night my blood sugars took a dive. I went to sleep and around 11 I woke up sweating and clammy. I took my blood sugar # and it had dropped to 57. My wife quickly brought me sugar tablets and I was able to stablize them at 80. I think this is to much medication."

XSL•FO
RenderX

Table 1. Performance of 3 variants of HypoDetect systems on the evaluation set.

Systems	AUC-ROC[a]	Precision	Sensitivity (recall)	Specificity	F1 score	Accuracy
Rule-based method	0.815	0.284	0.491	0.951	0.360	0.934
Linear support vector machines						
Baseline	0.945	0.614	0.377	0.991	0.467	0.966
Class weighting	0.952	0.529	0.561	0.980	0.545	0.964
RUS-ensemble[b]	0.949	0.198	0.921	0.852	0.326	0.855
ROS-ensemble[c]	0.950	0.559	0.500	0.984	0.528	0.966
SMOTE-ensemble[d]	0.951	0.564	0.500	0.985	0.530	0.966
Random forest						
Baseline	0.942	0.000	0.000	1.000	0.000	0.962
Class weighting	0.927	0.428	0.570	0.970	0.489	0.955
RUS-ensemble	0.928	0.143	0.904	0.787	0.248	0.791
ROS-ensemble	0.931	0.318	0.728	0.938	0.443	0.930
SMOTE-ensemble	0.942	0.486	0.596	0.975	0.535	0.961
Logistic regression						
Baseline	0.947	0.660	0.307	0.994	0.419	0.968
Class weighting	0.954	0.513	0.693	0.974	0.590	0.963
RUS-ensemble	0.946	0.192	0.912	0.849	0.318	0.851
ROS-ensemble	0.951	0.536	0.526	0.982	0.531	0.965
SMOTE-ensemble	0.951	0.566	0.552	0.983	0.559	0.943

[a]AUC-ROC: area under the receiver operating characteristic curve.

[b]RUS-ensemble: ensemble models using random undersampling.

[c]ROS-ensemble: ensemble models using random oversampling.

[d]SMOTE-ensemble: ensemble models using synthetic minority oversampling technique.

XSL•FO
RenderX

Table 2. Performance of different HypoDetect systems implemented by using all types of features or by respectively dropping each individual type of feature.

Systems	AUC-ROC[a]	Precision	Sensitivity (recall)	Specificity	F1 score	Accuracy
Linear support vector machines with class weighting						
All	0.952	0.529	0.561	0.980	0.545	0.964
Without TF-IDF[b]	0.920	0.263	0.737	0.919	0.388	0.912
Without topic	0.949	0.569	0.579	0.983	0.574	0.967
Without domain relevance	0.928	0.348	0.623	0.954	0.447	0.941
Random forest with SMOTE-ensemble[c]						
All	0.942	0.486	0.596	0.975	0.535	0.961
Without TF-IDF	0.938	0.364	0.632	0.956	0.462	0.944
Without topic	0.935	0.392	0.640	0.961	0.487	0.949
Without domain relevance	0.901	0.365	0.237	0.984	0.287	0.955
Logistic regression with class weighting						
All	0.954	0.513	0.693	0.974	0.590	0.963
Without TF-IDF	0.917	0.248	0.754	0.910	0.373	0.904
Without topic	0.950	0.500	0.640	0.975	0.561	0.962
Without domain relevance	0.901	0.437	0.579	0.971	0.498	0.956

[a]AUC-ROC: area under the receiver operating characteristic curve.

[b]TF-IDF: term frequency-inverse document frequency.

[c]SMOTE-ensemble: ensemble models using synthetic minority oversampling technique.

We used the best model—that is, logistic regression with class weighting—to analyze fine-grained feature effects. The results from this analysis showed that “low,” “sweating,” “shaking,” “sugar,” and “took” were among the top-10 features. Other top-10 features included 2 topic features (corresponding to topics 37 and 49 in the first table in Multimedia Appendix 3) and 3 domain-relevance features (corresponding to domain-specific topics 3, 4, and 5 in the second table in Multimedia Appendix 3).

Discussion

Principal Findings

We developed HypoDetect, an NLP system that automatically detects patient-reported hypoglycemia incidents from secure messages to facilitate early response from health care providers. Despite the challenges caused by imbalanced data and informal language use by patients, HypoDetect using logistic regression with class weighting achieved an AUC-ROC of 0.954 and F1 score of 0.590 on the evaluation set. This system had a high specificity (0.974) and a moderate sensitivity (0.693).

The F1 score is often used to assess the system’s capability to tolerate data imbalance because it is sensitive to data imbalance. This score is usually much lower on imbalanced datasets than on balanced ones [28,54]. F1 scores reported by previous studies on highly imbalanced datasets typically ranged between 0.3 and 0.5 [31,51]. Therefore, the F1 score of 0.590 achieved by our system is very promising.

Our work has clinical relevance. As introduced previously, secure messaging is intended for exchanging nonurgent information. Secure messaging also follows a triage process. The messages are viewed first by a nurse. If he or she determines it is necessary, the message will be forwarded to the clinician for review. This process can lead to both underreporting of and delayed responses to hypoglycemia events. First, the nurse may address a secure message about hypoglycemia and then close out the message, so that the clinician is not aware of the incident and the incident is not recorded in the patient’s record. Second, the triage process means that responses from clinical teams could be delayed, putting patients at higher risk for severe consequences. Our system has the potential to serve as a surveillance tool to support a proactive and timely response in such situations and, therefore, improve patient safety.

Previous work predicted the occurrence of hypoglycemia in a future period by learning from physiological data, such as monitored glucose levels and heart rate variability [55-60]. In contrast, our goal was to identify hypoglycemia events that have already happened and have been reported by patients. The inputs for our system were patients’ descriptions about the adverse events, which were mainly symptoms and often did not contain information about blood sugar levels.

Previous work on automatic classification of secure messages focused on information needs and did not address data imbalance [25-27]. Our work contributed to this literature by introducing a new task and by investigating strategies for treating data imbalance. Paragraph embeddings were shown useful for classifying information needs in secure messages [27]. In the

XSL•FO
RenderX

future, we will study the effects of using document embeddings as learning features for our task.

There has been active research in using NLP to detect or facilitate manual review of adverse drug events in unstructured electronic health record notes [40,61-64]. The prior work identified adverse events at the entity (eg, medical terms representing side effects of a drug) or relation (eg, a pair of terms that represent a drug and its side effects) level. In this study, we annotated hypoglycemia events at the message-thread level because patient-reported hypoglycemia events often lacked pivot terms and were composed of a set of symptoms and pertinent context. Sentence-level annotation may further improve system performance, which we will explore in the future.

Effects of Treating Data Imbalance

Our results showed that cost-sensitive learning (ie, class weighting) and SMOTE-ensemble were most effective in boosting system performance on imbalanced data (Table 1). Without treating data imbalance, the baseline systems failed to detect most positive examples. Class weighting and oversampling improved the sensitivity of all 3 variants of HypoDetect substantially. As a tradeoff, the specificity and precision decreased to a certain extent, but the overall effects (as measured by F1 score and AUC-ROC) were positive. Because oversampling increases the size of training data, it is computationally more expensive than class weighting.

Oversampling (ROS-ensemble and SMOTE-ensemble in Table 1) performed much better than undersampling (RUS-ensemble in Table 1) on our task when measured by the AUC-ROC and F1 score that consider both false positives and false negatives. Like typical text classifiers that use word features, our system uses thousands of features. However, the training set created by undersampling contained only about 200 examples, which was likely too few to train the system.

Patient-reported hypoglycemia needs to be evaluated quickly to avoid severe consequences. Therefore, systems with low sensitivity (eg, the baseline systems in our study) cannot be used for surveillance. On the other hand, systems with high sensitivity but low precision (eg, the systems using ensembled undersampling) would generate many false alarms, adding undue burden on already time-strapped health care providers. An ideal system for hypoglycemia detection would have high sensitivity and precision. To achieve this goal, we will explore ensemble methods that combine different types of systems (eg, systems with high sensitivity and systems with high precision) in our future work.

Effects of Features

Our results showed that TF-IDF, topic features, and domain relevance all contributed to system performance.

TF-IDF has been widely used for text classification. However, one disadvantage of TF-IDF is that it ignores semantic information and treats words with the same or similar meanings as separate entities. As a result, there are often thousands of TF-IDF features, posing challenges for machine learning when the training set is of small or moderate size.

Topic features cluster terms into a small set of semantically related groups, which helps alleviate the data sparseness problem to a certain extent. Topic features and their variants have proved useful for text classification, including categorizing clinical reports [65-67]. However, automatically induced topics may not be accurate and may lose fine-grained information for document classification. Therefore, a combination of both types of features is likely more robust.

The domain-relevance features are new features that we designed for this task. Our results suggested that knowledge-driven features can effectively improve system performance for domain-specific classification tasks.

Error Analysis

First, the systems often failed on cases that required discourse-level comprehension or human knowledge. For example, they tended to classify positive messages as negative if the messages contained irrelevant information. They also often classified negative messages as positive if the messages mentioned “blood sugar” and symptoms that looked similar to but were not caused by hypoglycemia (eg, example 4 in the second table in Multimedia Appendix 1). We expect that more annotated data will help reduce this type of error.

Second, our systems did not have specific treatments on negation and questions and therefore could be confused by messages that mentioned symptomatic terms in a negative mode or mentioned “low sugar level” or “hypoglycemia” in hypothetical questions (eg, examples 2 and 3 in the second table in Multimedia Appendix 1). Negation and question detection systems could be integrated to reduce this type of error.

Third, the systems often failed to extract glucose testing results when patients reported these numbers in informal ways. Example 4 in the first table in Multimedia Appendix 1 is a typical example, where the patient mentioned “This morning I had a 66.” One way to reduce this type of error is to develop another classifier to judge whether the blood sugar level is normal or abnormal and then use the classification results as a feature for our hypoglycemia classifier.

Limitations

Within the scope of this study, we annotated hypoglycemia incidents based on information solely in secure messages and treated instances that lacked clear context as negative. Clinic visit notes from around the time of the message may provide more information to reduce uncertainty, which we will study in the future. In this study, we used keywords to sample more positive examples because the positive examples retrieved by random sampling were too few to train and evaluate supervised systems. This strategy may affect the system’s performance in a real-world setting. We included the original, randomly sampled 1000 messages (with 1 positive example only) into our training data as a way to alleviate this problem.

Conclusions and Future Work

We developed HypoDetect, an NLP system to automatically identify patient-reported hypoglycemia incidents from secure messages to facilitate early response and corrective actions by clinical teams. Despite the challenge of data imbalance,

XSL•FO
RenderX

HypoDetect using class weighting or SMOTE-ensemble achieved promising results on this task. In future, we will investigate advanced data-driven methods, including active learning and document embeddings, to improve HypoDetect.

Acknowledgments

This work was supported by Institutional National Research Service Award (T32) 5T32HL120823-02 from the US National Institutes of Health (NIH), NIH grant R01HL125089, and NIH implementation research training program (R25t) grant 5R25CA172009-4. We also acknowledge the support of resources and facilities from Bedford VA Medical Center, Bedford, MA, USA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, the US Department of Veterans Affairs, or the US Government.

Authors' Contributions

HY and JC designed the study and are co-corresponding authors for this work. JC developed and evaluated the HypoDetect system, conducted the experiments and results analysis, and drafted the manuscript. JL, HY, ED, and WL contributed to manuscript writing. HY provided important intellectual inputs on experimental design and results analysis. JL contributed to the development of baseline systems and learning features. WL contributed to data collection and feature generation. ED and EG contributed to corpus creation, design of the rule-based method, and data analysis. VGV provided important intellectual input on the clinical relevance of the system. All authors contributed substantially to paper revision. Please contact HY (hong.yu@umassmed.edu) for data queries and JC for other queries.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Annotation guideline and examples of positive and negative instances for hypoglycemia incidents in our dataset.

[PDF File (Adobe PDF File), 141KB - jmir_v21i3e11990_app1.pdf]

Multimedia Appendix 2

Machine learning algorithms used by the HypoDetect system.

[PDF File (Adobe PDF File), 166KB - jmir_v21i3e11990_app2.pdf]

Multimedia Appendix 3

Supplemental information about features used in this study.

[PDF File (Adobe PDF File), 105KB - jmir_v21i3e11990_app3.pdf]

Multimedia Appendix 4

Keywords used by the rule-based method.

[PDF File (Adobe PDF File), 87KB - jmir_v21i3e11990_app4.pdf]

Multimedia Appendix 5

Supplemental results from this study.

[PDF File (Adobe PDF File), 303KB - jmir_v21i3e11990_app5.pdf]

References

1. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017 Jun;128:40-50. [doi: 10.1016/j.diabres.2017.03.024] [Medline: 28437734]
2. Miller CD, Phillips LS, Ziemer DC, Gallina DL, Cook CB, El-Kebbi IM. Hypoglycemia in patients with type 2 diabetes mellitus. Arch Intern Med 2001 Jul 09;161(13):1653-1659. [Medline: 11434798]
3. Gabriely I, Shamoon H. Hypoglycemia in diabetes: common, often unrecognized. Cleve Clin J Med 2004 Apr;71(4):335-342. [Medline: 15117175]

XSL•FO
RenderX

4. Zammitt NN, Frier BM. Hypoglycemia in type 2 diabetes: pathophysiology, frequency, and effects of different treatment modalities. Diabetes Care 2005 Dec;28(12):2948-2961. [Medline: 16306561]
5. UK Hypoglycaemia Study Group. Risk of hypoglycaemia in types 1 and 2 diabetes: effects of treatment modalities and their duration. Diabetologia 2007 Jun;50(6):1140-1147. [doi: 10.1007/s00125-007-0599-y] [Medline: 17415551]
6. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
7. Osborn CY, Rosenbloom ST, Stenner SP, Anders S, Muse S, Johnson KB, et al. MyHealthAtVanderbilt: policies and procedures governing patient portal functionality. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i18-i23 [FREE Full text] [doi: 10.1136/amiajnl-2011-000184] [Medline: 21807648]
8. 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;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
9. Griffin A, Skinner A, Thornhill J, Weinberger M. Patient portals: Who uses them? What features do they use? And do they reduce hospital readmissions? Appl Clin Inform 2016;7(2):489-501. [doi: 10.4338/ACI-2016-01-RA-0003] [Medline: 27437056]
10. Cronin RM, Davis SE, Shenson JA, Chen Q, Rosenbloom ST, Jackson GP. Growth of secure messaging through a patient portal as a form of outpatient interaction across clinical specialties. Appl Clin Inform 2015;6(2):288-304 [FREE Full text] [doi: 10.4338/ACI-2014-12-RA-0117] [Medline: 26171076]
11. Shenson JA, Cronin RM, Davis SE, Chen Q, Jackson GP. Rapid growth in surgeons' use of secure messaging in a patient portal. Surg Endosc 2016 Apr;30(4):1432-1440 [FREE Full text] [doi: 10.1007/s00464-015-4347-y] [Medline: 26123340]
12. Goldzweig CL, Towfigh AA, Paige NM, Orshansky G, Haggstrom DA, Beroes JM, et al. Systematic Review: Secure Messaging Between Providers and Patients, and Patients' Access to Their Own Medical Record: Evidence on Health Outcomes, Satisfaction, Efficiency and Attitudes. Washington, DC: Department of Veterans Affairs (US); Jul 2012.
13. Simon GE, Ralston JD, Savarino J, Pabiniak C, Wentzel C, Operskalski BH. Randomized trial of depression follow-up care by online messaging. J Gen Intern Med 2011 Jul;26(7):698-704. [doi: 10.1007/s11606-011-1679-8] [Medline: 21384219]
14. Zhou YY, Kanter MH, Wang JJ, Garrido T. Improved quality at Kaiser Permanente through e-mail between physicians and patients. Health Aff (Millwood) 2010 Jul;29(7):1370-1375 [FREE Full text] [doi: 10.1377/hlthaff.2010.0048] [Medline: 20606190]
15. Ralston JD, Hirsch IB, Hoath J, Mullen M, Cheadle A, Goldberg HI. Web-based collaborative care for type 2 diabetes: a pilot randomized trial. Diabetes Care 2009 Feb;32(2):234-239 [FREE Full text] [doi: 10.2337/dc08-1220] [Medline: 19017773]
16. Green BB, Cook AJ, Ralston JD, Fishman PA, Catz SL, Carlson J, et al. Effectiveness of home blood pressure monitoring, Web communication, and pharmacist care on hypertension control: a randomized controlled trial. JAMA 2008 Jun 25;299(24):2857-2867 [FREE Full text] [doi: 10.1001/jama.299.24.2857] [Medline: 18577730]
17. McMahon GT, Gomes HE, Hickson HS, Hu TM, Levine BA, Conlin PR. Web-based care management in patients with poorly controlled diabetes. Diabetes Care 2005 Jul;28(7):1624-1629 [FREE Full text] [Medline: 15983311]
18. Ross SE, Moore LA, Earnest MA, Wittevrongel L, Lin C. Providing a web-based online medical record with electronic communication capabilities to patients with congestive heart failure: randomized trial. J Med Internet Res 2004 May 14;6(2):e12 [FREE Full text] [doi: 10.2196/jmir.6.2.e12] [Medline: 15249261]
19. Wade-Vuturo AE, Mayberry LS, Osborn CY. Secure messaging and diabetes management: experiences and perspectives of patient portal users. J Am Med Inform Assoc 2013 May 1;20(3):519-525 [FREE Full text] [doi: 10.1136/amiajnl-2012-001253] [Medline: 23242764]
20. Shimada SL, Petrakis BA, Rothendler JA, Zirkle M, Zhao S, Feng H, et al. An analysis of patient-provider secure messaging at two Veterans Health Administration medical centers: message content and resolution through secure messaging. J Am Med Inform Assoc 2017 Sep 01;24(5):942-949. [doi: 10.1093/jamia/ocx021] [Medline: 28371896]
21. Lanham HJ, Leykum LK, Pugh JA. Examining the complexity of patient-outpatient care team secure message communication: qualitative analysis. J Med Internet Res 2018 Jul 11;20(7):e218 [FREE Full text] [doi: 10.2196/jmir.9269] [Medline: 29997107]
22. Nazi KM, Woods SS, Woods SS. MyHealtheVet PHR: a description of users and patient portal use. AMIA Annu Symp Proc 2008:1182. [Medline: 18999142]
23. My HealtheVet metrics: 1st quarter 2018 (January-March). Washington, DC: US Department of Veterans Affairs; 2018. URL: https://www.myhealth.va.gov/mhv-portal-web/webusers [accessed 2018-04-14] [WebCite Cache ID 6yg7RqfkN]
24. Robinson JR, Valentine A, Carney C, Fabbri D, Jackson GP. Complexity of medical decision-making in care provided by surgeons through patient portals. J Surg Res 2017 Jun 15;214:93-101 [FREE Full text] [doi: 10.1016/j.jss.2017.02.077] [Medline: 28624066]
25. Cronin RM, Fabbri D, Denny JC, Jackson GP. Automated classification of consumer health information needs in patient portal messages. AMIA Annu Symp Proc 2015;2015:1861-1870 [FREE Full text] [Medline: 26958285]

XSL•FO
RenderX

26. Cronin RM, Fabbri D, Denny JC, Rosenbloom ST, Jackson GP. A comparison of rule-based and machine learning approaches for classifying patient portal messages. Int J Med Inform 2017 Dec;105:110-120. [doi: 10.1016/j.ijmedinf.2017.06.004] [Medline: 28750904]
27. Sulieman L, Gilmore D, French C, Cronin RM, Jackson GP, Russell M, et al. Classifying patient portal messages using Convolutional Neural Networks. J Biomed Inform 2017 Oct;74:59-70. [doi: 10.1016/j.jbi.2017.08.014] [Medline: 28864104]
28. He H, Garcia EA. Learning from imbalanced data. IEEE Trans Knowl Data Eng 2009 Sep;21(9):1263-1284. [doi: 10.1109/TKDE.2008.239]
29. Zahirnia K, Teimouri M, Rahmani R, Salaq A. Diagnosis of type 2 diabetes using cost-sensitive learning. 2015 Presented at: 5th International Conference on Computer and Knowledge Engineering; Oct 29-30, 2015; Mashhad, Iran p. 158-163.
30. Krawczyk B, Galar M, Jeleń L, Herrera F. Evolutionary undersampling boosting for imbalanced classification of breast cancer malignancy. Appl Soft Comput 2016 Jan;38:714-726. [doi: 10.1016/j.asoc.2015.08.060]
31. Ramezankhani A, Pournik O, Shahrabi J, Azizi F, Hadaegh F, Khalili D. The impact of oversampling with SMOTE on the performance of 3 classifiers in prediction of type 2 diabetes. Med Decis Making 2016 Dec;36(1):137-144. [doi: 10.1177/0272989X14560647] [Medline: 25449060]
32. Blagus R, Lusa L. Joint use of over- and under-sampling techniques and cross-validation for the development and assessment of prediction models. BMC Bioinformatics 2015 Nov 04;16:363 [FREE Full text] [doi: 10.1186/s12859-015-0784-9] [Medline: 26537827]
33. National Institute of Diabetes and Digestive and Kidney Diseases. Low blood glucose (hypoglycemia). Bethesda, MD: US Department of Health and Human Services; 2016 Aug. URL: https://www.niddk.nih.gov/health-information/diabetes/overview/preventing-problems/low-blood-glucose-hypoglycemia [accessed 2018-08-08] [WebCite Cache ID 71WXtbFPp]
34. Workgroup on Hypoglycemia, American Diabetes Association. Defining and reporting hypoglycemia in diabetes: a report from the American Diabetes Association Workgroup on Hypoglycemia. Diabetes Care 2005 May;28(5):1245-1249. [Medline: 15855602]
35. Weiss GM. Mining with rarity: a unifying framework. SIGKDD Explor Newsl 2004 Jun 01;6(1):7. [doi: 10.1145/1007730.1007734]
36. Krawczyk B. Learning from imbalanced data: open challenges and future directions. Prog Artif Intell 2016 Apr 22;5(4):221-232. [doi: 10.1007/s13748-016-0094-0]
37. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002 Jun 01;16:321-357. [doi: 10.1613/jair.953]
38. Fernandez A, Garcia S, Herrera F, Chawla NV. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. J Artif Intell Res 2018 Apr 20;61:863-905. [doi: 10.1613/jair.1.11192]
39. Goldberg Y. A primer on neural network models for natural language processing. J Artif Intell Res 2016 Nov 20;57:345-420. [doi: 10.1613/jair.4992]
40. Munkhdalai T, Liu F, Yu H. Clinical relation extraction toward drug safety surveillance using electronic health record narratives: classical learning versus deep learning. JMIR Public Health Surveill 2018 Apr 25;4(2):e29 [FREE Full text] [doi: 10.2196/publichealth.9361] [Medline: 29695376]
41. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: 10.1007/BF00994018]
42. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
43. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12(Oct):2825-2830 [FREE Full text]
44. Arora R, Ravindran B. Latent Dirichlet allocation based multi-document summarization. 2008 Presented at: Second Workshop on Analytics for Noisy Unstructured Text Data; July 24, 2008; Singapore p. 91-97. [doi: 10.1145/1390749.1390764]
45. McCallum AK. MALLET: machine learning for language toolkit. 2002. URL: http://mallet.cs.umass.edu [accessed 2016-07-04] [WebCite Cache ID 6il7RNCwf]
46. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed representations of words and phrases and their compositionality. 2013 Presented at: 27th Conference on Neural Information Processing Systems (NIPS 2013); Dec 5-10, 2013; Lake Tahoe, NV, USA p. 3111-3119.
47. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. 2013 Jan 16. URL: http://arxiv.org/abs/1301.3781 [accessed 2016-11-18] [WebCite Cache ID 6m6NhZqFz]
48. Jagannatha AN, Chen J, Yu H. Mining and ranking biomedical synonym candidates from Wikipedia. 2015 Presented at: Sixth International Workshop on Health Text Mining and Information Analysis (Louhi); Sep 17, 2015; Lisbon, Portugal p. 142-151.
49. Pyysalo S, Ginter F, Moen H, Salakoski T, Ananiadou S. Distributional semantics resources for biomedical text processing. 2013 Presented at: Proceedings of the 5th International Symposium on Languages in Biology and Medicine (LBM 2013); December 12-13, 2013; Tokyo, Japan p. 39-43 URL: http://bio.nlplab.org/pdf/pyysalo13literature.pdf
50. Chen J, Zheng J, Yu H. Finding important terms for patients in their electronic health records: a learning-to-rank approach using expert annotations. JMIR Med Inform 2016 Nov 30;4(4):e40 [FREE Full text] [doi: 10.2196/medinform.6373] [Medline: 27903489]

XSL•FO
RenderX

51. Liu X, Wu J, Zhou Z. Exploratory undersampling for class-imbalance learning. IEEE Trans Syst Man Cybern B Cybern 2009 Apr;39(2):539-550. [doi: 10.1109/TSMCB.2008.2007853] [Medline: 19095540]
52. López V, Fernández A, García S, Palade V, Herrera F. An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics. Inf Sci 2013 Nov;250:113-141. [doi: 10.1016/j.ins.2013.07.007]
53. Sun Y, Wong AK, Kamel MS. Classification of imbalanced data: a review. Int J Pattern Recognit Artif Intell 2009 Jun;23(04):687-719. [doi: 10.1142/S0218001409007326]
54. Jeni LA, Cohn JF, De La Torre F. Facing imbalanced data recommendations for the use of performance metrics. Int Conf Affect Comput Intell Interact Workshops 2013;2013:245-251 [FREE Full text] [doi: 10.1109/ACII.2013.47] [Medline: 25574450]
55. Elvebakk O, Tronstad C, Birkeland KI, Jenssen TG, Bjørgaas MR, Frøslie KF, et al. Evaluation of hypoglycaemia with non-invasive sensors in people with type 1 diabetes and impaired awareness of hypoglycaemia. Sci Rep 2018 Oct 03;8(1):14722 [FREE Full text] [doi: 10.1038/s41598-018-33189-1] [Medline: 30283093]
56. Sudharsan B, Peeples M, Shomali M. Hypoglycemia prediction using machine learning models for patients with type 2 diabetes. J Diabetes Sci Technol 2015 Jan;9(1):86-90 [FREE Full text] [doi: 10.1177/1932296814554260] [Medline: 25316712]
57. Cichosz SL, Frystyk J, Hejlesen OK, Tarnow L, Fleischer J. A novel algorithm for prediction and detection of hypoglycemia based on continuous glucose monitoring and heart rate variability in patients with type 1 diabetes. J Diabetes Sci Technol 2014 Jul;8(4):731-737 [FREE Full text] [doi: 10.1177/1932296814528838] [Medline: 24876412]
58. Nguyen LL, Su S, Nguyen HT. Identification of hypoglycemia and hyperglycemia in type 1 diabetic patients using ECG parameters. Conf Proc IEEE Eng Med Biol Soc 2012;2012:2716-2719. [doi: 10.1109/EMBC.2012.6346525] [Medline: 23366486]
59. Cameron F, Niemeyer G, Gundy-Burlet K, Buckingham B. Statistical hypoglycemia prediction. J Diabetes Sci Technol 2008 Jul;2(4):612-621 [FREE Full text] [doi: 10.1177/193229680800200412] [Medline: 19885237]
60. Murata GH, Hoffman RM, Shah JH, Wendel CS, Duckworth WC. A probabilistic model for predicting hypoglycemia in type 2 diabetes mellitus: the Diabetes Outcomes in Veterans Study (DOVES). Arch Intern Med 2004 Jul 12;164(13):1445-1450. [doi: 10.1001/archinte.164.13.1445] [Medline: 15249354]
61. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. J Am Med Inform Assoc 2005;12(4):448-457 [FREE Full text] [doi: 10.1197/jamia.M1794] [Medline: 15802475]
62. Haerian K, Varn D, Vaidya S, Ena L, Chase HS, Friedman C. Detection of pharmacovigilance-related adverse events using electronic health records and automated methods. Clin Pharmacol Ther 2012 Aug;92(2):228-234 [FREE Full text] [doi: 10.1038/clpt.2012.54] [Medline: 22713699]
63. Jagannatha A, Yu H. Bidirectional RNN for medical event detection in electronic health records. 2016 Presented at: 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; Jun 12-17, 2016; San Diego, CA, USA p. 473-482 URL: https://www.aclweb.org/anthology/N/N16/N16-1056.pdf
64. Li F, Liu W, Yu H. Extraction of information related to adverse drug events from electronic health record notes: design of an end-to-end model based on deep learning. JMIR Med Inform 2018 Nov 26;6(4):e12159 [FREE Full text] [doi: 10.2196/12159] [Medline: 30478023]
65. Sriurai W. Improving text categorization by using a topic model. Adv Comput 2011 Nov 30;2(6):21-27. [doi: 10.5121/acij.2011.2603]
66. Sarioglu E, Yadav K, Choi HA. Topic modeling based classification of clinical reports. 2013 Presented at: 51st Annual Meeting of the Association for Computational Linguistics Proceedings of the Student Research Workshop; Aug 5-7, 2013; Sofia, Bulgaria p. 67-73.
67. Pavlinek M, Podgorelec V. Text classification method based on self-training and LDA topic models. Expert Syst Appl 2017 Sep;80:83-93. [doi: 10.1016/j.eswa.2017.03.020]

Abbreviations

AUC-ROC: area under the receiver operating characteristic curve
HypoDetect: Hypoglycemia Detector
ICD-9: International Classification of Diseases, Ninth Revision
ICD-10: International Classification of Diseases, Tenth Revision
MALLET: Machine Learning for Language Toolkit
NLP: natural language processing
SMOTE: synthetic minority oversampling technique
SVM: support vector machine
TF-IDF: term frequency-inverse document frequency
VA: Veterans Affairs

XSL•FO
RenderX

Edited by G Eysenbach; submitted 21.08.18; peer-reviewed by R Cronin, L Sulieman; comments to author 11.10.18; revised version received 19.01.19; accepted 10.02.19; published 11.03.19.

Please cite as:
Chen J, Lalor J, Liu W, Druhl E, Granillo E, Vimalananda VG, Yu H
Detecting Hypoglycemia Incidents Reported in Patients' Secure Messages: Using Cost-Sensitive Learning and Oversampling to Reduce Data Imbalance
J Med Internet Res 2019;21(3):e11990
URL: http://www.jmir.org/2019/3/e11990/
doi:10.2196/11990
PMID:30855231

©Jinying Chen, John Lalor, Weisong Liu, Emily Druhl, Edgard Granillo, Varsha G Vimalananda, Hong Yu. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 11.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Artificial Intelligence and the Future of Primary Care: Exploratory Qualitative Study of UK General Practitioners' Views

Charlotte Blease[1,2], PhD; Ted J Kaptchuk[1], OMD; Michael H Bernstein[3], PhD; Kenneth D Mandl[4,5,6], MD, MPH; John D Halamka[7,8], MD, PhD; Catherine M DesRoches[9], DrPH

[1]General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[2]School of Psychology, University College Dublin, Dublin, Ireland

[3]School of Public Health, Brown University, Providence, RI, United States

[4]Computational Health Informatics Program, Boston Children's Hospital, Harvard Medical School, Boston, MA, United States

[5]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, United States

[6]Department of Pediatrics, Harvard Medical School, Boston, MA, United States

[7]Beth Israel Deaconess Medical Center, Boston, MA, United States

[8]Brigham and Women's Hospital, Boston, MA, United States

[9]Open Notes, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Charlotte Blease, PhD
General Medicine and Primary Care
Beth Israel Deaconess Medical Center
Harvard Medical School
330 Brookline Avenue
Boston, MA, MA 02215
United States
Phone: 1 617 754 1457
Email: cblease@bidmc.harvard.edu

Abstract

Background: The potential for machine learning to disrupt the medical profession is the subject of ongoing debate within biomedical informatics and related fields.

Objective: This study aimed to explore general practitioners' (GPs') opinions about the potential impact of future technology on key tasks in primary care.

Methods: In June 2018, we conducted a Web-based survey of 720 UK GPs' opinions about the likelihood of future technology to fully replace GPs in performing 6 key primary care tasks, and, if respondents considered replacement for a particular task likely, to estimate how soon the technological capacity might emerge. This study involved qualitative descriptive analysis of written responses ("comments") to an open-ended question in the survey.

Results: Comments were classified into 3 major categories in relation to primary care: (1) limitations of future technology, (2) potential benefits of future technology, and (3) social and ethical concerns. Perceived limitations included the beliefs that communication and empathy are exclusively human competencies; many GPs also considered clinical reasoning and the ability to provide value-based care as necessitating physicians' judgments. Perceived benefits of technology included expectations about improved efficiencies, in particular with respect to the reduction of administrative burdens on physicians. Social and ethical concerns encompassed multiple, divergent themes including the need to train more doctors to overcome workforce shortfalls and misgivings about the acceptability of future technology to patients. However, some GPs believed that the failure to adopt technological innovations could incur harms to both patients and physicians.

Conclusions: This study presents timely information on physicians' views about the scope of artificial intelligence (AI) in primary care. Overwhelmingly, GPs considered the potential of AI to be limited. These views differ from the predictions of biomedical informaticians. More extensive, stand-alone qualitative work would provide a more in-depth understanding of GPs' views.

(J Med Internet Res 2019;21(3):e12802) doi:10.2196/12802

XSL•FO
RenderX

KEYWORDS

artificial intelligence; attitudes; future; general practice; machine learning; opinions; primary care; qualitative research; technology

Introduction

Background

Although debate about the future of medicine persists, much of the discussion still focuses on recurrent themes such as how health care is paid for and organizational management. In recent years, however, researchers working in the fields of artificial intelligence (AI) and biomedical informatics have begun to raise questions about the potential impact of technology on the medical workforce [1-3]. Although a minority of experts in these fields remain more skeptical that health care is on the cusp of a technological revolution [4], the overwhelming majority of informaticians predict that big data, machine learning, and innovations in AI are poised to significantly overhaul the delivery of medicine [5]. The consensus among these researchers is that core functions of medical professionals—including, but not limited to patient monitoring, diagnostics, and prognostics—will be transformed by technology [6-8].

How might these innovations affect the medical professions? To date, many informaticians forecast that the work carried out by radiologists and anatomical pathologists is likely to be outsourced to algorithms and moreover, that, "The time scale for these disruptions is years, not decades" [9]. When it comes to primary care, however, the predictions appear to be less clear-cut. Many AI researchers suggest that future technology will augment current work practices and eliminate the need for many routine patient visits, but in complex clinical cases, physicians will still be needed to coordinate care and provide empathic support to patients [10-12]. Other AI experts hint at a stronger forecast, suggesting that, in the long term, primary care is vulnerable to disintermediation with physicians eventually being replaced by machine learning algorithms and paraprofessionals [3,8,13].

In opposition to these views are the perspectives of primary care physicians, some of whom claim that the threat of AI is exaggerated. For example, Verghese and colleagues forcefully argue that technological innovation will not significantly encroach on general medicine and "concerns about physician 'unemployment' and 'de-skilling' are overblown" [14]. They argue that, even if computers provide more accurate diagnoses and prognoses than humans, physicians' clinical judgments will still be necessary in decision-making processes—as will their expertise in explaining medical information to patients and in the provision of care [14]. Even more emphatically, others have argued that the medical community must guard against, what are perceived to be, the damaging effects of current (and future) technology in general medicine [15]; the use of technology, it is claimed, threatens the quality of patient-centered care which, according to these commentators, necessitates a dyadic, face-to-face interaction between physicians and patients [16,17].

Objectives

Amid the debate and uncertainties surrounding the impact of AI on the future of the medical professions, we observe that limited attention has so far been paid to the views of practicing physicians [18-22]. Addressing this research gap, we employed quantitative methods to investigate UK-based general practitioners' (GPs') opinions about the potential impact of future technologies on primary care. Given the potential for finer-grained insights to be acquired using qualitative methods, we incorporated a single open-ended question into the survey. Our aim was to provide a preliminary investigation into GPs' perspectives on the themes of the quantitative survey: namely, the bearing of technological innovations on the future of their profession. To our knowledge, this is the first such exploratory investigation of GPs' opinions and attitudes about AI and the future of a medical specialism.

Methods

Main Survey

A complete description of the survey methods and quantitative results has been published previously [22]. In summary, we conducted an anonymous nationwide Web-based survey of UK GPs (response rate=48.84% [720/1474]). All procedures were approved by the ethics review board of University College Dublin; in addition, ethical exemption was approved by the institutional review board of Beth Israel Deaconess Medical Center, Harvard Medical School.

Participants were randomly sampled from membership of the clinician marketing service Doctors.net.uk [23]. This is the largest online medical network in the United Kingdom, and 86.95% (53,670/61,724) of registered and licensed UK GPs are members. We predicted a response rate of around 46% based on previous surveys using this platform [24-26]. Invitations were emailed and displayed on the Doctors.net.uk home pages of 1474 randomly selected GPs for 1 week (June 12-18, 2018). The sample was stratified according to gender and age using up-to-date demographic information about UK GPs provided by the General Medical Council [27]. Invited participants were advised that their identity would not be disclosed to the research team, and all respondents gave informed consent before participating. On completion, respondents were recompensed for the time with £10 (US $13, €11) worth of credit exchangeable for online shopping.

The study team devised an original survey instrument specifically designed to investigate GPs' opinions about the impact of future technologies on primary care (see Multimedia Appendix 1). We avoided terms such as "algorithms" and "machine learning" in favor of generic descriptors such as "machines" and "future technology." This was in part to avoid any confusion among physicians unfamiliar with this terminology and to avert technical debates about the explanatory adequacy of specific AI terms of art. The survey was developed in consultation with UK-based GPs (n=6) and primary care physicians in the United States (n=6) to ensure face validity.

XSL•FO
RenderX

Qualitative Component

To maximize response rate for the qualitative component, the survey instrument included a single open-ended question that allowed participants to respond in more detail on the topic of the questionnaire. Specifically, respondents were asked to "...provide any comments on the survey topic." Descriptive qualitative data analysis was used to investigate these responses [28,29].

We carried out inductive thematic coding of the data; in light of the limitations of the dataset, a full thematic analysis was not applicable [30]. Responses were collated and imported into QCAmap (coUnity Software Development GmbH) for analysis. The comment transcripts were initially read numerous times by CB to achieve familiarization with the participant responses. Afterward, an inductive coding process was employed in which brief descriptive labels ("codes") were applied to each comment. Multiple codes were applied to comments with multiple meanings. Comments and codes were reviewed and compared with investigate similarities and differences. CB and TJK met to discuss coding decisions, and subsequent revisions were made. First-order codes were grouped into second-order categories based on the commonality of their meaning to provide a descriptive summary of the responses.

Results

Overview

The survey had an overall response rate of 48.84% (of the 1474 GPs invited to participate, 720 responded) [22]. As outlined in the published quantitative survey, respondents were representative of UK GPs in terms of age and gender and from all regions of the United Kingdom [22]. A total of 9.1% of respondents (66/720) left comments (2096 words), which were typically brief (1 phrase or 1 or 2 sentences). GPs who submitted comments were not significantly different from those who did not both in terms of gender and whether they worked part time; however, older GPs (aged 45 years and older) were more likely to leave comments than younger GPs: 83% (55/66) of older GPs left comments, compared with 48.0% (314/654) of older GPs who did not; Table 1). A series of Mann-Whitney *U* tests verified that those who provided qualitative feedback did not differ from those who did not provide qualitative feedback, on the perceived likelihood that future technology would replace human doctors for any of the 6 tasks: p s>.19.

As a result of the iterative process of thematic analysis, 3 major categories of GPs' views were identified in relation to primary care: (1) limitations of future technology, (2) potential benefits of future technology, and (3) social and ethical issues. These categories were further subdivided into 8 themes, which are described below with illustrative comments; numbers in parentheses are identifiers that ascribe comments to individual participants.

Limitations of Future Technology in Primary Care

Empathy and Communication

Respondents' comments encompassed a number of perspectives reflecting perceived limitations of future technology in general practice. One of the main themes identified was the importance of face-to-face, human communication, with the concurrently held belief that technology cannot substitute for humans in providing empathic care. GPs claimed, for example:

> *It is unlikely that the human element of empathy and the subtlety of human communication and non-verbal cues can be detected by robots or machines.* [Participant 517]

> *Technology will never attain a personal relationship with patents. We are essentially a people business. It's personal relationships that count.* [Participant 45]

More strongly, some participants argued that physicians would be necessary to provide this care, and moreover, that patients would desire it. For example:

> *Technology will never be satisfactory as patients are looking for that interaction and dopamine squirt (doctor is the drug) which can only be achieved through empathic continuity of care of highly experienced General Practitioner specialists.* [Participant 686]

Clinical Reasoning

Multiple participants expressed skepticism about the capacity of technology to undertake processes related to clinical reasoning and diagnostic judgments. Some suggested that there was an ineffable aspect to medical decision making that renders it an intrinsically human pursuit:

> *Technology cannot replace doctors. There is definitely a 6th sense.* [Participant 635]

Notably, some GPs also implied—either directly or indirectly—that physicians are necessary to gather patient information; a process which, in turn, was deemed indispensable to diagnostics:

> *The other issue is the inability of a machine/AI to be able to skillfully ascertain the data required from a patient for correct analysis.* [Participant 453]

XSL•FO
RenderX

Table 1. Comparison of general practitioners who did (n=66) and did not (n=654) supply qualitative comments.

Characteristic	Qualitative comments	No qualitative comments	Comparison		
			Chi-square (*df*)	*t* test (*df*)	*P* value
Female (N=720), n (%)	28 (42)	295 (45)	0.2 (1)	Not applicable	.68
Part time, (N=720), n (%)	36 (55)	353 (54)	0.01 (1)	Not applicable	.93
Year of qualification, mean (SD)	1990.1 (9.81)	1996.1 (9.04)	Not applicable	5.15 (718)	<.001
Age≥45 years (N=720), n (%)	55 (83)	314 (48)	29.9 (1)	Not applicable	<.001

> *They [patients] are humans and subject to the vagaries of human recall, memory, and interpretation. AI may make it easier to interpret a blood result follow a protocol or order a test. But AI will always struggle when the same human can score 1/10 for a symptom today and 10/10 tomorrow.* [Participant 201]

Patient-Centeredness

Comments expressed varying degrees of cynicism about how technology might provide care that is respectful of, and responsive to, individual patient preferences, needs, and values. Some participants considered it improbable that AI could oversee shared decision making, deeming these aspects of patient care to require human competencies. For example:

> *Medicine, particularly general practice, is an art; listening to ideas concerns and expectations and negotiating a shared plan with the patient. Often doing nothing other than listening is required. I wonder how well a computer will be able to do this?* [Participant 285]

> *Technology won't replace GPs as patient management is about negotiation and managing risks and different patients have different views.* [Participant 703]

In summary: many participants appeared unpersuaded about the potential for technology to overhaul general practice. This was supported by comments that accentuated the intricacies of the work carried out by physicians. For example:

> *So much of ill health is vague, complicated and psychological, and the lack of IT to any time soon pick up on individualized, non-verbal cues etc. I feel will still leave a huge role for people.* [Participant 457]

Benefits of Future Technology in Primary Care

Although comments encompassed perceived benefits of AI advancements in primary care, notably these were frequently couched in wider perspectives about the predicted shortcomings of future technology. In the main, opinions suggested that such innovations would be restricted to improving performance within traditional GP roles.

Improved Efficiencies

A broad consideration was that future technology would improve GPs' productivity and workflow efficiencies. Some individuals expressed this viewpoint forcefully:

> *Please hurry up with the technological advances to take away some of the crap that I still have to sort out–then I will be able to get back to proper diagnosing and doctoring.* [Participant 693]

In this way, some GPs identified a positive paraprofessional role for technology in streamlining access to physicians:

> *Machines should be good at initial triage of uncomplicated patients presenting to primary care.* [Participant 88]

Supporting this perspective, some respondents were adamant that advancements in AI would buttress rather than replace the core roles of GPs and "help GPs with workload issues" [Participant 517].

An Administrative Role

Extending this viewpoint, many GPs expressed the idea that AI would reduce the burdens of paperwork and provide clerical assistance. For example:

> *Be useful to develop AI to do analyses of pathology returns, and read all the letters, to provide another presence in the consulting room, and to write the referral letters, organize investigations and the like, ie, act like a personal assistant might do.* [Participant 135]

> *I think technology's place is more about informing patients about conditions and management booking appointments ordering prescriptions contacting the surgery via the internet rather than the phone.* [Participant 683]

The dominant viewpoint was that technology will reduce the onuses of form-filling and provide administrative support.

Social and Ethical Concerns in Present and Future Primary Care

Understaffing

Interestingly, many participants chose not to interpret the question as directly asking about the impact of AI on the future of primary care, and instead, commented on the growing pressures on the GP workforce, including the risks that this was believed to pose to professionals and patients:

> *The only reason that I'm not burned out is that I reduced my workload and traded money for sanity.* [Participant 280]

> *I changed job due to stress as where I worked I had an unsafe workload* [Participant 608]

> *I hate the current stress due to understaffing which is so dangerous* [Participant 516]

XSL•FO
RenderX

Perhaps, consistent with the prevalent viewpoint that technology would play a limited role in primary care, some respondents stated that increased recruitment of physicians would relieve current demands on the GP workforce:

> *General Practice in the U.K. is teetering on the brink of collapse due to years of under investment in training adequate numbers of doctors to deal with an ageing population.* [Participant 331]
>
> *Not enough is done to retain experienced GPs. Would save NHS a fortune.* [Participant 635]

Acceptability of Artificial Intelligence

The social implications for patients of possible advancements in AI received much less attention: comments instead were focused on whether the public would be satisfied with, or open to new technology or different ways of obtaining primary care:

> *The question is whether they will be acceptable to patients although they may be very accessible compared to the current system.* [Participant 88]
>
> *The somewhat blunt tool of technology as it stands will need to evolve some way before the culture of clinicians and patients will accept it.* [Participant 453]

The Ethics of Innovation

Some GPs conveyed greater certainty that AI would lead to major disruptions within general practice though they were unspecific about the nature of these disruptions; for example:

> *Medicine will be unrecognizable compared to its present form in 25 years.* [Participant 282]

In respect of this, concerns about safety and accountability were touched on but not described in detail; for example:

> *Technology will be supporting clinicians in the very near future – the issue is responsibility and liability in legal terms for such tools.* [Participant 453]

Striking, and in opposition to the overriding outlook among participants, a few respondents considered the preservation of current work practices to be a source of present and future harm. For example:

> *Risk-taking is not admired or valued yet without it – or AI – general practice will be destroyed.* [Participant 464]
>
> *Technology and non-medical clinicians can replace GPs easily. My burnout is because of my frustration with colleagues and their Luddite working practices.* [Participant 495]

Discussion

Principal Findings

This initial qualitative study affords new insights into how UK GPs view the impact of future technology on primary care. The dominant perspective among respondents was one of skepticism—most GPs believed that the scope of technological innovations will be considerably restricted within general practice. Empathy and communication, in particular, were viewed as quintessentially human skills, and some respondents were adamant that patients will always desire physician-mediated care. Other participants considered doctor-patient interaction as necessary to the process of efficient medical information gathering; similarly, clinical acumen was often assumed to be a uniquely human expertise. GPs viewed the provision of patient-centered care as an interpersonal process that is unlikely to be threatened by automation. Reflecting these themes, expected benefits of AI were generally limited to efficiencies within current models of practice and in particular, to reducing administrative burdens.

Weighing in on wider social and ethical concerns, many GPs reported high levels of burnout, stress, and fears about unsafe workloads. Some comments forcefully expressed the view that greater investment in physician manpower is required. Taking a different perspective, other comments predicted that radical change to primary care was imminent with some GPs claiming that embracing technological innovations is an ethical responsibility to reduce workloads and prevent patient harm.

This exploratory survey suggests that GPs and informaticians are far apart in their views about the impact of machine learning in primary care. In contrast with many of the comments in our study, biomedical informaticians forecast that—both in the short- and long-term—the key functions of primary care will be radically transformed by AI [3,5,6,8,13,31,32]. Indeed, evidence challenges the assumption that physicians are necessary to gather health information [33,34]. Mobile health (mHealth) apps already allow patients to track and monitor a growing number of their own signs and symptoms (eg, blood glucose levels, blood pressure, and levels of physical activity) without the need for traditional checkups with their physician. For example, recent research indicates that home monitoring may be preferable for controlling and preventing chronic conditions: evidence from systematic reviews and meta-analyses of patients with type 1 and type 2 diabetes suggests that mHealth provides clear improvement over clinical and nonmobile interventions in glycemic control and patient self-management [33,34]. Similarly, a 10-year multicenter study of home monitoring for high blood pressure found that ambulatory tracking was not only more accurate but also safer than readings conducted in doctors' offices: the authors concluded that "white-coat hypertension is not benign" and can mask risk of hypertension among patients [35]. In summary: in contrast with GPs, AI health researchers predict that wearable devices with the capacity for real-time monitoring will improve precision in information gathering while also driving down unnecessary appointments and health care costs [36].

GPs also expressed broad cynicism about the prospects for AI undertaking diagnoses. Again, this perspective is diametrically opposed to the views of biomedical informaticians who argue that the accumulation of "big data"—the collection of vast amounts of information about individual patients (from the genomic and molecular levels, to information about diet, lifestyle, and other environmental factors)—when combined with machine learning, will yield more precise patterns about our individual health and medical outcomes and do so more quickly than humans are able to discern [5-7,9-11,35]. According to AI experts, the capacity to extract novel insights from large health scale data is "where machine learning shines,"

XSL•FO
RenderX

with the promise to afford discoveries of hitherto undetected subtypes of diseases [9]. Mining this information for regularities and patterns, and applying algorithmic predictions to new data, it is claimed, will lead to unprecedented personalized precision in diagnostics, prognostics, and treatment recommendations [7,9]. Indeed, aside from medical histories and patient reports, an exponentially increasing volume of health-related information generated from social media posts, apps, purchases, and credit card usage is already being used to support predictions about patient behavior and well-being [37]. In short, beyond the intentional use of digital health devices to undertake diagnostic and prognostic assessments, a vast range of nonmedical data are beginning to yield inferences about patient health, thereby challenging the traditional boundaries of medical expertise [35].

Despite these differences in outlook, GPs shared with many informaticians the view that technological advancements are unlikely to substitute physicians in the provision of empathic patient care [9,10]. Many AI experts argue that humanistic care will be improved with developments in machine learning: they suggest that by outsourcing precision clinical decision making to machine learning algorithms, physicians will be set to invest greater time attending to the needs of their patients. On the other hand, not all AI experts are sanguine about the future role of physicians, or indeed of people, in overseeing humanistic aspects of care [13]. For example, some researchers working in the field of affective computing point to findings that computers can already outperform humans when it comes to accurate discernment of facial expressions [38] and judgments about personality [39].

Comments also incorporated assumptions about patients' preferences. Some GPs' assumed that patients would prefer to receive medical care from physicians and raised concerns about the acceptability of future technology among health care users. Although there is limited research either to support or negate these claims, one recent US survey found significant differences between consumers and health providers with respect to their views about mobile technologies [19]. Boeldt and colleagues concluded that consumers were more likely to prefer, and to feel comfortable about, the use of technology for diagnoses than health providers [19].

Interestingly, few comments touched on safety issues: compared with GPs studied, AI researchers identify a wide range of serious concerns related to the design and use of machine learning algorithms. These include the risk of unfairness associated with "algorithmic biases," which can arise when demographic groups are underrepresented in training phases of machine learning [40]; the reliability and validity of medical algorithms [40]; problems of transparency in determining how algorithms reach decisions [41]; the adequate regulation of apps and mobile technologies [42]; and issues related to privacy and security with respect to patients' sensitive health information [43,44].

Finally, GPs' comments about levels of burnout raise important questions about how AI might mitigate unsafe workloads. Even if new machine learning technologies aim to augment human clinical decision making, it is unclear whether these tools will alleviate current levels of physician burnout [8,45]. Conceivably, in the short- and medium-term, if such clinical decision support systems are not suitably designed for human clinical workflow, they may result in "alert fatigue," undermining their utility [45,46]. Therefore, AI applications that aim to augment clinical judgments need to be designed with human and ergonomic factors in mind.

Strengths and Limitations

The study has a number of limitations. Comments were often brief, and because of the restrictions of online surveys, it was not possible to probe participants' responses to obtain a richer understanding of their views. Focus groups or in-depth qualitative interviews would have allowed finer-grained analysis of GPs' opinions. Given the often short, yet diverse range of opinions articulated in this survey, and in light of omissions of key questions about the potential impact of technology on primary care and the professional roles of physicians, further qualitative work is warranted.

Notwithstanding these limitations, this study provides a foundational exploration of physicians' views about the future of a medical specialism. The themes support the results of the earlier published quantitative survey while providing more nuanced perspectives on GPs' opinions about the future of primary care. We recommend that additional qualitative research focus on the attitudes of physicians working in other medical specialties as well as the views of nurse practitioners and physician assistants about how AI may encroach on both the future of patient care and the medical workforce. Finally, when it comes to technological advancements, we urge that greater attention be given to the attitudes and opinions of the users of primary care—patients.

Conclusions

This preliminary descriptive analysis provides insights into the ways in which GPs think about the impact of technological advancements on primary care. Perceived limitations and benefits of technology and social and ethical concerns about the future of general medicine have been elucidated. A variety of opinions were expressed reflecting some divergence in perspectives; overwhelmingly, however, participants were skeptical about the scope of technology to encroach on the traditional role of the GP (ie, in undertaking patient examinations; performing diagnoses; and providing personalized, empathic care) with only a few considering changes to current practices likely. Notably, this outlook stands in opposition to the predictions of biomedical informaticians and experts working in health care AI.

As we consider these findings, we cannot help but reflect on GPs' contrastive expectations that future technology will procure benefits in reducing administrative tasks. Such views raise questions about the equanimity of participants' forecasts. Similarly, some GPs were adamant that technology could not replace physicians in the delivery of empathic care. Although many AI experts appear to agree with this outlook, we suggest that this shared viewpoint does not yet future-proof the role of GPs in overseeing this task: conceivably, if informaticians' predictions are borne out, nurse practitioners, or a new occupation of medical empathizers, may emerge to undertake humanistic tasks. Relatedly, comments that greater investment

XSL•FO
RenderX

in primary care physicians could address workloads are challenged by findings of the World Health Organization, which claims that there will be a worldwide shortage of 18 million health care workers by 2030, over twice the current shortfall [47]. Increasing numbers of patients suffering from chronic illness and aging populations have therefore led many commentators to suggest that new strategies will be required to cope with growing national as well as global health care needs [9,48]. Nonetheless, we caution that the issue of physicians being overworked is unlikely to abate easily. As physicians and providers are currently overwhelmed, burnout itself may be a barrier in the implementation of new tools that are aimed at streamlining the medical care problem [45,49].

Although we cannot provide evidence of explanations for GPs' skepticism, it seems plausible that respondents' beliefs may reflect a level of disengagement from the literature on health care AI [50]. Our findings carry implications about the capacity of physicians to contribute in meaningful and objective ways to the many cutting-edge ethical and policy issues in relation to the advancement of AI in medicine [8,51,52]. Therefore, we conclude that our survey results raise important questions about the adequacy of medical curricula to equip future physicians for potential changes to clinical practice and, thereby, to lead and shape crucial debates about the future of patient care. Improvements in education, we suggest, may go some way to closing the rift between current AI health researchers and practitioners.

Acknowledgments

This work was conducted with support from Harvard Catalyst, The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR002541, and financial contributions from Harvard University and its affiliated academic health care centers); it was also supported by an Irish Research Council-Marie Skłodowska Curie Global Fellowship (CLNE/2017/226) and the University of Basel. Additional funding: CB was supported by a Fulbright Award, MHB was supported by T32DA016184. The study funders played no role in the study design, writing of the report, or the decision to submit the manuscript for publication. Researchers were independent of influence from study funders.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questions.

[PDF File (Adobe PDF File), 105KB - jmir_v21i3e12802_app1.pdf]

References

1. Wachter R. The Digital Doctor: Hope, Hype and at the Dawn of Medicines Computer Age. Pennsylvania: McGraw-Hill Education; 2015.
2. Susskind RE, Susskind D. The Future Of Professions. Oxford: Oxford University Press; 2017.
3. Coiera E. The fate of medicine in the time of AI. The Lancet 2018 Dec;392(10162):2331-2332. [doi: 10.1016/S0140-6736(18)31925-1]
4. Shah ND, Steyerberg EW, Kent DM. Big data and predictive analytics. J Am Med Assoc 2018 Jul 03;320(1):27. [doi: 10.1001/jama.2018.5602]
5. Kohane I, Drazen J, Campion E. A Glimpse of the Next 100 Years in Medicine. N Engl J Med 2012 Dec 27;367(26):2538-2539. [doi: 10.1056/NEJMe1213371]
6. Hinton G. Deep learning-a technology with the potential to transform health care. J Am Med Assoc 2018;320(11):1101-1102. [doi: 10.1001/jama.2018.11100] [Medline: 30178065]
7. Murdoch TB, Detsky AS. The inevitable application of big data to health care. J Am Med Assoc 2013 Apr 03;309(13):1351-1352. [doi: 10.1001/jama.2013.393]
8. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018;2(10):719.
9. Obermeyer Z, Emanuel EJ. Predicting the future-big data, machine learning, and clinical medicine. N Engl J Med 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]
10. Steinhubl S, Muse ED, Topol EJ. Can mobile health technologies transform health care? J Am Med Assoc 2013 Dec 11;310(22):2395-2396. [doi: 10.1001/jama.2013.281078] [Medline: 24158428]
11. Naylor CD. On the Prospects for a (deep) learning health care system. J Am Med Assoc 2018 Sep 18;320(11):1099. [doi: 10.1001/jama.2018.11103]
12. Steinhubl SR, Kim K, Ajayi T, Topol EJ. Virtual care for improved global health. The Lancet 2018;391(10991):419. [doi: 10.1016/S0140-6736(18)30109-0] [Medline: 29407024]
13. Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. J Am Med Assoc 2016 Feb 09;315(6):551. [doi: 10.1001/jama.2015.18421] [Medline: 26864406]

XSL•FO
RenderX

14. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. J Am Med Assoc 2018;319(1):19-20. [doi: 10.1001/jama.2017.19198] [Medline: 29261830]
15. Alpert J. Will physicians stop performing physical examinations? Am J Med 2017;130(7):759-760. [doi: 10.1016/j.amjmed.2017.03.013]
16. Rosenthal DI, Verghese A. Meaning and the nature of physicians' work. New Engl J Med 2016;10(375):1813-1815. [doi: 10.1056/NEJMp1609055]
17. Khorana A. Physician as typist. J Clin Oncol 2010 Aug 20;28(24):3899-3900. [doi: 10.1200/JCO.2010.29.4504] [Medline: 20547988]
18. Collado-Mesa F, Alvarez E, Arheart K. The role of artificial intelligence in diagnostic radiology: a survey at a single radiology residency training program. J Am Coll Radiol 2018;15(12):1753-1757. [doi: 10.1016/j.jacr.2017.12.021] [Medline: 29477289]
19. Boeldt DL, Wineinger NE, Waalen J, Gollamudi S, Grossberg A, Steinhubl SR, et al. How consumers and physicians view new medical technology: comparative survey. J Med Internet Res 2015;17(9):e215. [doi: 10.2196/jmir.4456] [Medline: 26369254]
20. Goldhahn J, Rampton V, Spinas GA. Could artificial intelligence make doctors obsolete? Br Med J 2018;363:k4563. [doi: 10.1136/bmj.k4563] [Medline: 30404897]
21. Pinto Dos Santos D, Giese D, Brodehl S, Chon S, Staab W, Kleinert R, et al. Medical students' attitude towards artificial intelligence: a multicentre survey. Eur Radiol 2018 Jul 06:- (forthcoming). [doi: 10.1007/s00330-018-5601-1] [Medline: 29980928]
22. Blease C, Bernstein MH, Gaab J, Kaptchuk TJ, Kossowsky J, Mandl KD, et al. Computerization and the future of primary care: a survey of general practitioners in the UK. PLoS One 2018;13(12):e0207418 [FREE Full text] [doi: 10.1371/journal.pone.0207418] [Medline: 30540791]
23. Doctors.net. Doctors.net.uk. URL: http://www.doctors.net.uk/about.html [accessed 2019-01-31] [WebCite Cache ID 75qXHXO0T]
24. Howick J, Bishop FL, Heneghan C, Wolstenholme J, Stevens S, Hobbs FD, et al. Placebo use in the United kingdom: results from a national survey of primary care practitioners. PLoS One 2013;8(3):e58247 [FREE Full text] [doi: 10.1371/journal.pone.0058247] [Medline: 23526969]
25. Turner PJ, Van den Bruel A, Jones CH, Plüddemann A, Heneghan C, Thompson MJ, et al. Point-of-care testing in UK primary care: a survey to establish clinical needs. Fam Pract 2016 Dec;33(4):388-394 [FREE Full text] [doi: 10.1093/fampra/cmw018] [Medline: 27048525]
26. Gonçalves-Bradley DC, Boylan AM, Koshiaris C, Vazquez Montes M, Ford GA, Lasserson DS. GPs' adherence to guidelines for structured assessments of stroke survivors in the community and care homes. Fam Pract 2015 Dec;32(6):659-663 [FREE Full text] [doi: 10.1093/fampra/cmv074] [Medline: 26424724]
27. General Medical Council. 2018. GMC data explorer URL: https://data.gmc-uk.org/gmcdata/home/ [accessed 2019-01-31] [WebCite Cache ID 75qV3hlkp]
28. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
29. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [Medline: 10940958]
30. Joffe H, Yardley L. Content and thematic analysis. In: Marks DF, Yardley L, editors. Research Methods for Clinical and Health Psychology. London: SAGE; 2004:56-58.
31. Mandl KD, Bourgeois FT. The evolution of patient diagnosis: from art to digital data-driven science. J Am Med Assoc 2017 Nov 21;318(19):1859-1860. [doi: 10.1001/jama.2017.15028] [Medline: 29075757]
32. Beam AL, Kohane IS. Translating artificial intelligence into clinical care. J Am Med Assoc 2016 Dec 13;316(22):2368-2369. [doi: 10.1001/jama.2016.17217] [Medline: 27898974]
33. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mHealth interventions for patients with diabetes: an overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]
34. Cui M, Wu X, Mao J, Wang X, Nie M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. PLoS One 2016;11(11):e0166718 [FREE Full text] [doi: 10.1371/journal.pone.0166718] [Medline: 27861583]
35. Banegas JR, Ruilope LM, de la Sierra A, Vinyoles E, Gorostidi M, de la Cruz JJ, et al. Relationship between clinic and ambulatory blood-pressure measurements and mortality. N Engl J Med 2018 Apr 19;378(16):1509-1520. [doi: 10.1056/NEJMoa1712231] [Medline: 29669232]
36. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. NPJ Digital Med 2018 Mar 14;1(1):5. [doi: 10.1038/s41746-017-0012-2]
37. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Dec;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]

XSL•FO
RenderX

38. Bartlett MS, Littlewort GS, Frank MG, Lee K. Automatic decoding of facial movements reveals deceptive pain expressions. Curr Biol 2014 Mar 31;24(7):738-743 [FREE Full text] [doi: 10.1016/j.cub.2014.02.009] [Medline: 24656830]
39. Youyou W, Kosinski M, Stillwell D. Computer-based personality judgments are more accurate than those made by humans. Proc Natl Acad Sci U S A 2015 Jan 27;112(4):1036-1040 [FREE Full text] [doi: 10.1073/pnas.1418680112] [Medline: 25583507]
40. Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, et al. Genetic misdiagnoses and the potential for health disparities. N Engl J Med 2016 Aug 18;375(7):655-665 [FREE Full text] [doi: 10.1056/NEJMsa1507092] [Medline: 27532831]
41. Price W. Regulating Black-Box Medicine. Mich Law Rev 2017;116(3):421-474. [Medline: 29240330]
42. Cohen IG, Graver HS. Cops, docs, and code: a dialogue between big data in health care and predictive policing. UCD Law Rev 2017;51:437 [FREE Full text]
43. Cohen I, Mello MM. HIPAA and protecting health information in the 21st Century. J Am Med Assoc 2018 Dec 17;320(3):231-232. [doi: 10.1001/jama.2018.5630] [Medline: 29800120]
44. Gostin L, Halabi SF, Wilson K. Health data and privacy in the digital era. J Am Med Assoc 2018 Jul 17;320(3):233-234. [doi: 10.1001/jama.2018.8374] [Medline: 29926092]
45. Yu KH, Kohane IS. Framing the challenges of artificial intelligence in medicine. BMJ Qual Saf 2018 Oct 05:1-4 (forthcoming). [doi: 10.1136/bmjqs-2018-008551] [Medline: 30291179]
46. Singh H, Spitzmueller C, Petersen NJ, Sawhney MK, Sittig DF. Information overload and missed test results in electronic health record-based settings. JAMA Intern Med 2013 Apr 22;173(8):702-704 [FREE Full text] [doi: 10.1001/2013.jamainternmed.61] [Medline: 23460235]
47. World Health Organization. World Health Organization. 2016. Draft global strategy on human resources for health: Workforce 2030 URL: http://who.int/gb/ebwah/pdf_files/WHA69/A69_38-en.pdf [accessed 2018-10-05] [WebCite Cache ID 73t79UuBn]
48. Darzi A, Evans T. The global shortage of health workers-an opportunity to transform care. Lancet 2016 Dec 26;388(10060):2576-2577. [doi: 10.1016/S0140-6736(16)32235-8] [Medline: 27894651]
49. Imo UO. Burnout and psychiatric morbidity among doctors in the UK: a systematic literature review of prevalence and associated factors. BJPsych Bull 2017 Aug;41(4):197-204 [FREE Full text] [doi: 10.1192/pb.bp.116.054247] [Medline: 28811913]
50. Obermeyer Z, Lee TH. Lost in thought-the limits of the human mind and the future of medicine. N Engl J Med 2017 Sep 28;377(13):1209-1211 [FREE Full text] [doi: 10.1056/NEJMp1705348] [Medline: 28953443]
51. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS Med 2018 Nov;15(11):e1002689 [FREE Full text] [doi: 10.1371/journal.pmed.1002689] [Medline: 30399149]
52. Fraser H, Coiera E, Wong D. Safety of patient-facing digital symptom checkers. Lancet 2018 Nov 24;392(10161):2263-2264. [doi: 10.1016/S0140-6736(18)32819-8] [Medline: 30413281]

Abbreviations

AI: artificial intelligence
GPs: general practitioners
mHealth: mobile health

Edited by T Kool; submitted 13.11.18; peer-reviewed by L Laranjo, D Boeldt, B Chaudhry; comments to author 14.12.18; revised version received 16.12.18; accepted 17.12.18; published 20.03.19.

Please cite as:
Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, DesRoches CM
Artificial Intelligence and the Future of Primary Care: Exploratory Qualitative Study of UK General Practitioners' Views
J Med Internet Res 2019;21(3):e12802
URL: http://www.jmir.org/2019/3/e12802/
doi:10.2196/12802
PMID:30892270

©Charlotte Blease, Ted J Kaptchuk, Michael H Bernstein, Kenneth D Mandl, John D Halamka, Catherine M DesRoches. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of

XSL•FO
RenderX

Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Physician Confidence in Artificial Intelligence: An Online Mobile Survey

Songhee Oh[1], MD; Jae Heon Kim[2], MD; Sung-Woo Choi[3], MD; Hee Jeong Lee[1], MD; Jungrak Hong[4], MD; Soon Hyo Kwon[1], MD

[1]Division of Nephrology, Department of Internal Medicine, Soonchunhyang University Hospital, Seoul, Republic of Korea

[2]Department of Urology, Soonchunhyang University Hospital, Seoul, Republic of Korea

[3]Department of Orthopedic Surgery, Soonchunhyang University Seoul Hospital, Seoul, Republic of Korea

[4]Department of Internal Medicine, New York Medical College, New York Health Hospital, New York, NY, United States

Corresponding Author:
Soon Hyo Kwon, MD
Division of Nephrology
Department of Internal Medicine
Soonchunhyang University Hospital
Daesagwanro 59 Youngsangu
Seoul, 04401
Republic of Korea
Phone: 82 01034413147
Email: ksoonhyo@schmc.ac.kr

Abstract

Background: It is expected that artificial intelligence (AI) will be used extensively in the medical field in the future.

Objective: The purpose of this study is to investigate the awareness of AI among Korean doctors and to assess physicians' attitudes toward the medical application of AI.

Methods: We conducted an online survey composed of 11 closed-ended questions using Google Forms. The survey consisted of questions regarding the recognition of and attitudes toward AI, the development direction of AI in medicine, and the possible risks of using AI in the medical field.

Results: A total of 669 participants completed the survey. Only 40 (5.9%) answered that they had good familiarity with AI. However, most participants considered AI useful in the medical field (558/669, 83.4% agreement). The advantage of using AI was seen as the ability to analyze vast amounts of high-quality, clinically relevant data in real time. Respondents agreed that the area of medicine in which AI would be most useful is disease diagnosis (558/669, 83.4% agreement). One possible problem cited by the participants was that AI would not be able to assist in unexpected situations owing to inadequate information (196/669, 29.3%). Less than half of the participants(294/669, 43.9%) agreed that AI is diagnostically superior to human doctors. Only 237 (35.4%) answered that they agreed that AI could replace them in their jobs.

Conclusions: This study suggests that Korean doctors and medical students have favorable attitudes toward AI in the medical field. The majority of physicians surveyed believed that AI will not replace their roles in the future.

(J Med Internet Res 2019;21(3):e12422) doi:10.2196/12422

KEYWORDS

artificial intelligence; AI; awareness; physicians

Introduction

Research into and usage of artificial intelligence (AI) has been gaining popularity in the field of computer science [1-3]. Recently, various kinds of AI programs have been developed based on "big data" collected through the Internet of Things. AI programs have been widely used in the manufacturing sector, the information-communications industry [4], and the medical field [5-7]. The development and utilization of AI programs in the medical field are currently entering the stage of commercialization [8,9]. AI is defined as the ability of computer systems to perform tasks that would usually require human levels of intelligence. A subfield of AI is machine learning, which can be used to teach a computer to analyze a vast amount

XSL•FO
RenderX

of data in a rapid, accurate, and efficient manner through the use of complex computing and statistical algorithms [10,11].

In the past, it was thought that AI would replace doctors in many areas [12-15]. However, this has not occurred. Recent scientific advances have been rapid, suggesting that this might be a possibility. IBM's Watson was developed from a huge database of published literature and millions of medical records [16,17]. Based on this, it can assist in the establishment of precise diagnoses and proper treatment plans [16,17]. Furthermore, Watson provides advice on the best treatments for cancer and conducts genome analyses [18]. Similarly, Google's DeepMind software is being used to test the feasibility of the automated grading of digital fundus photographs using optical coherence tomography [19]. Recently, AI has been used to predict genetic variations in low-grade gliomas [20], identify genetic phenotypes in small cell lung carcinoma [21], decrease false-positive rates in screening mammography computer-aided detection [22], improve pathologic mediastinal lymph node detection [23], and automatically perform bone age assessment [24]. These examples demonstrate the influence of AI in medicine. The application of AI will be further extended to other areas in the future, leading to fundamental changes in the role of physicians and the way they practice medicine [25].

Korea is regarded as a technologically advanced country. Among people aged 18 to 24 years in the Republic of Korea, mobile phone penetration is 97.7%. Of the approximately 19 million households in Korea, 99.2% have internet access via an optical local area network, digital subscriber line (xDSL), cable modem, mobile device, or other media compared to approximately 75% in the United States [26].

There are differing perspectives on the future of AI. A pessimistic view of AI is that AI will replace humans in many industries. Optimistic views also exist in which humans will have more opportunities to benefit from clinical advances in the future with AI support [27]. Recently the AlphaGo AI program defeated a human Go professional, which shocked Korean society and provoked controversy in Korea [28,29].

Recent news reports revealed that Korean patients would follow AI advice over a doctor's advice about their cancer treatment [30]. However, there is no research on the opinions and attitudes of Korean physicians toward the application of AI programs in the medical field. Current medical students and young physicians will be affected by AI before they retire. Therefore, physicians need to be prepared for these changes to use AI effectively as a tool.

The purpose of this study is to investigate the awareness of AI programs among Korean medical doctors and to provide basic information about physicians' awareness of and reactions to the introduction of AI in the future.

Methods

Participants

This study was approved by an institutional review board at Soonchunhyang Medical College Hospital Seoul (no. 2017-05-014). Using Google Forms, we surveyed medical students, doctors who graduated from Soonchunhyang Medical College, and doctors at hospitals affiliated with Soonchunhyang University. The survey was administered online through a mobile phone invitation. Demographic and professional information on the medical students and doctors were obtained. Each participant was sent a unique link to the online survey. Participants were informed about the goal of the survey (medical research) in the preface of the questionnaire. By voluntarily participating in the survey after being given adequate information on its purpose, informed consent was implied. We confirm that participation was voluntary; participants could not be identified from the material presented and no plausible harm to participating individuals could arise from the study. Responses were made on a single Web page with one "submit" button that only allowed submissions through these unique links, thus making noninvited responses extremely unlikely.

Measurement Instruments

Survey

In May 2017, our online survey, consisting of 11 closed-ended questions, was conducted (Textbox 1 and Multimedia Appendix 1). Survey content validity was reviewed by study researchers (n=5) and a panel of physicians (n=5) who were accepting patients at their sites. Following this, pilot testing was performed by medical college students (n=20) and physicians (n=80) who did not participate in developing the survey. Our survey was in accordance with the Checklist for Reporting Results of Internet E-survey (CHERRIES) [31]. The contents of the survey consisted of a questionnaire regarding the recognition of and attitudes toward AI, the direction of AI development in medicine, and the possible risks of using AI in the medical field. Three internal medicine physicians consulted the latest journals on AI and composed the questionnaire [5,6,9,16-18,32-34]. We sent 3000 doctors and medical students Web links to the questionnaire. These potential participants were almost entirely alumni of Soonchunhyang Medical College or were employed at hospitals affiliated with Soonchunhyang University.

The answers to five questions (Q1-Q5) were assessed using a five-point ordinal Likert scale (1=strongly disagree to 5=strongly agree). For three additional questions (Q6, Q8, Q10), 50 sample respondents were given to the questions in an open-ended format, and the five most commonly given answers were selected to be the five possible answer choices for survey participants.

XSL•FO
RenderX

Textbox 1. Questions asked in the online survey regarding artificial intelligence (AI) in the medical field. The answers to questions 1-5 were assessed with a five-point Likert scale (1=strongly disagree to 5=strongly agree).

Attitudes

Q1. Do you agree that you have good familiarity with artificial intelligence?

Q2. Do you agree that artificial intelligence has useful applications in the medical field?

Q3. Do you agree that the diagnostic ability of AI is superior to the clinical experience of a human doctor?

Q4. Do you agree that artificial intelligence could replace your job?

Q5. Do you agree that you would always use AI when making medical decisions in the future?

Q6. What are the advantages of using artificial intelligence?

- AI can speed up processes in health care
- AI can help reduce medical errors.
- AI can deliver vast amounts of clinically relevant high-quality data in real time
- AI has no space-time constraint
- AI has no emotional exhaustion nor physical limitation

Q7. If your medical judgment and an artificial intelligence's judgments differ, which will you follow?

- Doctor's opinion
- Artificial intelligence's opinion
- Patients' choice

Expected Applications in Medicine

Q8. In which field of medicine do you think artificial intelligence will be most useful?

- Making a diagnosis
- Making treatment decisions
- Direct treatment (including surgery)
- Biopharmaceutical research and development
- Providing medical assistance in underserved areas
- Development of social insurance program

Q9. Which sector of health care do you think will be the first to commercialize artificial intelligence?

- Public primary care such as public health centers
- Primary care in private clinics
- Specialized clinics (spine, knee, obstetrics and gynecology, etc)
- University hospitals

Possible Risks

Q10. What are you concerned about application of AI in medicine? It cannot be used to provide opinions in unpredicted situations due to inadequate

- Information
- It is not flexible enough to be applied to every patient
- It is difficult to apply to controversial subjects
- The low ability to sympathize and consider the emotional well-being of the patient
- It is developed by a specialist with little clinical experience in medical practice

Q11. Who do you think will be liable for legal problems caused by artificial intelligence?

- Doctor in charge
- Company that created the artificial intelligence
- Patient who consented to follow artificial intelligence's input

XSL•FO
RenderX

Questionnaire

1. Attitudes: the first part of the survey asked about the physician's attitude toward the medical application of AI. The questions and possible answer choices (if applicable) are detailed in Multimedia Appendix 1. A total of seven closed-ended questions were included (Q1-Q7).
2. Expected applications in medicine: medical students and physicians were asked about the medical fields in which AI could be applied. The questions and possible answer choices are detailed in Multimedia Appendix 1. A total of two closed-ended questions were included (Q8 and Q9).
3. Possible risks: medical students and physicians were asked which problems they were concerned about regarding the application of AI in medicine. It is not clear who is liable when there are adverse clinical outcomes between humans and AI; therefore, we included a question about liability for AI decisions in medicine. The questions and possible answer choices for each question are detailed in Multimedia Appendix 1. A total of two closed-ended questions were included (Q10 and Q11).

Subgroup Analyses: Specialty, Working Status, and Medical Experience

We investigated whether attitudes differed regarding the medical applications of AI depending on the respondent's specialty degree of medical experience, working status, and work location.

For this study, the categories for department were medical student, physician, surgeon, or other. The categories for working status were medical student, training physician (intern, resident, or clinical fellow), university professor, or nonuniversity physician. The categories for amount of medical experience were the number of years licensed from medical school graduation: less than 10 years, between 10 and 20 years, or more than 20 years. The categories for working location were in and around Seoul, large cities outside of Seoul, small cities outside of Seoul, or small towns or rural areas.

Statistical Analysis

Basic statistics (mean and standard deviation or total number and percent) were computed for all covariates. In the subgroup analysis, Kruskal-Wallis tests served for evaluating the effect of gender factors of questionnaire items. The differences in the questionnaire responses according to working state, location, licensed years, and medical specialty were analyzed using the Mann-Whitney test. For all tests, the level of significance was set at $P \leq .05$.

Results

Participants

During the study period, 669 participants, out of approximately 3000, completed the survey (22.3% rate of return). There were 121 medical students, 162 training physicians, and 386 physicians. Among these participants, 22.4% (150/669) were younger than 30 years, and 22.1% (148/669) were female. The demographic and professional characteristics of the participants are listed in Table 1.

Questionnaire

The results of the questionnaire are summarized in Table 2.

Responses to the Questionnaire

Attitudes

Generally, familiarity with AI was low. Only 40 of 669 respondents (6.0%) answered that they had good familiarity with AI (Figure 1). Many participants considered AI useful in the medical field (73.4%, 558/669). The respondents agreed that the advantages of using AI were its ability to quickly obtain vast amounts of clinically relevant, high-quality data in real time (62.3%, 417/669), speed up processes in health care (19.1%, 128/669), and decrease the number of medical errors (9.6%, 64/669) (Figure 1). However, fewer than half of the participants agreed that "AI is superior to a doctor's experience" (44%, 294/669), "AI could replace a doctor" (35.4%, 237/669), or "AI would be used whenever medical decisions need to be made" (42%, 281/669) (Figures 1 and 2). If there were differences between an AI's decision and a doctor's opinion regarding a medical decision, 79% (528/669) of participants would follow the doctor's opinion. The results from the attitudes section of the questionnaire are summarized in Table 2.

Expected Application in Medicine

Respondents felt the areas in medicine where AI would be most useful in the future were reaching a diagnosis (83.4%, 558/669) and forming a treatment plan (53.8%, 360/669). Fewer than 10% felt it would be useful in providing medical assistance in underserved areas (9.6%, 64/669), treating patients independently (eg, performing surgery, 9.0%, 60/669), or developing medical insurance guidelines (6.1%, 41/669). Additionally, most participants (66.2%, 443/669) thought that AI would be first commercialized at a university hospital. The results from the expected fields section of the questionnaire are summarized in Table 2.

Possible Risks

According to the respondents, the possible problems with AI are that AI would be unable to provide an opinion in an unpredicted situation owing to inadequate information (29.3%, 196/669) and that it would not be applied to every patient (34.1%, 228/669). In the case of a medical problem caused by AI, respondents felt responsibility should lie with the doctors (49.3%, 330/669), patients who consented to the use of AI (31.2%, 209/669), and the company that created the AI (19.4%, 130/669). The results from the possible risks section of the questionnaire are summarized in Table 2.

Subgroup Analysis

The results of subgroup analysis according to the demographic characteristics of participants are summarized in Table 3.

XSL•FO
RenderX

Table 1. Demographic characteristics of participants surveyed about physicians and artificial intelligence (N=669).

Characteristics	n (%)
Age (years)	
<30	150 (22.4)
31-40	197 (29.4)
41-50	159 (23.8)
51-60	137 (20.5)
61-70	18 (2.7)
≥71	8 (1.2)
Gender	
Male	514 (76.8)
Female	148 (22.1)
No response	7 (1.0)
Working status	
Medical student	121 (18.1)
Training physicians (intern, residents, fellows)	112 (16.7)
University professors	90 (13.5)
Nonuniversity physicians	346 (51.7)
Licensed years	
Medical student	121 (18.1)
<10 years	177 (26.5)
10-20 years	170 (25.4)
>40 years	201 (30.0)
Medical specialty	
Medical student	121 (18.1)
Medical department	284 (42.5)
Surgical department	204 (30.5)
Extra department	60 (9.0)
Hospital status	
Medical school	121 (18.1)
University hospital	162 (24.2)
District general hospital	67 (10.0)
Solo practice	217 (32.4)
Group practice	30 (4.5)
Long-term care hospital	24 (3.6)
Community health center or military hospital	29 (4.3)
Others	19 (2.8)
Location of the clinics	
Seoul (Capital city)	278 (41.6)
Seoul Metropolitan Area (Capital area)	162 (24.2)
Regional Metropolitan City	44 (6.6)
Cities	128 (19.1)
Rural	57 (8.5)

XSL•FO
RenderX

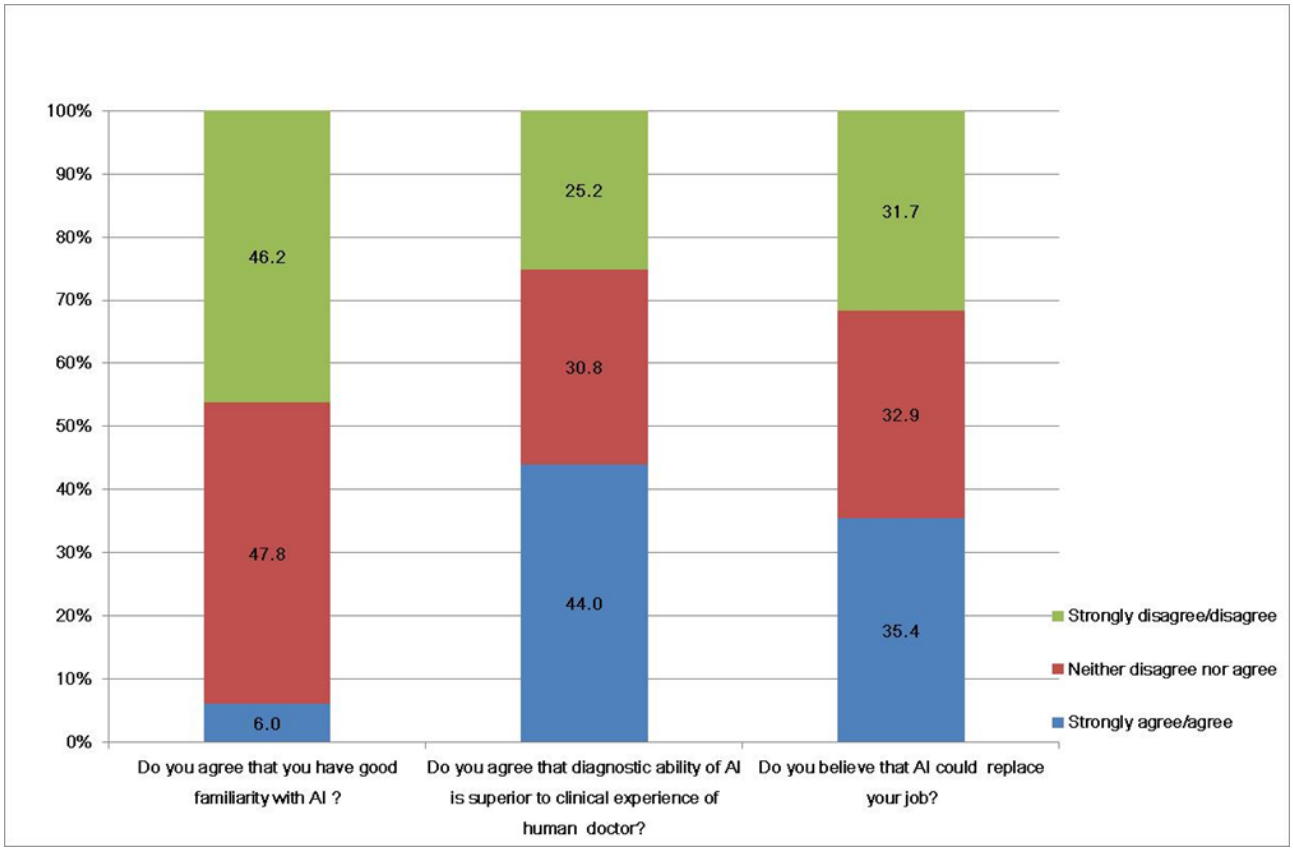
Table 2. Participant's attitudes on artificial intelligence (AI), the expected applications in medicine, and possible risks (N=669).

Question	n (%)
Attitudes	
Q1. Do you agree that you have good familiarity with artificial intelligence?	
Strongly agree/agree	40 (6.0)
Neither disagree nor agree	320 (47.8)
Strongly disagree/disagree	309 (46.2)
Q2. Do you agree that AI has useful applications in the medical field?	
Strongly agree/agree	558 (73.4)
Neither disagree nor agree	97 (14.5)
Strongly disagree/disagree	14 (2.1)
Q3. Do you agree that the diagnostic ability of AI is superior to the clinical experience of human doctors?	
Strongly agree/agree	294 (44.0)
Neither disagree nor agree	206 (30.8)
Strongly disagree/disagree	169 (25.2)
Q4. Do you agree that AI could replace you in your job?	
Strongly agree/agree	237 (35.4)
Neither disagree nor agree	220 (32.9)
Strongly disagree/disagree	212 (31.7)
Q5. Do you agree that you will always use AI to make medical judgments in the future?	
Strongly agree/agree (=always/often)	281 (42.0)
Neither disagree nor agree (=occasionally)	87 (13.0)
Strongly disagree/disagree (=never/seldom)	301 (45.0)
Q6. What are the advantages of using AI?	
AI can speed up the process in health care	128 (19.1)
AI can help in reducing the number of medical errors	64 (9.6)
AI can deliver clinically relevant, vast amounts of high-quality data in real time	417 (62.3)
AI has no space-time constraint	12 (1.8)
AI has no emotional exhaustion or physical limitation	3 (0.4)
Q7. If your judgment and AI judgments differ, which will you follow?	
Doctor's opinion	528 (78.9)
Artificial intelligence's opinion	110 (16.4)
Patients' choice	31 (4.6)
Expected fields	
Q8. In which field of medicine do you think AI will be most useful?	
Making diagnoses	558 (83.4)
Making the decision for treatment	360 (53.8)
Direct treatment (including surgery)	60 (9.0)
Biopharmaceutical research and development	84 (12.6)
Provide medical assistance in underserved areas	64 (9.6)
Development of social insurance program	41 (6.1)
Q9. Which sector of health care do you think will be the first to commercialize AI?	
Public primary care such as public health centers	98 (14.6)
Primary care in private clinics	31 (4.6)

XSL•FO
RenderX

Question	n (%)
Specialized clinics (spine, knee, obstetrics, and gynecology, etc)	97 (14.5)
University hospitals	443 (66.2)
Possible risks	
Q10. Which problems are you concerned about regarding the application of AI in medicine?	
It cannot be used to provide opinions in unexpected situations owing to inadequate stored information	196 (29.3)
It is not flexible enough to be applied to every patient	228 (34.1)
It is difficult to apply to controversial subjects	38 (5.7)
Low ability to sympathize and consider the emotional well-being of the patient	179 (26.8)
It was developed by a specialist with little clinical experience in medical practice	19 (2.8)
Q11. Who do you think will be responsible for medical problems caused by AI?	
Doctor in charge	330 (49.3)
Company that created the artificial intelligence	130 (19.4)
Patients who agreed to follow artificial intelligence's input	209 (31.2)

Figure 1. Major results of the questionnaire. AI: artificial intelligence.



Specialty

There was no significant difference in attitude toward AI according to the department of the respondent for any of the questions (Q1-Q5).

Working Status

There were no significant differences according to working status for three questions (Q2, Q4, Q5). There were significant differences for two questions (Q1, Q3). For the question about the recognition of AI (Q1), the outcomes for training physicians are summarized in Table 4.

Amount of Medical Experience

There was no statistical difference according to the degree of medical experience for three questions (Q2, Q4, Q5). There were significant differences between two questions (Q1, Q3). For the question about the recognition of AI (Q1), the outcomes for physicians licensed for less than than 10 years, physicians licensed between 10 and 20 years, and physicians licensed more than 20 years were higher than for medical students. For the

XSL•FO
RenderX

question about the superiority of AI in diagnostic ability (Q3), the outcomes of medical students and physicians licensed for fewer than 10 years were higher than for physicians licensed between 10 and 20 years and physicians licensed more than 20 years. The results of the subgroup analysis according to the amount of medical experience and age are summarized in Table 4. Age showed a similar result as the analysis according to license year.

Working Location

There was no significant difference in attitudes toward AI according to working location.

Figure 2. Responses about the advantage of artificial intelligence (AI) in medicine.

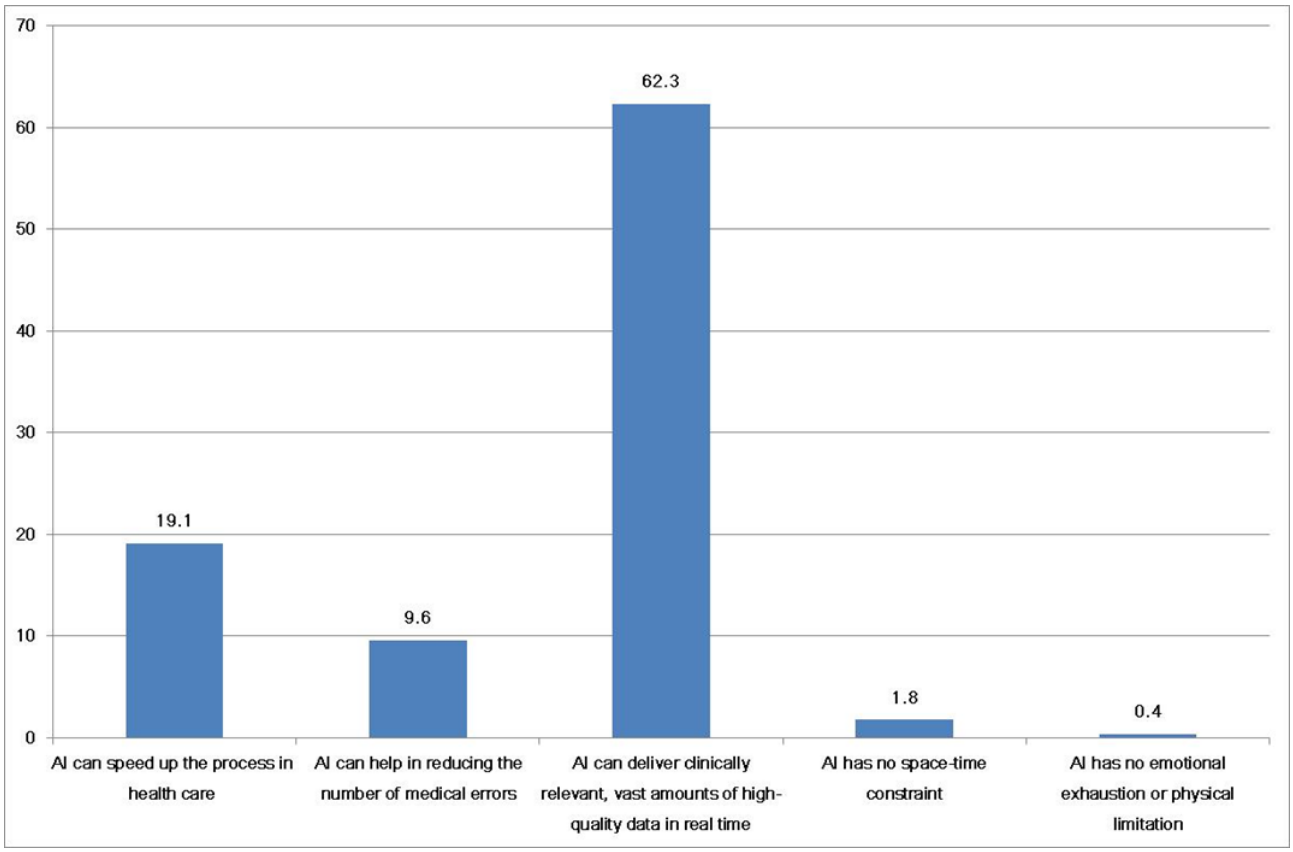


Table 3. Subgroup analysis according to the demographic characteristics of participants.

Question	*P* value[a]				
	Department	Working status	License year	Age	Location
Q1. Familiarity of AI[b]	.06	<.001	<.001	<.001	.54
Q2. Usefulness of AI	.07	.11	.24	.24	.10
Q3. Diagnostic ability of AI	.38	.001	<.001	<.001	.07
Q4. Replacement human job (doctor)	.46	.19	.35	.32	.52
Q5. Frequency of using AI	.92	.95	.92	.43	.17

[a]*P* values for categorical variables are based on Kruskal-Wallis tests.

[b]AI: artificial intelligence.

XSL•FO
RenderX

Table 4. Subgroup analysis according to working status.

Subgroup and question	Median (IQR[a])	*P* value[b]	Post hoc
Working status			
Q1. Familiarity with AI[c]		<.001	A<B, C, D
A. Students	2 (2-3)		
B. Training physician	3 (2-3)		
C. Professor	3 (2-3)		
D. Clinical physicians	3 (2-3)		
Q3. Diagnostic ability of AI		.001	A=B>C=D
A. Students	4 (3-4)		
B. Training physician	4 (3-4)		
C. Professor	3 (2-3)		
D. Clinical physicians	3 (2-4)		
License year			
Q1. Familiarity with AI		<.001	A<B=C=D
A. Students	2 (2-3)		
B. <10 years	3 (2-3)		
C. 10-20 years	3 (2-3)		
D. >20 years	3 (2-3)		
Q3. Diagnostic ability of AI		<.001	A=B>C=D
A. Students	4 (3-4)		
B. <10 years	4 (3-4)		
C. 10-20 years	3 (2-4)		
D. >20 years	3 (2-4)		
Age			
Q1. Familiarity with AI		<.001	A<B=C=D=E
A. 20-29 years	2 (2-3)		
B. 30-39 years	3 (2-3)		
C. 40-49 years	3 (2-3)		
D. 50-59 years	3 (2-3)		
E. >60 years	3 (2-3)		
Q3. Diagnostic ability of AI		<.001	A=B>C=D=E
A. 20-29 years	3 (3-4)		
B. 30-39 years	4 (3-4)		
C. 40-49 years	3 (2-4)		
D. 50-59 years	3 (2-4)		
E. >60 years	3 (2-4)		

[a]IQR: interquartile range.

[b]*P* values for categorical variables are based on Mann-Whitney tests.

[c]AI: artificial intelligence.

XSL•FO
RenderX

Discussion

Principal Results and Comparison With Prior Work

To the best of our knowledge, this study is the first survey of the attitudes of physicians toward AI. The results of this survey suggest that the recognition of AI by medical students and doctors is low. However, they regarded AI to be useful in the medical field. Physicians and medical students felt that AI would be most useful for reaching a diagnosis and formulating a treatment in the future. The majority of Korean doctors do not believe that AI will replace them.

Precision medicine is "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person" [35]. This approach allows doctors to choose treatment and prevention strategies for their patients. It requires significant computing power and algorithms that can learn by themselves at an unprecedented rate. Therefore, there is no precision medicine without AI. In our study, most physicians expected that AI would be helpful with diagnoses and in planning treatment by providing the latest clinically relevant data.

We asked the participants about the diagnostic superiority of AI compared to that of doctors. Fewer than half of the participants agreed that AI would be diagnostically superior. In the subgroup analysis, doctors in academic positions and office clinicians who had more clinical experience were less likely than medical students and training physicians to agree that AI is diagnostically superior. Additionally, experienced clinicians (licensed for more than 10 years) were less likely to agree that AI has superior diagnostic ability. Our questions were about general clinical practice. Although pathologists and radiologists were among the respondents, there were relatively few (27/669, 4.0%).

In contrast to our study, recent studies have shown that image recognition technology might make predictions or recognize diseases as effectively as or even better than physicians [17,36]. Liu and colleagues [36] from Google used an AI technique called convolutional neural network machine learning and demonstrated that AI achieves image-level area under the curve scores greater than 97% on both the Camelyon16 test set (metastasis detection of lymph nodes) and an independent set of 110 slides compared to a human pathologist, who achieved 73.2% sensitivity. Metastasis detection is currently performed by pathologists when reviewing large expanses of biological tissue. This process is labor intensive and error prone. However, AI machine learning saves time and is less likely to make errors [37]. In the case of radiology and pathology, some believe that AI will replace doctors based on diagnostic superiority [9,25,38]. Furthermore, AI could be able to extract fine information about tissues invisible to the human eye and process these data quickly and accurately [39,40].

Generally, AI has been used in imaging and pathology and is considered favorably in these fields [41-45]. Pathology and radiology have a common destiny as "informational specialists" with regard to images and pathology [38]. However, we did not investigate the reasons for their choices. They might believe that technical progress in the field of AI will not reach the level of human intelligence. It is also possible that Korean physicians have not examined the recent data on AI in the medical field.

In our study, 35.4% of participants agreed that doctors will be replaced by AI. This is not consistent with previous studies about AI. A 2017 survey by the Pew Research Center conducted with 4135 participants found that the public is roughly twice as likely to express worry (72%) than enthusiasm (33%) about a future in which robots and computers are capable of doing many human jobs [46]. Unlike other occupations, doctors felt that there would be difficulties in replacing doctors. Krittanawong [40] argued that AI cannot replace doctors yet at the bedside, given its limitations. First, AI cannot engage in high-level conversation with patients to gain their trust, reassure them, or express empathy [47]. These are all important parts of the doctor-patient relationship. Second, although AI sensors may glean valuable information to help with diagnosis, physicians will still be needed for interpretation in ambiguous situations to integrate medical histories, conduct physical exams, and facilitate further discussion [40]. It is possible that many Korean doctors believe this intuitively.

Skepticism can arise when applying AI to medical care. Regulations and principles of AI application need to be defined. AI can provoke ethical and legal problems in medicine. A regulatory authority should control AI algorithms for public safety. This issue will require debate from a social perspective.

Our survey response rate was 22.3%, which seems to be a lower response rate. However, previous studies demonstrated that electronic modalities often have lower response rates than paper mailed surveys [48-50]. Internet-based surveys demonstrated a lower response rate (45%) than the mail questionnaires (58%) (absolute difference 13%, 95% confidence interval 4%-22%, $P<.01$) [48]. A Cochrane review of randomized controlled trials identified numerous methods to increase response rates for both postal and electronic surveys [51]. With reference to this study, we made the questionnaire short, used a simple header, and gave a deadline. However, there were no monetary incentives, one of the major factors that increase the response rate. Plus, we could not provide a prenotification nor send reminder messages or have follow-up contact because we could not distinguish between those who already responded and the nonresponders.

Limitations

Some limitations of our study should be noted. First, we did not ask background questions concerning how much the individual participants technically understood AI. Each participant may have had different conceptualizations of AI. Second, there is the possibility of selection bias. Participants may have been more motivated and might have expressed more positive attitudes compared to nonparticipants. Because the data were self-reported, a bias owing to social desirability cannot be excluded. In addition, the selected participants may not have been a good representation of Korean doctors in general. However, our study did include various ages and clinical backgrounds. Third, the questionnaire about AI was created by doctors rather than AI experts.

XSL•FO
RenderX

Conclusion

This study found that physicians felt the application of AI to medicine would be useful. Physicians felt that the areas in medicine where AI would be most useful were diagnosis and treatment planning. However, more than half of the physicians did not believe AI would replace their role as health care providers. From a diagnostic point of view, doctors who had more experience favored a physician's experience over AI. Follow-up surveys and multinational studies should be conducted to further explore these issues.

Acknowledgments

This study was supported by a Soonchunhyang University research grant (20180005).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Online survey form.

[PDF File (Adobe PDF File), 250KB - jmir_v21i3e12422_app1.pdf]

References

1. Hastie TT, Friedman J, Tibshirani R. The Elements of Statistical Learning. New York: Springer Science & Business Media; 2009.
2. Abu-Mostafa Y, Magdon-Ismail M, Lin HT. Learning from Data. Learning From Data: AMLbook.com; 2012.
3. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med 2001 Aug;23(1):89-109. [Medline: 11470218]
4. Pyo S, Lee J, Cha M, Jang H. Predictability of machine learning techniques to forecast the trends of market index prices: hypothesis testing for the Korean stock markets. PLoS One 2017 Nov;12(11):e0188107 [FREE Full text] [doi: 10.1371/journal.pone.0188107] [Medline: 29136004]
5. Miller DD, Brown EW. Artificial intelligence in medical practice: the question to the answer? Am J Med 2018 Feb;131(2):129-133. [doi: 10.1016/j.amjmed.2017.10.035] [Medline: 29126825]
6. Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, et al. The coming of age of artificial intelligence in medicine. Artif Intell Med 2009 May;46(1):5-17 [FREE Full text] [doi: 10.1016/j.artmed.2008.07.017] [Medline: 18790621]
7. Ohno-Machado L. Research on machine learning issues in biomedical informatics modeling. J Biomed Inform 2004 Aug;37(4):221-223 [FREE Full text] [doi: 10.1016/j.jbi.2004.07.004] [Medline: 15465475]
8. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. Crit Care Med 2018 Apr;46(4):547-553. [doi: 10.1097/CCM.0000000000002936] [Medline: 29286945]
9. Dreyer KJ, Geis JR. When machines think: radiology's next frontier. Radiology 2017 Dec;285(3):713-718. [doi: 10.1148/radiol.2017171183] [Medline: 29155639]
10. Alsharqi M, Woodward W, Mumith A, Markham D, Upton R, Leeson P. Artificial intelligence and echocardiography. Echo Res Pract 2018 Dec 05:30400053. [doi: 10.1530/ERP-18-0056] [Medline: 30400053]
11. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? Heart 2018 Jan 19;104(14):1156-1164. [doi: 10.1136/heartjnl-2017-311198]
12. Lusted LB. Logical analysis in roentgen diagnosis. Radiology 1960 Feb;74:178-193. [doi: 10.1148/74.2.178] [Medline: 14419034]
13. Schwartz WB. Medicine and the computer. N Engl J Med 1970 Dec 03;283(23):1257-1264. [doi: 10.1056/nejm197012032832305]
14. Chockley K, Emanuel E. The end of radiology? Three threats to the future practice of radiology. J Am Coll Radiol 2016 Dec;13(12 Pt A):1415-1420. [doi: 10.1016/j.jacr.2016.07.010] [Medline: 27652572]
15. Jha S. Medscape. 2016 May 12. Will computer replace radiologist? URL: https://login.medscape.com/login/sso/getlogin?urlCache=aHR0cHM6Ly93d3cubWVkc2NhcGUuY29tL3ZpZXdhcnRpY2xlLzg2MzEyNw==&ac=401 [accessed 2019-02-11] [WebCite Cache ID 766qprzqW]
16. Bakkar N, Kovalik T, Lorenzini I, Spangler S, Lacoste A, Sponaugle K, et al. Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis. Acta Neuropathol 2018 Feb;135(2):227-247 [FREE Full text] [doi: 10.1007/s00401-017-1785-8] [Medline: 29134320]

XSL•FO
RenderX

17. Trivedi H, Mesterhazy J, Laguna B, Vu T, Sohn JH. Automatic determination of the need for intravenous contrast in musculoskeletal MRI examinations using IBM Watson's natural language processing algorithm. J Digit Imaging 2018 Apr;31(2):245-251. [doi: 10.1007/s10278-017-0021-3] [Medline: 28924815]
18. Malin JL. Envisioning Watson as a rapid-learning system for oncology. J Oncol Pract 2013 May;9(3):155-157 [FREE Full text] [doi: 10.1200/JOP.2013.001021] [Medline: 23942497]
19. De Fauw FJ, Keane P, Tomasev N, Visentin D, van den Driessche G, Johnson M, et al. Automated analysis of retinal imaging using machine learning techniques for computer vision. F1000Res 2016 Jun;5:1573 [FREE Full text] [doi: 10.12688/f1000research.8996.2] [Medline: 27830057]
20. Korfiatis P, Kline TL, Lachance DH, Parney IF, Buckner JC, Erickson BJ. Residual deep convolutional neural network predicts MGMT methylation status. J Digit Imaging 2017 Oct;30(5):622-628 [FREE Full text] [doi: 10.1007/s10278-017-0009-z] [Medline: 28785873]
21. Rios VE, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. Cancer Res 2017 Dec 15;77(14):3922-3930 [FREE Full text] [doi: 10.1158/0008-5472.CAN-17-0122] [Medline: 28566328]
22. Mordang J, Gubern-Mérida A, Bria A, Tortorella F, den Heeten HG, Karssemeijer N. Improving computer-aided detection assistance in breast cancer screening by removal of obviously false-positive findings. Med Phys 2017 Apr;44(4):1390-1401. [doi: 10.1002/mp.12152] [Medline: 28182277]
23. Liu J, Hoffman J, Zhao J, Yao J, Lu L, Kim L, et al. Mediastinal lymph node detection and station mapping on chest CT using spatial priors and random forest. Med Phys 2016 Jul;43(7):4362 [FREE Full text] [doi: 10.1118/1.4954009] [Medline: 27370151]
24. Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, et al. Fully automated deep learning system for bone age assessment. J Digit Imaging 2017 Aug;30(4):427-441 [FREE Full text] [doi: 10.1007/s10278-017-9955-8] [Medline: 28275919]
25. Diprose W, Buist N. Artificial intelligence in medicine: humans need not apply? N Z Med J 2016 May 06;129(1434):73-76. [Medline: 27349266]
26. Ramirez E. Forbes. 2017 Jan 31. Nearly 100% of households in South Korea now have internet access, thanks to seniors URL: https://www.forbes.com/sites/elaineramirez/2017/01/31/nearly-100-of-households-in-south-korea-now-have-internet-access-thanks-to-seniors/#1c05a9c55572 [accessed 2019-02-06] [WebCite Cache ID 75yHsuhVW]
27. Johnston SC. Anticipating and training the physician of the future: the importance of caring in an age of artificial intelligence. Acad Med 2018 Aug;93(8):1105-1106. [doi: 10.1097/ACM.0000000000002175] [Medline: 29443717]
28. Metz C. Wired. 2016 Mar 11. The sadness and beauty of watching Google's AI play Go URL: https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/ [accessed 2019-02-06] [WebCite Cache ID 75yI9Hsj9]
29. Kim H. Lessons from the Lee Sedol vs AlphaGo match. Korea IT Times 2016 Mar 14:1.
30. Kim C. Chosun.com. Dr Watson and the medical staff...patient URL: http://news.chosun.com/site/data/html_dir/2017/01/12/2017011200289.html [accessed 2019-02-06] [WebCite Cache ID 75yIZJ9WE]
31. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
32. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. Artif Intell Med 2014 Sep;62(1):1-10. [doi: 10.1016/j.artmed.2014.06.004] [Medline: 25059820]
33. Knight BA, Potretzke AM, Larson JA, Bhayani SB. Comparing expert reported outcomes to national surgical quality improvement program risk calculator-predicted outcomes: do reporting standards differ? J Endourol 2015 Sep;29(9):1091-1099. [doi: 10.1089/end.2015.0178] [Medline: 25936420]
34. Dawes T, de Marvao A, Shi W, Fletcher T, Watson G, Wharton J, et al. Machine learning of three-dimensional right ventricular motion enables outcome prediction in pulmonary hypertension: a cardiac MR imaging study. Radiology 2017 Dec;283(2):381-390 [FREE Full text] [doi: 10.1148/radiol.2016161315] [Medline: 28092203]
35. James R. National Institute of on Minority Health and Health Disparities. 2017. National Institutes of Health Precision medicine initiative URL: https://www.nimhd.nih.gov/about/legislative-info/clips/pmi.html [accessed 2019-02-05] [WebCite Cache ID 75yNuNGg5]
36. Liu Y, Gadepalli K, Norouzi M, Dahl G, Kohlberger T, Boyko A, et al. Detecting cancer metastases on gigapixel pathology images. arVix 2017:1703-02442v2 [FREE Full text]
37. Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities. Br J Gen Pract 2018 Mar;68(668):143-144. [doi: 10.3399/bjgp18X695213] [Medline: 29472224]
38. Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. JAMA 2016 Dec 13;316(22):2353-2354. [doi: 10.1001/jama.2016.17438] [Medline: 27898975]
39. Kayser K, Görtler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E, et al. AI (artificial intelligence) in histopathology--from image analysis to automated diagnosis. Folia Histochem Cytobiol 2009 Jan;47(3):355-361 [FREE Full text] [doi: 10.2478/v10042-009-0087-y] [Medline: 20164018]

XSL•FO
RenderX

40. Krittanawong C. The rise of artificial intelligence and the uncertain future for physicians. Eur J Intern Med 2018 Dec;48:e13-e14. [doi: 10.1016/j.ejim.2017.06.017] [Medline: 28651747]
41. Zhou S, Greenspan H, Shen D. Deep Learning for Medical Image Analysis. Cambridge: Elsevier; 2017.
42. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016 Dec 13;316(22):2402-2410. [doi: 10.1001/jama.2016.17216] [Medline: 27898976]
43. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Jan 25;542(7639):115-118. [doi: 10.1038/nature21056]
44. Ehteshami BB, Veta M, Johannes VD, van Diest B, Karssemeijer N, Litjens G, the CAMELYON16 Consortium, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017 Dec 12;318(22):2199-2210 [FREE Full text] [doi: 10.1001/jama.2017.14585] [Medline: 29234806]
45. Yu K, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 2016 Dec 16;7:12474 [FREE Full text] [doi: 10.1038/ncomms12474] [Medline: 27527408]
46. Smith A, Anderson M. Pew Research Center. 2017 Oct 4. Automation in everyday life URL: http://www.pewinternet.org/2017/10/04/automation-in-everyday-life/ [accessed 2019-02-05] [WebCite Cache ID 75yShocGY]
47. 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. [doi: 10.2196/12106]
48. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison (2). J Med Internet Res 2004 Oct 29;6(4):e39. [doi: 10.2196/jmir.6.4.e39]
49. Crouch S, Robinson P, Pitts M. A comparison of general practitioner response rates to electronic and postal surveys in the setting of the National STI Prevention Program. Aust N Z J Public Health 2011 Apr;35(2):187-189. [doi: 10.1111/j.1753-6405.2011.00687.x] [Medline: 21463418]
50. VanGeest JB, Johnson TP, Welch VL. Methodologies for improving response rates in surveys of physicians: a systematic review. Eval Health Prof 2007 Dec;30(4):303-321. [doi: 10.1177/0163278707307899] [Medline: 17986667]
51. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009(3):MR000008. [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]

Abbreviations

AI: artificial intelligence
IQR: interquartile range

Edited by M Focsa; submitted 15.10.18; peer-reviewed by Y Liu, J Rey-Martinez, S Zheng; comments to author 03.12.18; revised version received 09.01.19; accepted 17.01.19; published 25.03.19.

Please cite as:
Oh S, Kim JH, Choi SW, Lee HJ, Hong J, Kwon SH
Physician Confidence in Artificial Intelligence: An Online Mobile Survey
J Med Internet Res 2019;21(3):e12422
URL: http://www.jmir.org/2019/3/e12422/
doi:10.2196/12422
PMID:30907742

©Songhee Oh, Jae Heon Kim, Sung-Woo Choi, Hee Jeong Lee, Jungrak Hong, Soon Hyo Kwon. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design and Evaluation of Personalized Motivational Messages by a Virtual Agent that Assists in Post-Traumatic Stress Disorder Therapy

Myrthe L Tielman[1], PhD; Mark A Neerincx[1,2], PhD; Willem-Paul Brinkman[1], PhD

[1]Delft University of Technology, Delft, Netherlands

[2]Netherlands Organisation for Applied Scientific Research (TNO), Soesterberg, Netherlands

Corresponding Author:
Myrthe L Tielman, PhD
Delft University of Technology
Van Mourik Broekmanweg 6
Delft,
Netherlands
Phone: 31 643545851
Email: m.l.tielman@tudelft.nl

Abstract

Background: Systems incorporating virtual agents can play a major role in electronic-mental (e-mental) health care, as barriers to care still prevent some patients from receiving the help they need. To properly assist the users of these systems, a virtual agent needs to promote motivation. This can be done by offering motivational messages.

Objective: The objective of this study was two-fold. The first was to build a motivational message system for a virtual agent assisting in post-traumatic stress disorder (PTSD) therapy based on domain knowledge from experts. The second was to test the hypotheses that (1) computer-generated motivating messages influence users' *motivation to continue* with therapy, *trust in a good therapy outcome*, and the *feeling of being heard* by the agent and (2) personalized messages outperform generic messages on these factors.

Methods: A system capable of generating motivational messages was built by analyzing expert (N=13) knowledge on what types of motivational statements to use in what situation. To test the 2 hypotheses, a Web-based study was performed (N=207). Participants were asked to imagine they were in a certain situation, specified by the progression of their symptoms and initial trust in a good therapy outcome. After this, they received a message from a virtual agent containing either personalized motivation as generated by the system, general motivation, or no motivational content. They were asked how this message changed their *motivation to continue* and *trust in a good outcome* as well as how much they felt they were *being heard by the agent*.

Results: Overall, findings confirmed the first hypothesis, as well as the second hypothesis for the measure *feeling of being heard* by the agent. Personalization of the messages was also shown to be important in those situations where the symptoms were getting worse. In these situations, personalized messages outperformed general messages both in terms of *motivation to continue* and *trust in a good therapy outcome*.

Conclusions: Expert input can successfully be used to develop a personalized motivational message system. Messages generated by such a system seem to improve people's *motivation* and *trust* in PTSD therapy as well as the user's *feeling of being heard* by a virtual agent. Given the importance of motivation, trust, and therapeutic alliance for successful therapy, we anticipate that the proposed system can improve adherence in e-mental therapy for PTSD and that it can provide a blueprint for the development of an adaptive system for persuasive messages based on expert input.

(J Med Internet Res 2019;21(3):e9240) doi:10.2196/jmir.9240

KEYWORDS

mental health; motivation; trust; user-computer interface; PTSD; computer assisted therapy

XSL•FO
RenderX

Introduction

Untreated mental disorders account for 13% of the total global burden of disease, and the gap between treatment needed and treatment received is wide [1]. Electronic health (eHealth) can help remove some of the barriers to care for mental health, offering a cost-effective, accessible, and privacy-sensitive solution to mental health problems [2,3]. A broad range of e-mental health systems exists, differing in purpose and domain. Some systems are only meant for monitoring [4], whereas others offer full therapeutic interventions [5,6]. The latter may be classified as behavior change support systems (BCSS), aiming to change unhealthy behavior patterns. Such systems can be developed for a range of different disorders such as substance abuse [7], anxiety [8], depression [9], or post-traumatic stress disorder (PTSD) [10]. For the success of such BCSS for mental health, the users need to stay motivated to use the system, because behavior change is difficult [11]. Motivational messages can, therefore, be a useful addition to e-mental health systems.

In this study, we focus on an eHealth system for PTSD treatment. PTSD patients have experienced one or more traumatic experiences and have symptoms such as intrusive memories, a persistent negative state, and active avoidance of anything related to their trauma [12]. Of the most common types of therapy for PTSD, 1 is cognitive behavioral therapy including an exposure component. During this exposure, patients need to break the pattern of avoidance by actively recollecting their traumatic experiences [13]. The idea behind exposure is that although tension will initially rise, the automatic fear reaction will eventually fade away [14]. This initial tension means, however, that the situation might temporarily get worse, something that is reflected in treatment manuals [13] and patient reports [15]. This potential worsening also highlights the importance of trust and motivation; it is important that patients still believe they will get better if they persevere. Although shown to be effective [16], exposure therapy is therefore also challenging. This means that in an eHealth therapy system for PTSD treatment such as in the study by Tielman et al [10], motivational messages can be particularly useful.

A therapy system for PTSD can present motivational messages in several ways. Increasingly, BCSS for mental health care include virtual agents to present content to patients [17]. These agents are virtual characters that communicate with the users of a system in some way, and they have been applied, for instance, for dementia [18], substance abuse [19], and PTSD [20]. Virtual agents are perceived more positively by users than a text-based interface [21]. Moreover, they have the potential to positively affect treatment compliance and outcome [22]. This implies that virtual agents can successfully influence users simply by being present. Their impact might partially be explained by self-determination theory, which describes the concept of motivation [12]. It distinguishes between extrinsic motivation, coming from external sources, and intrinsic motivation, which stems from an internal drive and is more powerful of the two. Motivation from BCSS is always extrinsic, but extrinsic motivation might be internalized, a process that is supported by relatedness [23]. When a person feels closer in relation to the source of the extrinsic motivation, the motivation is more likely to be internalized. Virtual agents have the potential to increase relatedness more than simple text messages [24], and are, therefore, suitable to present motivational messages.

Besides the presentation mode of motivational messages, their content is also an important factor. An important question when designing this content is how to tailor it to the user so it is sufficiently effective. This tailoring can be done based on several different factors. One possible factor is readiness to change, which is the underlying concept of the transtheoretical model (TTM) [25]. This is a motivational model that identifies 6 stages of change for health behavior, from where people are not ready yet to make a change to where a change has occurred but needs to be maintained. To change behavior, TTM states a person should move from one stage to the next. Motivational interviewing (MI) is a motivational tactic also based on the readiness to change concept, focusing on increasing motivation by highlighting the discrepancies between the current and the ideal situation [26]. In this way, readiness to change is slowly increased. Both MI and TTM have been applied to virtual agents in various applications [27-30]. However, these models are best considered as full motivational strategies as they focus less on the content of individual motivational messages. Moreover, people in different stages of change do not also prefer different motivational messages [31], making this a less suitable factor for tailoring.

Another possible way to tailor motivation to the user is to consider personality traits. Studies have shown that this influences what type of motivational messages is found preferable [32]. Similar results were found for the effect of masculinity and femininity [33]. Interpersonal differences seem to influence what type of motivation is preferred, but effect sizes in these studies are small. Other factors might, therefore, also be useful to tailor messages in an e-mental health system for PTSD. Moreover, studies with personality traits are commonly performed in the general population and are not focused on specific situations. In PTSD therapy, however, motivation is necessary for a specific task. It seems more useful, therefore, to tailor the messages to the specific health situation of the user instead.

To tailor motivational messages for PTSD therapy to the user's situation, specific parameters to describe a situation need to be established. In the remainder of this paper, *situation* will, therefore, be defined in terms of the *progression of PTSD symptoms* and the user's *trust in a good therapy outcome.* These 2 parameters can be related to several theories on health behavior and motivation. Protection motivation theory proposes threat appraisal and coping appraisal as main predictors of health behavior [34]. Progression of PTSD symptoms relates to threat appraisal, as progression allows one to predict how successful therapy will be. Trust in a good therapy outcome, on the other hand, relates to coping appraisal, that is, how much a person feels they will be able to handle what is coming. Similar concepts are found in other theories such as the health belief model [35], which describes parameters such as perceived threat, perceived barriers, and self-efficacy.

XSL•FO
RenderX

To actually develop situation-based motivational messages for a virtual agent assisting in PTSD therapy, their specific content needs to be established. Some systems base the content of their motivational messages on concepts from the literature [36]. However, literature appropriate to the domain is not always available. If larger amounts of data are available on the motivational messages used by therapists, machine-learning techniques could, in principle, also be used to classify which messages are used in which situations. However, such data are rare, given the private nature of conversations between therapists and patients. Another way to build a motivational message system is to code human-human interactions for the use of a virtual agent [37]. This assumes, however, that the virtual agent has the same input from the user as a human would. This is usually not the case as it includes visual cues such as facial expression and posture, which are not always available for analysis. A final option to gather content for motivational messages is to rely on experts to write the messages [8]. This results in an appropriate list, but experts need to write many full messages to allow the system to vary its messages. A way to bypass this final problem would be to break up the messages into smaller, categorized statements that can be recombined into a new message. This allows for the generation of more variation in messages and avoids repetitiveness, which is especially desirable for agents used for long-term interactions [38]. Moreover, by also asking experts to write messages for specific situations, the categorized statements in these messages could be matched to those situations in which they are most suitable.

Therefore, the first objective of this paper is to introduce a method for designing a situation-based motivational system in the domain of PTSD therapy based on expert input. However, the question then remains what effect these motivational messages have on users when they are presented by a virtual agent. Therefore, the second objective of this paper is to empirically study the effect of these motivational messages.

The Methods section first presents our vision for the motivational system and introduces clear hypotheses for the empirical study. It then first outlines the methods for actually creating this message system using expert input. Second, it describes the methods of an empirical study, which was conducted to test the effect of personalized motivational messages presented by a virtual agent. The Results section describes the results of this empirical study, and the Discussion section discusses our findings.

Methods

Vision

This paper envisions a situation-based motivational message system for a virtual agent for PTSD therapy, as shown in Figure 1. The system generates motivational messages based on a patient's *situation*, which is defined by the *PTSD symptom trend* and the patient's *trust in a good therapy outcome* in this context. Such messages consist of different types of statements. These *statement types* are sentences distinguished by their content, for instance, pointing out the situation, asking a question, or giving a compliment. Although questions are not technically statements, we are categorizing them here as such, as we use term "statement" to refer to a sentence that is a part of a motivational message. As the experts knew the virtual agent would never be able to understand any answers to the questions, we interpret these as rhetorical questions. To enable the system to process PTSD symptom trend data, the brief PTSD-check-list (PCL) is used [39]. Through showing PCL scores for past sessions in a graph, *PCL trend* can be defined as dropping (symptoms improve), stable (symptoms stay roughly the same over sessions), or rising (symptoms worsen). *Initial trust in a good therapy* outcome is defined at 3 levels as well and can be either low, medium, or high. This results in a total of 9 possible situations for the system to adapt to.

Given these situations, a suitable *motivational message* can be composed of different *motivational statements*, which are the shorter motivational phrases, categorized by *statement type.* This is done using a database with different types of such motivational statements. The database includes likelihoods that express how suitable a motivational statement of a certain type is, given the situation. By combining specific motivational statements from each type, the system can generate diverse motivational messages suitable for a patient's situation.

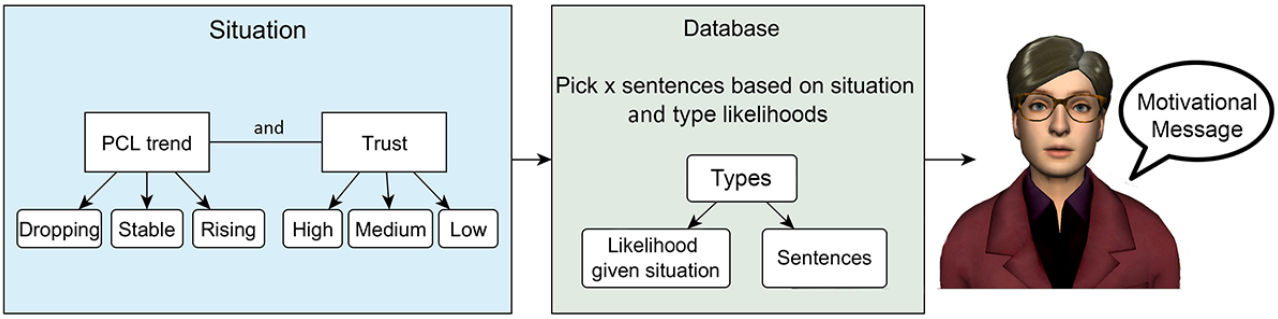
This database, therefore, consists of specific motivational statements, their type, and their suitability for a given situation. This paper envisions that this database can be composed using input from experts familiar with PTSD therapy. Consequently, the first objective is to *build a system that composes motivational messages personalized to user situation based on input from PTSD therapy experts.*

This objective is addressed in the Expert Study section, describing the expert study that was performed and how the eventual motivation system was composed based on the results. In addition, we are also interested in testing the effect of these motivational messages on the people receiving them. The motivational message system is first meant to increase a patient's *motivation to continue* with therapy. Second, it should also increase their *trust in a good therapy outcome*. Finally, given the importance of therapeutic alliance on the therapeutic process [40], the personalization also serves to increase the patients' feeling that they are being taken seriously and heard by the virtual agent. The 2 hypotheses were therefore formulated. First, this paper hypothesizes that motivational messages can increase the 3 variables mentioned above. Second, it hypothesizes that by personalizing the motivational messages to a user's situation, the *motivation*, *trust*, and *feeling of being heard* will improve even further. The following 2 hypotheses were evaluated in a second study performed with users, as presented in the User Study section:

- H_1: Motivational messages improve trust in a good outcome, motivation to continue, and the feeling of being heard more than messages without motivational content.
- H_2: Personalized motivational messages as generated by the motivational system improve trust in a good outcome, motivation to continue, and the feeling of being heard more than general motivational messages.

XSL•FO
RenderX

Figure 1. Vision of the situation-based motivational message system. PCL: post-traumatic stress disorder-check-list.



Expert Study

To achieve the first objective, the motivational message system was created. A database was built including motivational statements, their types, and the likelihood they would be used by experts in a given situation.

Participants

A total of 13 therapists (5 male and 8 female) were recruited from 6 different mental health clinics. All experts had professional experience in treating patients with PTSD.

Procedure

All experts were presented with 9 different situations expressed by patient trust in therapy outcome and a graph representing PCL trend. Every situation was presented twice, using 2 different graphs, as shown in Figure 2. This was done to ensure the responses would not reflect 1 specific peculiarity of the graph, but the trend shown therein. For every situation, the expert was asked to write what they would say to the patient to increase their *motivation to continue* and their *trust in a good therapy outcome*. As context, the experts were given the example of the PTSD therapy system presented in the study by Tielman et al. [10]. All situations described a patient in session 8 out of the total 12, which is right in the middle of the most intense phase of treatment. The design of this expert study was approved by the university ethics committee, ID number 134.

Expert Answer Analysis

From this study, we received 234 *motivational answers*, that is, motivational messages written by the experts for the situations. These answers were first split into shorter *motivational statements* and categorized to provide them with a *statement type*, after which an analysis was done to calculate the likelihood of the statement types occurring in each situation.

Expert Input Categorization

The motivational answers given by the experts were between 1 and 143 words long (Mean 40.40, SD 24.68). The single word Dutch statement “Terecht,” which translates into English as “and you should!,” was used once in a positive situation with high trust. As this word was also included in longer statements, it was categorized as the other “terecht”s, which were part of larger answers and incorporated in the analysis in this manner. Most answers contained several sentences, including a number of messages to the patient. To allow for a more fine-grained categorization, the answers were split into *motivational statements* based on topic. Splitting generally happened after commas or between sentences. This relatively low-inference procedure was as follows: whenever a statement could be categorized in multiple ways and ambiguity could be resolved by splitting, the statement was split. Splitting yielded 844 *motivational statements*. The rounded mean number of *motivational statements* per motivational answer was 3 for the *dropping PCL-high trust*, *dropping PCL-medium trust*, and *stable PCL-high trust* situations, and 4 for the others. All motivational statements were manually sorted into types based on their topic. These *statement types* were not predefined but arose during the analysis. For some, additional subtypes were defined to better describe nuances.

After initial categorization, all statement types that included less than 10 motivational statements (in the total set of 844) were removed. After this selection, 97% of the motivational statements were left to be included in the analysis, namely, those with a statement type shown in Figure 3. To analyze the reliability of the coding process, a sample of 32 statements was selected (including 2 full motivational answers for every situation) and rated by a second coder. Comparison between coders showed a substantial interrater reliability (κ=.73, P<.001).

It is interesting to note that *medium trust* was not mentioned explicitly by the experts, and thus does not occur in the list of statement types. Although this medium level was included in the situations (defined as “doubts about a good outcome”), in categorizing the answers of experts, no distinction between low and medium trust could be made. Although differences between these situations can be seen in how much other statement types were used, the specific mentions of trust were similar in both situations. This might be because both were interpreted as low trust, one just even lower than the other.

XSL•FO
RenderX

Figure 2. Graphs describing post-traumatic stress disorder symptoms in terms of post-traumatic stress disorder-check-list trend, as presented to the experts in the expert study. PCL: post-traumatic stress disorder-check-list.

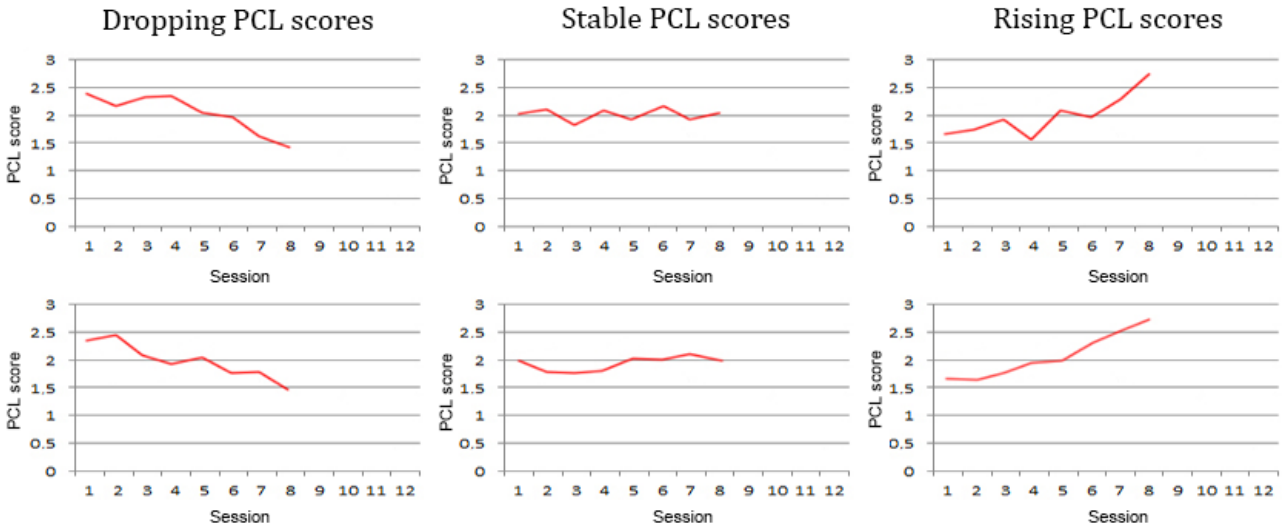
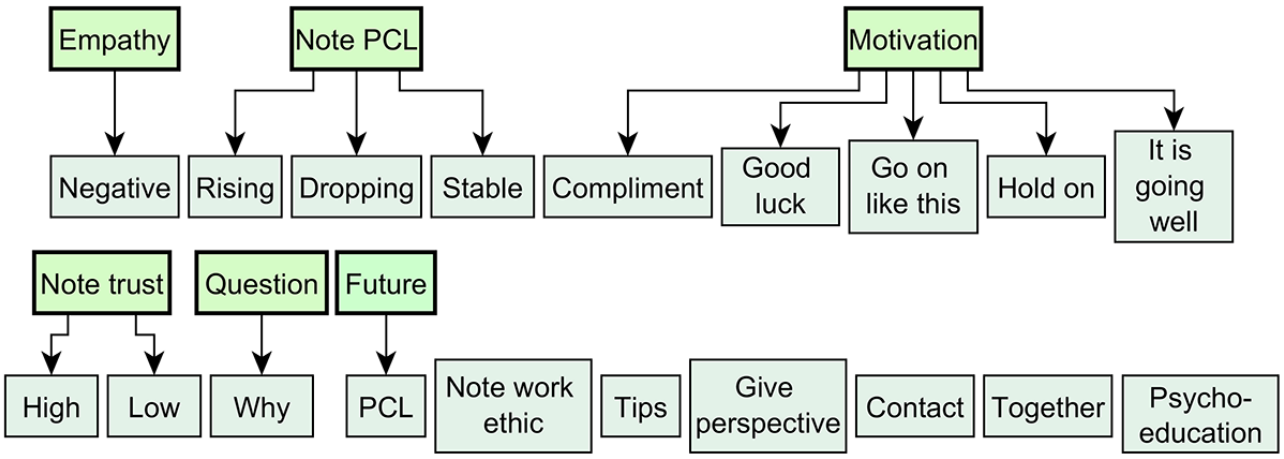


Figure 3. Types of motivational statements found in the expert survey. Thicker bordered green types are super types. PCL: post-traumatic stress disorder-check-list.



Extraction of Likelihood Rates

To calculate the likelihood of a statement type occurring in the motivational answer given a certain situation, the data were analyzed with R 3.3. First, it was noted whether a statement type was present for every motivational answer given by the experts. A multilevel analysis with *expert* as random intercept was done, showing that adding *statement type* as a fixed factor significantly improved the model predicting whether a statement type occurred in the motivational answers ($F_{49,11638}$=88.86; P<.001). This indicates that there were differences in how many different statement types were used. Second, an additional analysis showed that the model containing only *statement type* was significantly improved by adding the 3-way interaction effects for *PCL trend*, *trust*, and *statement type* as a fixed factor (χ^2_3=1954.80; P<.001). This shows that the situation influenced whether a statement type was used in the motivational answer. Therefore, a second multilevel logistic regression was fit for each statement type separately, predicting if a statement type occurred in the motivational answers for that situation. This analysis used *expert* as the random intercept and the 2-way interaction between *PCL trend* and *trust* as the fixed effect. This resulted in the probability of each statement type occurring in a motivational message in a given situation.

These probabilities form the basis of our database. It was supplemented with a list of motivational statements and their type. Motivational statements were included in this list only if the probability for their statement type was above .05 for at least one of the situations. So, those statements with rare statement types were excluded. The list of motivational statements in the system is, therefore, a subset of the motivational statements used in the analysis. All motivational statements were slightly rewritten from how they originally occurred in the motivational answers to ensure they could be used consecutively. Both the full probability table as well as the resulting motivational statements are available online [41].

Generating Motivational Messages

Our analysis resulted in a database with motivational statements, their statement types, and their likelihood in any situation. This section describes how this database is used to generate full motivational messages. We have developed an algorithm which selects motivational statements of certain types from the database, given a situation, and which combines them into 1

XSL•FO
RenderX

message. Throughout this section, we will use a running example of generating a message for the situation with *rising PCL* and *low trust.*

First, the length of the motivational message is determined. The motivational answers by the experts revealed that some situations might warrant more motivational statements than others. Therefore, either 3 or 4 motivational statements were combined, depending on the rounded mean number of motivational statements given by the experts in their motivational answers for that situation. For our example message for a situation with *rising PCL* and *low trust*, the motivational message will contain 4 motivational statements.

Second, the statement types are selected. In this, we follow the probability table for the situation. First, any statement types with probabilities above .5 are always included. For our example, this means that the statement types *motivation*, *give perspective*, and *note PCL* will be included.

If any of these statement types with a probability above .5 is a super-type, all subtypes with a probability above .5 are included; if none meet this criterion, one is chosen based on a weighted randomized selection strategy. This strategy entails that all probabilities are added, and every statement type is assigned an interval equal to their probability. By generating a random number between zero and the total sum of the probabilities, the statement type belonging to the interval in which the random number occurred is selected. For our example, the statement types *motivation* and *note PCL* are super types, so a subtype needs to be selected. For *motivation*, none of the subtypes have a probability over .5, so a choice is made with the weighted random selection strategy. For this example, let us say we end up choosing *hold on*. For the *note PCL* type, we do have a subtype with a probability over .5, namely, *rising*, so this will be chosen.

Then, if the expected number of statements is not met after this initial selection, additional statement types are added one by one using the weighted randomized selection strategy. Duplicate statement types are avoided so that if a statement type is already included, it cannot be chosen again. For our example, we still need one more statement type. A final type is chosen with the selection strategy, let us say we end up with the statement type *empathy.*

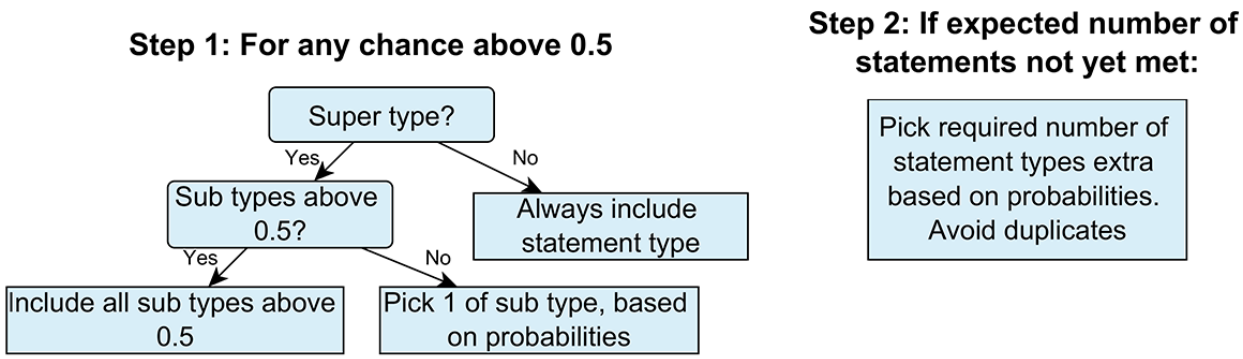
The full selection process for the statement types is described in Figure 4. This process was designed to fit different probability tables. For instance, for our probability table (see the resource by Tielman et al [41]), no statement type has 2 subtypes with probabilities higher than .5, but this algorithm could deal with this situation as well. For application to other probability tables, the only possible change would be the threshold of .5 to always include a statement type. This number was chosen because for our results it meant 2 or 3 statement types are always automatically selected, whereas at least one is always added via the weighted randomized strategy. This ensures variation while also ensuring the most important statement types are always included.

Third, to complete the motivational message, a motivational statement is selected at random for every statement type that is included. This means that even if the statement types for a certain situation are always the same, the specific motivational statements making up the motivational message can differ. The ordering of the motivational statements within the message is inspired by the original motivational answers given by the experts. They are sorted based on statement type, such that *noting the situation* always comes first, followed by *empathy*, *giving perspective*, and statements about tackling things *together*. *General motivation* is always near the end, only followed by statements about the *future*.

For example, this final step means randomly choosing motivational statements from the database for the statement types and putting them in the order: *note PCL-rising*, *empathy*, *give perspective*, and *motivation-hold on.* This could result in the following motivational message (translated from the original Dutch):

> *I see you indicate that your complaints have gotten substantially worse.* [note rising PCL]
>
> *I'm sorry to hear that.* [Empathy]
>
> *However, it's always hard work before we see any results.* [Give perspective]
>
> *Hold on!* [Motivation hold on]

Figure 4. Selection process for choosing the statement types for the motivational message system.



XSL•FO
RenderX

User Study

To test hypotheses 1 and 2, an empirical user study was done. The *personalized motivational messages* as generated by the motivational message system described in the previous section were compared with 2 other *message types*, namely, *general motivational messages* and messages *with no motivational content*.

The motivational message system is designed for PTSD patients following therapy assisted by a virtual agent. However, to evaluate the message system with patients in therapy would mean a full clinical trial, which traditionally mainly focuses on clinical outcome. Instead, eHealth systems are often evaluated in phases [42], and separate components are first tested using proximal outcome measures such as motivation [43,44]. Testing components separately also allows for a clearer picture of exactly what effect a certain manipulation, such as motivational message type (ie, personalized, generic or no motivational message), has on specific outcome variables, such as *motivation to continue, trust in a good therapy outcome, or* the feeling the user has of *being heard*. This abstraction strategy is, for instance, also used in game design to reduce surrounding and complicating factors and speed up the process [45]. The evaluation of the motivational message system takes this component-focused approach. As it was tested in isolation instead of incorporated into a full therapy, participants were asked to imagine they were in a certain situation during PTSD therapy, as defined by symptom trend and trust in therapy outcome. One can imagine, however, that people who have experience with PTSD exposure therapy will interpret the situations differently from people who do not. Participants were, therefore, also asked for their experience with PTSD symptoms and exposure therapy so that these factors could be taken into account as potential covariates.

The message system was tested in a 3×3×3 mixed design. Message type (personalized, general, or no motivation) was measured between-subject, whereas the 3 graphs describing PTSD symptom trend and the 3 levels of initial trust in therapy outcome were presented to all participants within-subject. After seeing the situation, participants were presented with a virtual agent presenting a personalized motivational message, a general motivational message, or a message with no motivational content. Following this, participants were asked to indicate changes in their *motivation to continue* and *trust in a good therapy outcome* for that situation. The design of this user study was approved by the university ethics committee as well, ID number 184.

Participants

To study the effect of the situation-based motivational message system, participants were recruited via Amazon Mechanical Turk. An a priori power analysis was performed (G*Power 3.1) to determine the necessary number of participants. Given a one-way Analysis of variance (ANOVA) with 3 groups, a medium effect size of 0.25, and power of 0.9, a preferred sample size of 207 was calculated. A total of 529 participants started the experiment. Participants were excluded if they incorrectly answered a control question (to check if they had read all situations and questions; n=189), if they did not complete the survey (n=128), or in case of administrative errors (n=5). Participants were excluded directly after completing the study, which kept running until the required number of 207 included participants was met. Of the excluded participants, 25.5% were in the personalized motivation condition, 27.6% in the general motivation condition, 29.5% in the no motivation condition, and 30.1% stopped before being assigned a condition. Of the 207 included participants, 34.3% were in the personalized motivation condition, 34.3% in the general motivation condition, and 31.4% in the no motivation condition. All participants were aged 18 years and above and native English speakers and received a small monetary compensation for their time. Due to an administrative error, specific age and gender data of participants were not collected.

Measures

A total of 5 different measures were collected via questionnaires, 2 of them repeatedly for every situation that was presented. A list of all questions is available online [41].

Primary Measures

Trust in a good therapy outcome was measured in terms of change. After each situation and message, the question of how much the message of the virtual agent changed *trust in a good therapy outcome* was posed. Answers ranged on an analog scale from −10 (*it decreased a lot*) via the neutral point of 0 (*nothing changed)* to 10 (*it increased a lot)*. This question was repeated at the end of the study, asking for the overall change in *trust*.

Motivation to continue was measured similarly to trust, in terms of change caused by the message of the virtual agent, asked for in every situation. The analog scale also ranged from −10 (*it decreased a lot)* via the neutral point of 0 (*nothing changed)* to 10 (*it increased a lot)*, and this question too was repeated at the end asking for the overall change in *motivation*.

Feeling of being heard was measured only at the end of the experiment and was included to get a measure of how much the participants felt the agent took them seriously and really heard them. This questionnaire consisted of items 1 and 4 of the short patient satisfaction questionnaire [46], and items 5 and 8 from the trust in physician scale [47]. Questions were slightly adapted to be positively phrased statements. Moreover, 3 additional questions were added to complete the questionnaire, stating that *the virtual agent took you seriously*, the *virtual agent replied appropriately to you*, and the *virtual agent listened to your preferences*.

Descriptive Measures

To get an indication of whether participants had experience with PTSD symptoms, a brief version of the PCL was presented at the beginning of the study, including questions 2, 3, 6, 7, 16, and 17. People were asked to keep in mind their most stressful or traumatic experience when answering these questions. Participants were also asked if they had experience in following exposure therapy, and if so for what disorder.

Procedure

After accepting the assignment on Amazon Mechanical Turk, participants were redirected to Qualtrics for the main experiment. They were first presented with an information form

XSL•FO
RenderX

outlining the procedure of the experiment and the consent form. Then, they were given the PCL and asked if they had ever followed exposure therapy. Before the main part of the experiment, all participants saw an example situation with a PCL graph that was not reused in the remainder of the experiment, a clip from the virtual agent introducing itself, and the motivation and trust questions. This example included the explicit instruction that a rising PCL trend meant symptoms were getting worse and vice versa. After the example, participants were randomly divided into the 3 conditions (with personalized, general, or no motivation messages). The virtual agent always appeared as a video clip and displayed idle and mouth movement, speaking with a computer-generated text-to-speech voice. The virtual agent is shown in Figure 5.

In all conditions, participants were asked to repeatedly imagine they were in therapy and had just answered the PCL resulting in the situation shown. All 9 situations representing the combinations of PCL trend (rising, stable, and dropping) and trust level (no trust, doubt, and high trust) were shown in a randomized order. To avoid overfitting, a new set of PCL graphs was used, slightly different than the one shown to the experts. These new graphs used in the user study are shown in Figure 6. After reading the situation, participants could view the message by the virtual agent. In the condition with the personalized motivational messages, 1 of the 3 comments generated by the motivational message system for that particular situation was used, translated from the original Dutch to English. For the condition with general motivational messages, 9 different general motivational messages were written by 3 different experts and translated into English. These general messages were matched with the situations in 3 different ways, such that each general message occurred once for each PCL trend and each trust level in the final set. Participants were randomly presented with 1 of these 3 combinations. The same type of randomization was performed for the messages without motivational content. In this condition, the virtual agent merely thanked the participant for filling in the PCL and announced the continuation of the session, using 9 different formulations. An example of a message from each condition is shown in Table 1. All messages for the different conditions can be found in the resource by Tielman et al [41].

After the message by the virtual agent, participants were presented with the question of how much their trust in a good outcome and *motivation to continue* had changed. Both the situation and the virtual agent were still visible at this point; so, participants had the option to review both the situation and the message. After all situations, these 2 questions were repeated for the overall experience, and the *feeling of being heard* questionnaire was presented. Figure 7 shows a full outline of the procedure.

Data Preparation and Analysis

Data were analyzed with R 3.3; the full data table and analyses are available online [41]. Following the guidelines for the PCL [39], a participant was classified as potentially having PTSD if question 2 or 3 was 3 or higher, question 6 or 7 was 3 or higher, and question 16 or 17 was 3 or higher. Reliability for the *feeling of being heard* scale was high with a Cronbach alpha of .96, so answers to the items of the *feeling of being heard* scale were averaged to form 1 score.

Covariates

To examine *potentially having PTSD* and *having experience with exposure* as potential covariate factors in the analysis, ANOVAs were done with these variables as predictors of overall change in *motivation*, *trust*, and *feeling of being heard score*. As these analyses did not find any effects, neither *potential PTSD* nor *exposure experience* were further taken into account as covariates.

Figure 5. Virtual agent.

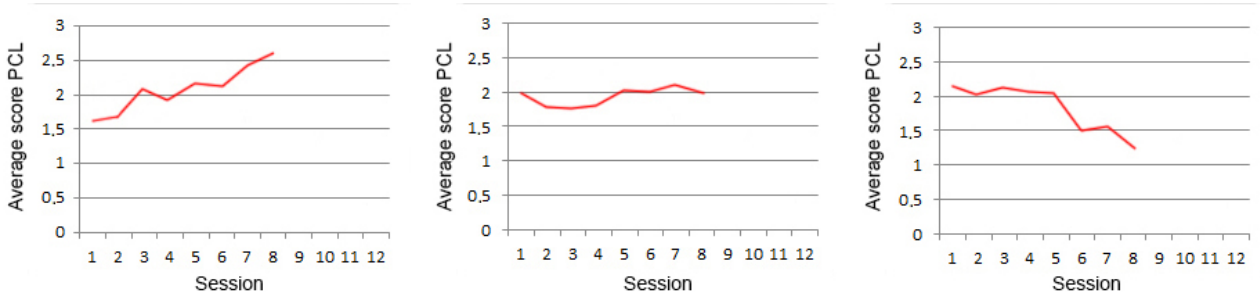


XSL•FO
RenderX

Table 1. Examples of the different types of messages presented.

Type	Message
Personalized to stable post-traumatic stress disorder-check-list scores and doubts	Your complaints show little consistency. Sometimes they get less, and then they get worse again. That makes sense, you're working on your past and that's tough. Stick to it now! Let's see together what you need to continue.
General	You've now been working on this treatment for a while. Perhaps you've noticed already that your complaints are slowly getting less. It's important to stick with it now, so we can achieve the largest result.
No motivation	Thank you for filling in the questionnaires. We'll now continue with the session.

Figure 6. Situations used in the experiment. From right to left: rising post-traumatic stress disorder-check-list, stable post-traumatic stress disorder-check-list, and dropping post-traumatic stress disorder-check-list. PCL: post-traumatic stress disorder-check-list.



Overall Measures

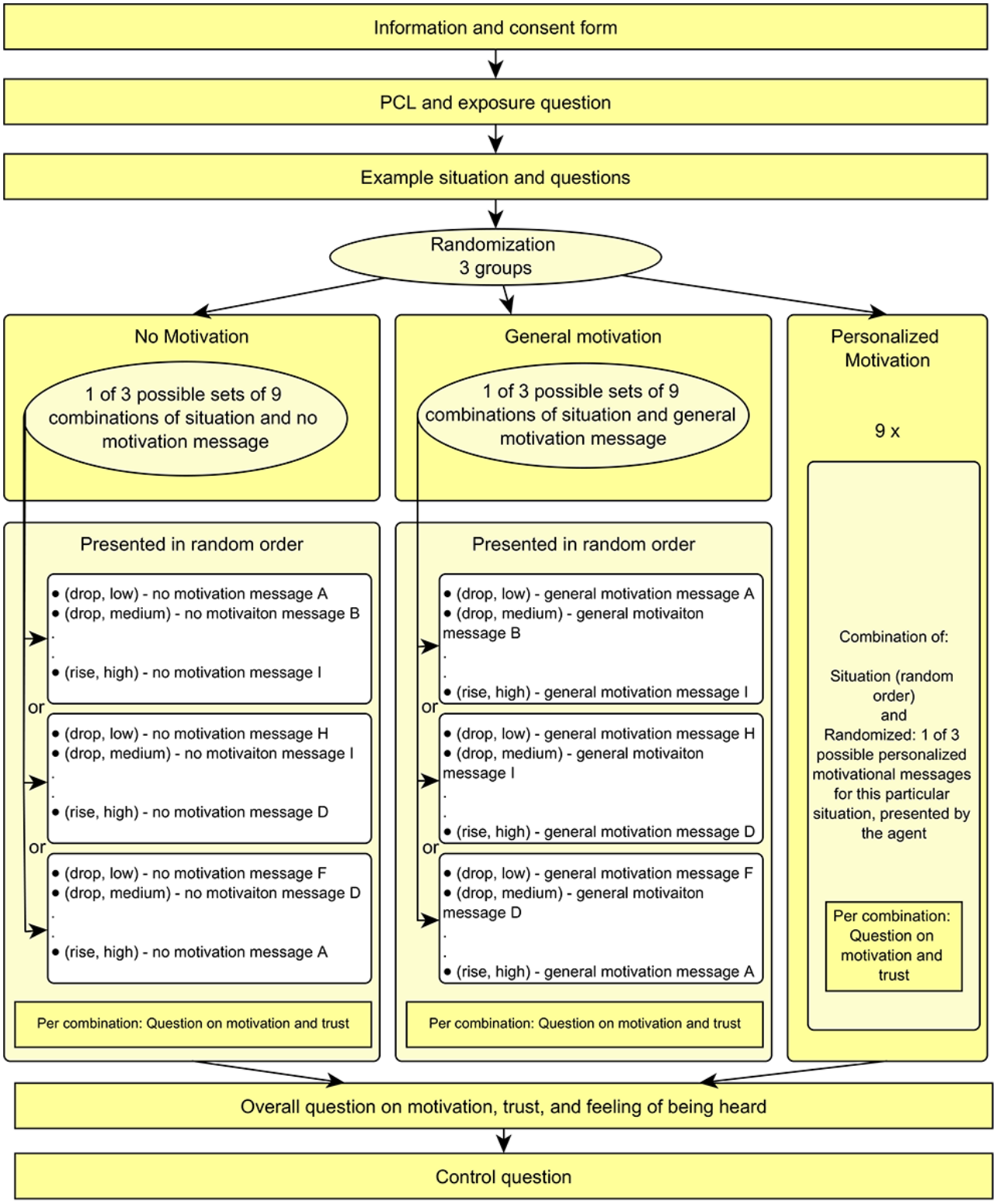
For all overall measures, a linear regression model with only the intercept was fit to establish a deviation from zero. *Type of message* was added to the model to establish if it affected the overall outcome measures. When this analysis showed a significant effect of *type of message*, individual post hoc analyses were run on data containing only 2 message types to test for differences between them.

Per Situation

To analyze the data per situation, every situation was treated as a separate data point. A multilevel regression was run with *participant* as random intercept. For both change in *motivation to continue* and *trust in a good outcome*, the null model with only a fixed intercept was compared with model 1, including *message type* (3 levels) as fixed effect. Model 2 built on model 1 and added *PCL trend* as fixed effect; model 3 added *initial trust level*. Further models were built, adding first the 2-way interaction effects and finally the 3-way interaction effect. Post hoc analyses were run on subsets of the data to find the differences between each of the 3 message types. This was done for the 3 possible *PCL trends*, the 3 levels *of initial trust*, and for the 3×3 combination of these measures resulting in all 9 situations.

XSL•FO
RenderX

Figure 7. Outline of the user study experiment procedure. PCL: post-traumatic stress disorder-check-list.



Results

Descriptive Measures

Table 2 shows the descriptive measures including their influence on the overall outcome measures. Scores on PCL show that almost a third of participants showed an indication of having PTSD. A total of 14 participants had experience with exposure therapy. Out of the 7 participants with experience with exposure for PTSD, 5 also showed an indication of PTSD. The other specific disorders mentioned were obsessive-compulsive disorder and both specific and general anxiety disorders. Table 2 also shows that *an indication of PTSD* or *exposure experience* were not found to affect overall change in *motivation*, *trust*, or *feeling of being heard score*. There was, therefore, no indication found that experience with PTSD or exposure therapy influenced the interpretation of the situations and the effect of the messages

XSL•FO
RenderX

by the virtual agent. This supports the generalizability of the results to clinical populations.

Overall Experiences

Table 3 shows the mean values of all overall outcome variables measured at the end of the experiment as well as their deviance from the neutral point of 0. Both change in *motivation* and *trust* were rated significantly above neutral, indicating that overall *motivation* and *trust* improved. The *feeling of being heard* did not deviate significantly from 0. For all 3 outcome measures, linear regression analyses revealed a significant effect for *message type*.

Post hoc analyses revealed the differences between each of the 3 *message types* for the overall variables. Figure 8 visualizes the results from this analysis, showing that *motivational messages* significantly improve overall increase in *motivation* and *trust*. Moreover, the *feeling of being heard* increases significantly not only if a *motivational message* is present but even more if this message is personalized to the participant's situation. The confidence intervals show that if a *personalized message* was presented, *motivation*, *trust*, and the *feeling of being heard* improved, whereas a message without motivational content always resulted in a decrease.

Per Situation

Table 4 shows the effects of *message type*, *PCL trend*, and *initial trust* on the outcome measures of *motivation to continue* and *trust in a good outcome*, as measured per situation. This table also includes the effects of the 2- and 3-way interactions.

This table shows that both *message type* and the 2 situational factors influence *motivation* and *trust* as measured per situation. Figure 9 visualizes the pair-wise comparison of the effect of *message type*, showing that *motivational messages* outperform messages with no motivational content. The differences between the 3 levels of *PCL trend* and *initial trust* are displayed in Figures 10 and 11, respectively, showing how these situational factors affected the outcome measures.

Besides individual effects, Table 4 also shows the existence of an interaction effect between *message type* and *PCL trend* for both *motivation* and *trust* and a further effect between *message type* and *initial trust* for *motivation*. Given these results, a pair-wise analysis was done comparing each of the message types given the 3 possible values of *initial trust* and *PCL trend*, shown in Figures 12 and 13, respectively. Figure 13 shows that *initial trust* level effects how much *motivation* and *trust* improve. Moreover, if the initial *trust* level is doubtful of low, receiving a message without motivation seems to even reduce *motivation to continue* and *trust in a good outcome*. Figure 13 shows that when the *PCL trend is rising* and the situation, therefore, is getting worse, *personalized motivational messages* significantly *outperformed general motivational messages* in terms of both, *motivation to continue* and *trust in therapy outcome*.

Finally, Table 4 shows the existence of a 3-way interaction effect of *message type*, *initial trust*, and *PCL trend* on *motivation to continue*. Figure 14 shows the effect of *message type* on *motivation to continue* for all 9 situations, combining *PCL trend* and *initial trust level*. These graphs show that the exact effect of *message type* differs per situation. In the situations where symptoms are getting worse, but *initial trust* is doubtful or high, *personalized motivational messages* show an advantage over *general motivational messages*. The same is true for the situation with dropping *PCL trend* and high *initial trust*, which shows not only that personalized motivation is best but also that general motivational messages do not even outperform messages without motivational content. In most other situations, motivational messages of some type outperform messages without motivational content. The only exception is the situation with dropping *PCL trend* and low *initial trust*, where *motivation* is improved most by general motivational messages.

Table 2. Descriptive characteristics of participants.

Characteristic	n (%)	Motivation		Trust		Feeling of being heard	
		F value (1, 205)	*P* value	*F value* (1, 205)	*P* value	*F value* (1, 205)	*P* value
Post-traumatic stress disorder indication	66 (32)	0.99	.32	0.13	.72	0.04	.83
Exposure experience	14 (7)	1.48	.22	1.18	.28	0.34	.56
Post-traumatic stress disorder	7 (3)	—[a]	—	—	—	—	—
Anxiety	6 (3)	—	—	—	—	—	—
Obsessive-compulsive disorder	1 (0)	—	—	—	—	—	—

[a]Not applicable.

Table 3. Description, overall measures, and effects of condition.

Variable	Mean (SD)	Deviance from 0		Effect of message type	
		t test (206)	*P* value	*F value* (2, 204)	*P* value
Motivation	1.06 (4.05)	3.76	<.001	20.58	<.001
Trust	0.74 (4.16)	2.56	.011	16.18	<.001
Feeling of being heard	−0.66 (4.83)	−1.96	.051	50.97	<.001

XSL•FO
RenderX

Figure 8. Effect of message type on the overall outcome measures. *P<.05, **P<.01. Error bars indicate 95% Confidence Intervals. FBH: feeling of being heard.

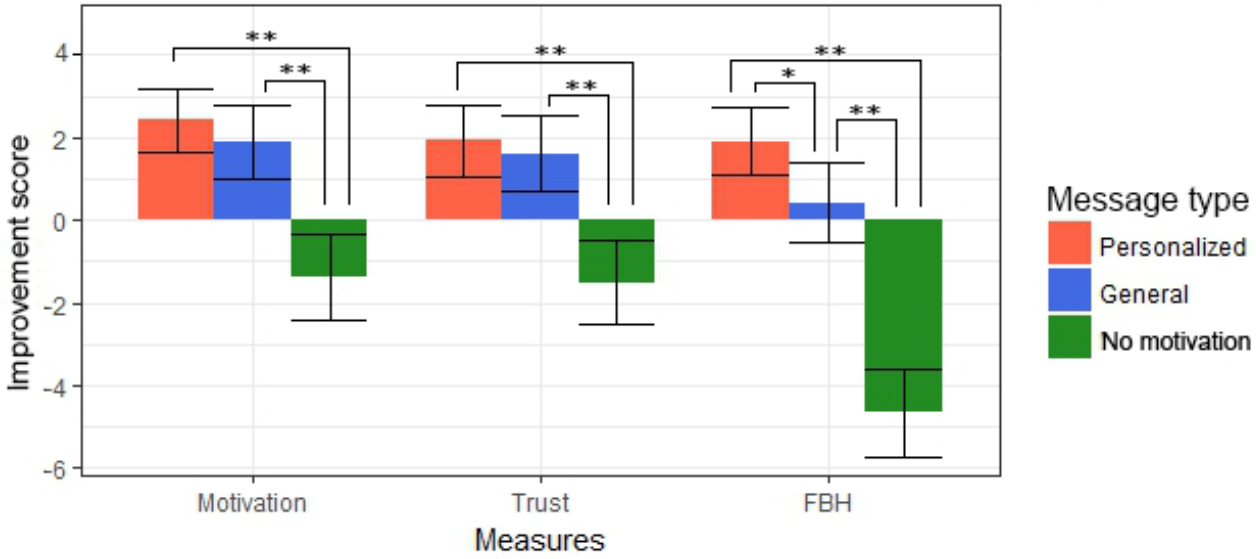


Table 4. Influence of situation and feedback type on *motivation to continue* and trust in therapy outcome (n=1863).

Model comparison	χ^2 (degrees of freedom)	P value
Motivation to continue		
Add Message type (M_0 vs M_1)	29.1 (5)	<.001
Add PCL[a] Trend (M_1 vs M_2)	27.5 (7)	<.001
Add initial trust (M_2 vs M_3)	40.5 (9)	<.001
Add initial trust × PCL Trend (M_3 vs M_4)	9.4 (13)	.05
Add message type × PCL Trend (M_4 vs M_5)	25.2 (17)	<.001
Add message type × Initial Trust (M_5 vs M_6)	11.2 (21)	.03
Add message type × Initial Trust × PCL Trend (M_6 vs M_7)	17.1 (29)	.03
Trust in outcome		
Add message type (M_0 vs M_1)	22.9 (5)	<.001
Add PCL trend (M_1 vs M_2)	55.7 (7)	<.001
Add initial trust (M_2 vs M_3)	59.9 (9)	<.001
Add initial trust × PCL Trend (M_3 vs M_4)	5.3 (13)	.26
Add message type × PCL Trend (M_4 vs M_5)	20.5 (17)	<.001
Add message type × initial trust (M_5 vs M_6)	7.4 (21)	.12
Add message type × initial trust × PCL trend (M_6 vs M_7)	12.7 (29)	.12

[a]PCL: post-traumatic stress disorder-check-list.

XSL•FO
RenderX

Figure 9. Influence of message type on motivation and trust per situation. **P<.01.

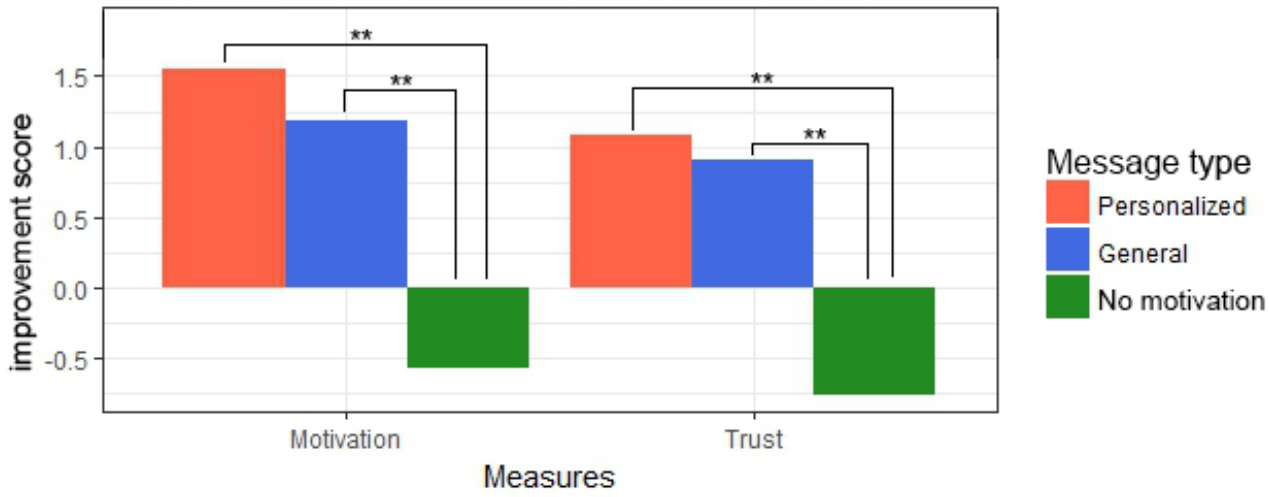


Figure 10. Influence of post-traumatic stress disorder-check-list trend on motivation and trust per situation. PCL: post-traumatic stress disorder-check-list. **P<.01.

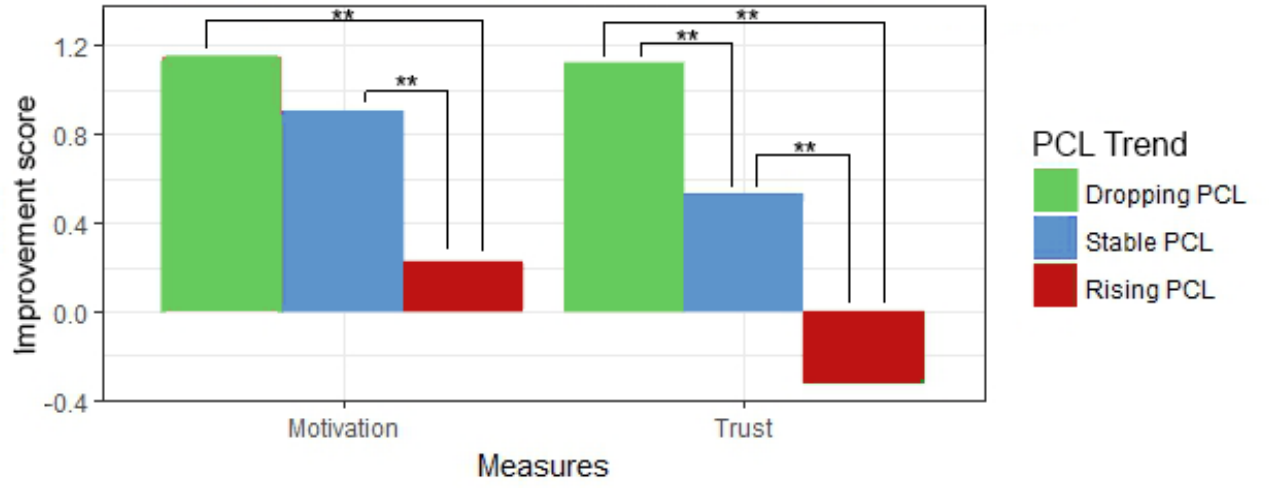
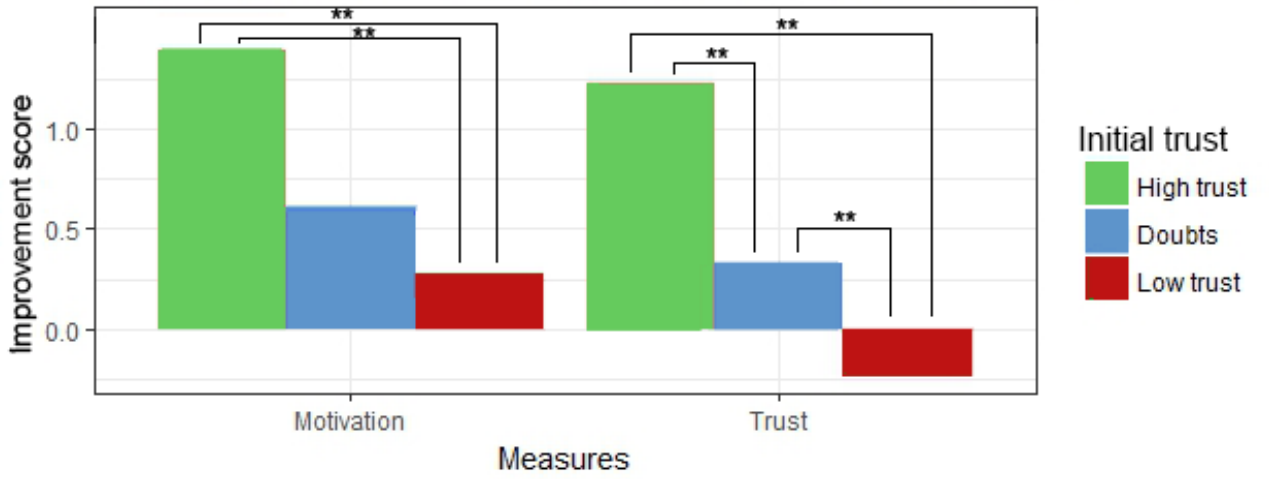


Figure 11. Influence of initial trust on motivation and trust per situation. **P<.01.



XSL•FO
RenderX

Figure 12. Effect of message type on motivation and trust, per initial trust level. *P<.05, **P<.01.



Figure 13. Effect of message type on motivation and trust, per post-traumatic stress disorder-check-list trend. *P<.05, and **P<.01. PCL: post-traumatic stress disorder-check-list.

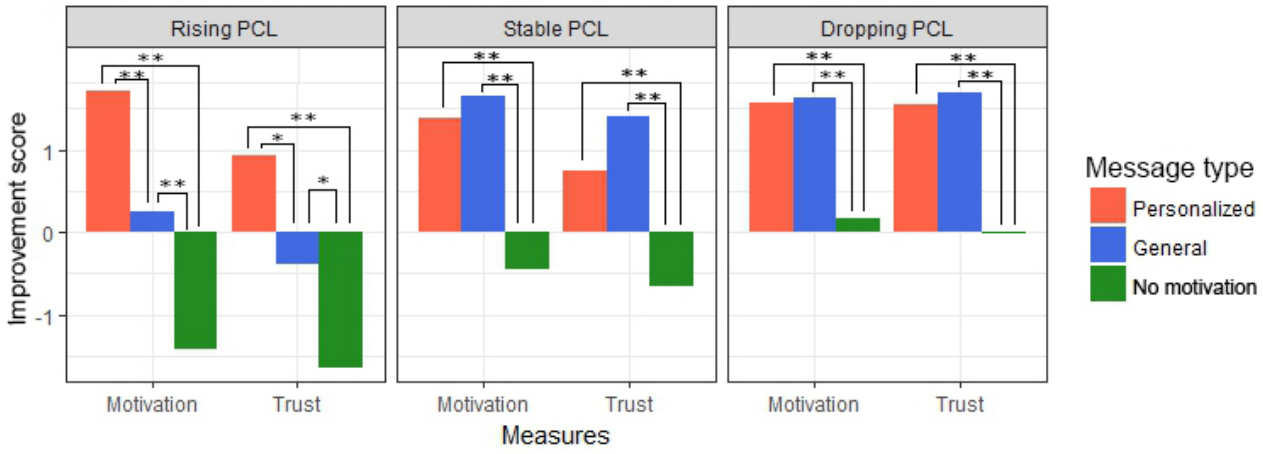
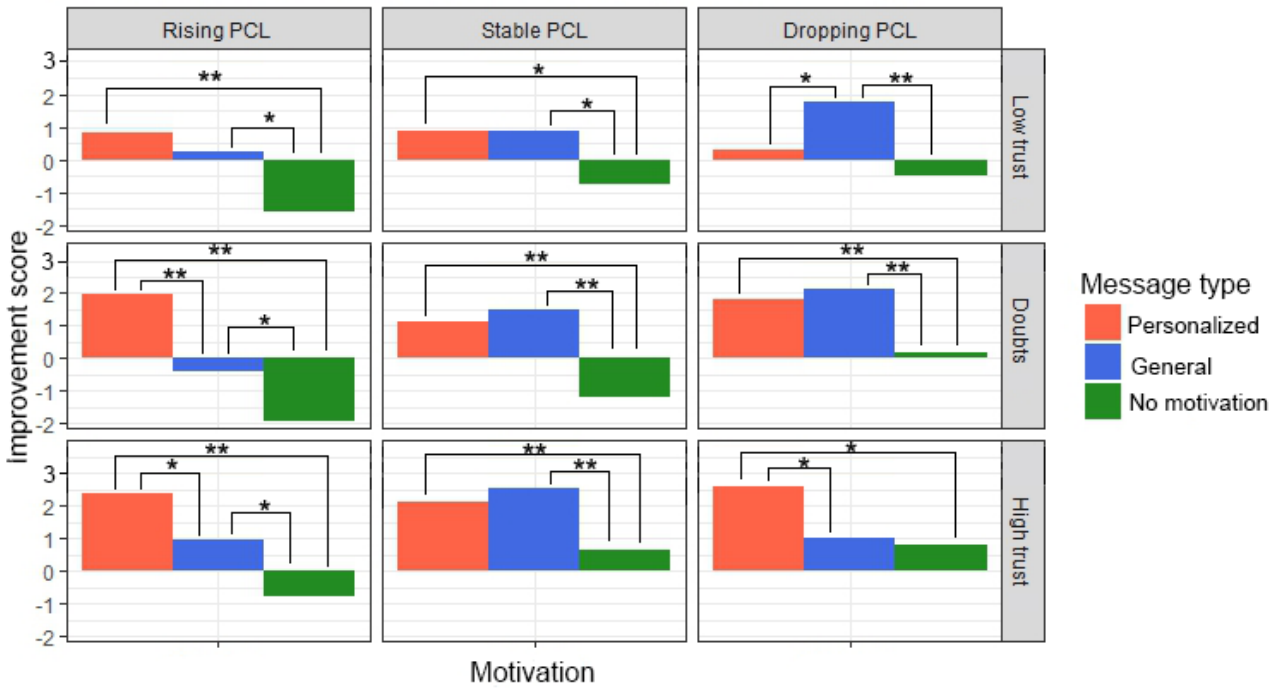


Figure 14. Effect of message type on motivation in all separate situations. *P<.05 and **P<.01. PCL: post-traumatic stress disorder-check-list.



XSL•FO
RenderX

Discussion

Results of the User Study

The aim of the user study was to test 2 hypotheses about the effect of the (personalized) motivational messages on the *motivation to continue*, *trust in a good therapy outcome*, and the *feeling of being heard by the agent*. The results show that the content of a message influences *motivation to continue*, *trust in a good outcome*, and how much participants *felt heard by the agent*. More specifically, messages with motivational content outperform messages without such content. This confirms hypothesis 1 and is in line with recommendations by experts that motivational messages are relevant [48].

The results regarding hypothesis 2 are a bit more complex. We hypothesized that motivational messages could be improved through personalization with our motivational system. Some findings support this hypothesis. This study shows that personalization can increase how much users feel the agent really *hears them* and takes them seriously. This result is important as it shows that personalization could increase a feeling of relatedness to an agent. Studies on motivation have shown that extrinsic motivation (such as our messages) is less powerful than intrinsic motivation and that for extrinsic motivation to be internalized, a feeling of relatedness is important [23]. More work is needed to identify to what extent the *feeling of being heard* increases relatedness exactly. However, the result that the content of a motivational message can influence this variable is a promising first step to further increase the impact of personalized motivational messages.

Our system aims to improve motivational messages by personalizing them based on the type of situation the user was in. This seems to have worked as these personalized messages were more effective in terms of *motivation to continue* and *trust in a good outcome* in the situation with worsening PTSD symptoms in particular. This is an interesting result as it indicates that personalization is more effective in those situations where motivation is most important.

However, results also indicate that the effect of personalization can only be truly understood if we consider not just the symptoms of the user but also their initial trust levels. These results give a slightly more nuanced picture and do not directly fully support hypothesis 2. For situations with worsening symptoms, personalization improved motivation only when the *initial trust* level was doubtful or high. Therefore, in the situation where symptoms were getting worse and initial trust was low, personalization did not have as much effect. A closer look at the results (see Figure 14) reveals that in this situation, neither type of motivational message (general or personalized) led to a very large improvement in motivation in the first place. An effect between the types of messages was shown, but mostly because a message without motivation led to a decrease in motivation. These results could indicate that if symptoms are worsening and trust is low, a simple message by a virtual agent is not enough to truly improve motivation of the user. Still, these messages might prevent a further decrease in motivation, as was the case with the nonmotivational message. They would, therefore, nevertheless be useful, but should perhaps be supplemented with other types of motivation.

The situations where PTSD symptoms are reducing also show a clear impact of *initial trust level* on the effect of personalization. Although personalization improves motivation in situations with high trust, situations with low trust show an opposite effect, with general motivational messages outperforming personalization. This result is surprising, but 1 possible explanation might be that the personalized messages will point out that scores are dropping, which is already a good sign. If trust level is low, however, people might instead prefer to get reassurance that it is normal to find therapy difficult. Nearly all general motivational messages included this reassurance, whereas the personalized messages for this situation did not, as things were actually going well. This could mean that the messages our experts wrote for this situation were not quite as suitable as believed. Obviously, this result warrants more study into why, exactly, our general messages outperformed those based on the experts' answers.

Although the overall results, as well as the results from other situations, confirm hypothesis 2, this final result that in 1 situation general motivation outperforms personalized motivation suggests a rejection. A full rejection might, however, be too strong, as it might also be the case that our particular way of personalizing was not effective in this situation. Instead, it might be more fitting to argue that for this specific situation with a reduction of PTSD symptoms and low trust, general motivation is more suitable. To explain why this is the case would require more research.

Implications for Motivational Models

The evaluation of the personalized motivational messages shows that they are capable of increasing *motivation* and *trust* as well as the *feeling of being heard* by a virtual agent. However, it also shows that personalization is more effective in some situations than in others. These results indicate that it is worthwhile to take a closer look at the motivational model itself and what aspects might be improved in the future.

Impact of Situational Variable on Statement Type Choice

The motivational model used 2 input variables to describe user situation, namely, *initial trust* in a good therapy outcome and *PCL trend*. These variables describe aspects of current motivation and symptom progress, respectively. The analysis of our results reveals that to predict which statement types are best, both of these factors are necessary. Moreover, an interaction effect was found, which shows that they cannot be seen separately from each other. However, a closer inspection of the probabilities for each statement type shows that for some, either *PCL trend* or *initial trust* seems more important than the other [41].

Most noticeable is that the *note PCL* statement type has a very high probability for all situations, but whether the subtype is *note stable*, *note dropping*, or *note rising* is highly dependent on the *PCL trend* and does not seem to be influenced as much by *initial trust*. This makes sense, as this statement type directly refers to the *PCL trend*. For the *note trust* statement type, the story seems a little more complex. Whether trust is noted at all

XSL•FO
RenderX

seems dependent on both *initial trust* and *PCL trend*. However, whether *high trust* is noted or *low trust*, strongly depends *on initial trust level* only, which again makes sense as that is what it refers to. Another interesting *statement type* is that of general motivational comments, where overall, the *initial trust level* seems to influence whether such a statement is used, whereas *PCL trend* does not have much impact. This is probably because most such general statements were positive in nature, so they were compliments, or they confirmed that it was going well. *Initial trust level* seems to influence whether these are more suitable than symptom progression.

Improving the Model

Looking at the empirical results, our model shows promise, but also that there is room for improvement. First, there was 1 situation in which the general messages outperformed the personalized messages. This means that for this situation, the personalized system needs to be improved. However, to do this properly, the first step would be to examine why the general motivation worked better. It is unclear whether this was truly, as suggested, because people prefer reassurance when trust is low even if symptoms are actually reducing. If this is the case, the likelihood table would need to be adapted for this situation.

Besides this specific situation, general improvements could also be made. These improvements could be made at different stages of building the model. These stages can be seen in Figure 15. First, the model might be improved by providing more information about the situation. Adding variables such as stage in the therapy or personality might allow for a more fine-grained personalization. Another option would be to include more expert participants to create a larger corpus of sentences as well as making the likelihood table more stable. Our current model is based completely on the answers of these experts, so their role in building this model is large. Then, the categorization might be improved further as well. One option would be to let the experts rate the categorization or provide the statement types they believe they have used themselves. To improve the statistical analysis, 1 possible method would be to let the experts rate each other's answers and use these as weights. Another method for improvement would be to analyze the order in which motivational statement types are used. Our current system orders the statements similarly for every situation, but perhaps, our input variables could be used to personalize this order as well. This would possibly improve the way the motivational messages are generated. All these options are aimed to eventually improve the content motivational system.

Something that is increasingly receiving attention in the literature is the goal to learn about the preferences of the user over time. However, this requires a measure of the success of the motivational message as well as enough time to learn. Given our current application domain, this might not be completely suitable. The only measure of success would be whether someone drops out of therapy, but that would mean no more messages are provided in any case. Moreover, participants only participate in 12 sessions, so by the time a learned model becomes good enough, the therapy might be over. Given the severity of PTSD and the importance of therapy, it is preferable to start with motivational messages, which are optimized beforehand.

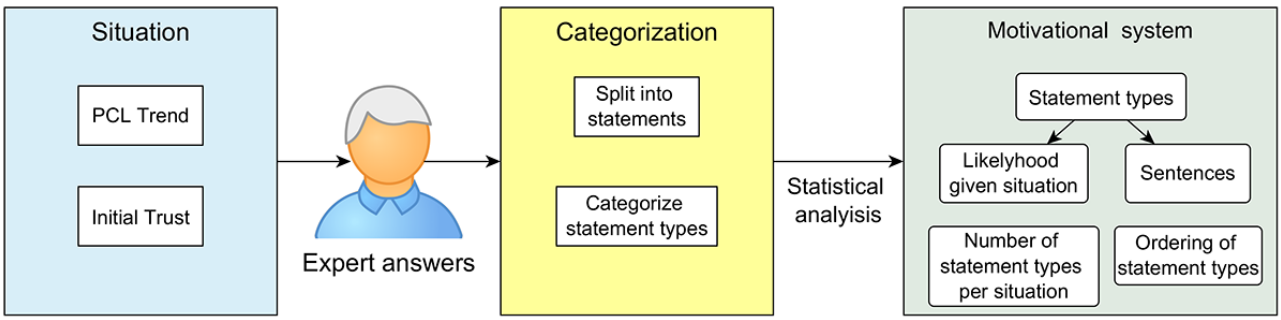
Limitations

To fully appreciate the system and results presented in this paper, it is important to also consider the limitations. Although the motivational message system is based on the patient's situation, it only considers 2 parameters, namely, *symptom* trend and *initial trust*. However, other factors such as interpersonal differences in personality and gender could, for instance, also be used [32]. Moreover, the system currently does not take into account the patient's stage in therapy, for instance, if they have just started or are near the end. However, this might also affect what statement types are most suitable to use in the motivational message [49]. Another limitation is that the data gathered from experts were all in Dutch, whereas the user study was performed in English. In both cases, this was the native language of our participants, and as Dutch and English are quite similar in structure, it was possible to make relatively direct translations. However, future work would be necessary to better study the effect of language on motivational effect. This ties in with possible cultural differences that might need to be reflected on in the content of the motivational messages.

In the user study design, all the participants were presented with all the different situations, something which would not happen in a real therapy. Responder bias might have caused participants to feel the need to reply differently to each situation, whereas some messages might actually have had the same effect. Although this might have caused more variance, that is, noise, it is unlikely to have caused any systematic bias as this was mitigated by randomization of the order in which participants were exposed to the situations. Another limitation is that the variables measured in this experiment relied on self-reports of the participant and responder bias might have, therefore, played a role. Moreover, they were asked to indicate what effect messages would have, given an imaginary situation. This imaginary aspect of our manipulation and measures could have affected our results. Similarly, the experiment was not performed with PTSD patients currently in therapy. However, we did not find indications that experience with exposure therapy influenced our results, which sheds a favorable light on the generalizability of the results. It is more difficult to say what the effect of the message system would be for a full therapy, as participants were only presented with situations in a single session. Further research would have to be done to study what results repeated motivation during the course of therapy would have. Such research could also shed light on the influence of a more interactive virtual agent. In the user study, video clips were used, but a more adaptive agent might further influence how motivation by this agent is perceived.

XSL•FO
RenderX

Figure 15. Different stages of building the motivational system. From the situation, we receive expert answers, which are then categorized, and then a statistical analysis follows, resulting in a system that matches statement types with situations using likelihood numbers and then chooses matching sentences. It also includes information on the number of statement types to include and the order in which to present them. PCL: post-traumatic stress disorder-check-list.



Contributions

The main scientific contribution of the work presented in this paper is the motivational message system and its empirical evaluation. The motivational messages generated by this system could be used in other applications for PTSD therapy. Moreover, with minimal rewording of the messages, it might also be usable for exposure therapy systems for other disorders. Although we would not recommend using the exact statements and likelihoods for other domains, the underlying development approach could be applied to many other health care systems as well. An implemented version of the model described in this paper can be found in the study by Tielman et al [50]. Our study shows that a method involving the analysis of expert knowledge can be used to produce motivational messages that work to motivate users. The use of such an approach could be of benefit to other systems that currently rely exclusively on theory due to lack of data [36].

Similar systems could be applied to many different interventions, as shown by the broad range of mental health applications using motivational statements [8,22,51]. Moreover, our system sheds light on exactly how to motivate patients during therapy. The different statement types reveal what topics domain experts would address and that the patient's personal situation has an effect on the type of language used. For instance, it was very common for experts to not just use purely motivational language but also to first explicitly refer to the user's situation. Many experts used sentences to explain that negative feelings were normal, but would eventually go away, indicating that it is important to refer to a positive future. Moreover, the exact type of motivational statement did differ per situation, with, for instance, compliments being more suitable when people trusted in a good outcome, regardless of their symptoms, whereas phrases such as "go on like this" were used much more when symptoms were decreasing and trust was also good. These insights could be used to better tailor motivational messages in other applications as well.

Conclusions

Google Assistant, Amazon Alexa, and Apple Siri are popular conversational agents that allow people to access a variety of services. Their capability and popularity seem to be growing. This paper shows how such agents could generate motivational messages personalized to the users' situation and successfully increase their trust in a good outcome and motivation to continue. Moreover, it outperforms general motivational messages in several situations as well as making the user feel more heard by the system. This successful application shows how domain knowledge can be gathered from experts to build smarter technology. Especially in domains such as mental health care, where data on how to give therapy are not always available, such methods can be valuable. Moreover, given the importance of motivation in mental health applications in general, a suitable motivational message system could have the potential to increase adherence and reduce dropouts. Moreover, we hope, this will eventually even improve the overall effectiveness of e-mental health therapy.

Acknowledgments

This study is part of the program Virtual e-Coaching and Storytelling Technology for Post-Traumatic Stress Disorder, which is financed by the Netherlands Organization for Scientific Research (pr. nr. 314-99-104).

Conflicts of Interest

None declared.

References

1. World Health Organization. 2012. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level URL: https://apps.who.int/iris/handle/10665/78898 [accessed 2019-02-07] [WebCite Cache ID 760Ro56uh]
2. Cartreine J, Ahern D, Locke S. A roadmap to computer-based psychotherapy in the United States. Harv Rev Psychiatry 2010;18(2):80-95. [doi: 10.3109/10673221003707702] [Medline: 20235773]

XSL•FO
RenderX

3. Donker T, Blankers M, Hedman E, Ljótsson B, Petrie K, Christensen H. Economic evaluations of Internet interventions for mental health: a systematic review. Psychol Med 2015 Dec;45(16):3357-3376. [doi: 10.1017/S0033291715001427] [Medline: 26235445]
4. Kasckow J, Zickmund S, Rotondi A, Mrkva A, Gurklis J, Chinman M, et al. Development of telehealth dialogues for monitoring suicidal patients with schizophrenia: consumer feedback. Community Ment Health J 2014 Apr;50(3):339-342. [doi: 10.1007/s10597-012-9589-8] [Medline: 23306676]
5. Kordy H, Backenstrass M, Hüsing J, Wolf M, Aulich K, Bürgy M, et al. Supportive monitoring and disease management through the internet: an internet-delivered intervention strategy for recurrent depression. Contemp Clin Trials 2013 Nov;36(2):327-337. [doi: 10.1016/j.cct.2013.08.005] [Medline: 23974036]
6. Kampmann I, Emmelkamp P, Hartanto D, Brinkman WP, Zijlstra B, Morina N. Exposure to virtual social interactions in the treatment of social anxiety disorder: a randomized controlled trial. Behav Res Ther 2016 Feb;77:147-156. [doi: 10.1016/j.brat.2015.12.016] [Medline: 26752328]
7. Yasavur U, Lisetti C, Rishe N. Let's talk! Speaking virtual counselor offers you a brief intervention. J Multimodal User Inter 2014;8:381-398. [doi: 10.1007/s12193-014-0169-9]
8. Hartanto D, Brinkman WP, Kampmann I, Morina N, Emmelkamp P, Neerincx M. Home-based virtual reality exposure therapy with virtual health agent support. Perva Comput Paradigm Mental Health 2015;204:85-98. [doi: 10.1007/978-3-319-32270-4_9]
9. Breso A, Martinez-Miranda J, Garcia-Gomez J. Leveraging adaptive sessions based on therapeutic empathy through a virtual agent. 2014 Presented at: ICAART - Doctoral Consortium; March 6-8, 2014; Angers, France.
10. Tielman M, Neerincx M, Bidarra R, Kybartas B, Brinkman WP. A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories. J Med Syst 2017;41:1-10. [doi: 10.1007/s10916-017-0771-y]
11. 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]
12. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. Washington DC, US: American Psychiatric Association; 2013.
13. Foa E, Hembree E, Olaslov RB. Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences Therapist Guide (Treatments That Work). New York: Oxford University Press; 2007.
14. Hofmann S, Otto MW. Cognitive Behavioral Therapy for Social Anxiety Disorder: Evidence-Based and Disorder-Specific Treatment Techniques. New York: Routledge; 2008.
15. Hundt N, Barrera T, Arney J, Stanley M. "It's Worth It in the End": Veterans' Experiences in Prolonged Exposure and Cognitive Processing Therapy. Cogn Behav Pract 2016;24:50-57. [doi: 10.1016/j.cbpra.2016.02.003]
16. Powers M, Halpern J, Ferenschak MP, Gillihan S, Foa E. A meta-analytic review of prolonged exposure for posttraumatic stress disorder. Clin Psychol Rev 2010 Aug;30(6):635-641. [doi: 10.1016/j.cpr.2010.04.007] [Medline: 20546985]
17. Provoost S, Lau H, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017 Dec 9;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]
18. Morandell MM, Hochgatterer A, Fagel S, Wassertheurer S. Avatars in assistive homes for the elderly a user-friendly way of interaction? In: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work. 2008 Presented at: USAB'08; November 20-21, 2008; Graz, Austria.
19. Zhou S, Bickmore T, Rubin A, Yeksigian C, Lippin-Foster R, Heilman M, et al. A Relational Agent for Alcohol Misuse Screening and Intervention in Primary Care. 2017 Presented at: CHI'17 Workshop on Interactive Systems in Healthcare (WISH); 2017; Denver, Colorado, USA URL: http://relationalagents.com/publications/WISH17-VA.pdf
20. Rizzo A, Lange B, Buckwalter J, Forbell E, Kim J, Sagae K, et al. An intelligent virtual human system for providing healthcare information and support. Stud Health Technol Inform 2011;163:503-509 [FREE Full text] [doi: 10.3233/978-1-60750-706-2-503] [Medline: 21335847]
21. Yee N, Bailenson J, Rickertsen K. A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2007 Presented at: CHI'07; April 28-May 3, 2007; San Jose, California, USA.
22. Bickmore T, Silliman RA, Nelson K, Cheng DM, Winter M, Henault L, et al. A randomized controlled trial of an automated exercise coach for older adults. J Am Geriatr Soc 2013 Oct;61(10):1676-1683. [doi: 10.1111/jgs.12449] [Medline: 24001030]
23. Ryan RM, Deci EL. A self-determination theory approach to psychotherapy: the motivational basis for effective change. Can Psychol 2008;49(3):186-193 [FREE Full text] [doi: 10.1037/a0012753]
24. Walter N, Ortbach K, Niehaves B. Designing electronic feedback – analyzing the effects of social presence on perceived feedback usefulness. Int J Hum Comput Stud 2015 Apr;76:1-11. [doi: 10.1016/j.ijhcs.2014.12.001]
25. Prochaska J, DiClemente C. The transtheoretical approach. In: Norcross JC, Goldfried MR, editors. Oxford Series in Clinical Psychology. Handbook of Psychotherapy Integration. New York, NY: Oxford University Press; 2005:147-171.
26. Miller W, Rollnick S. Motivational Interviewing: Preparing People to Change Addictive Behavior. New York: Guilford Press; 1991.

XSL•FO
RenderX

27. Friederichs S, Oenema A, Bolman C, Guyaux J, van Keulen HM, Lechner L. Motivational interviewing in a web-based physical activity intervention: questions and reflections. Health Promot Int 2015 Sep;30(3):803-815. [doi: 10.1093/heapro/dat069] [Medline: 24101160]
28. Blanson Henkemans OA, van der Boog PJ, Lindenberg J, van der Mast CA, Neerincx MA, Zwetsloot-Schonk BJ. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technol Health Care 2009;17(3):253-267. [doi: 10.3233/THC-2009-0545] [Medline: 19641261]
29. Imperato DC, Damus K, Hempstead M, Schulman D, Kopy M, Tran H, et al. Incorporating motivational interviewing into virtual contraceptive counseling. 2012 Presented at: American Public Health Association 2012 Annual Meeting & Expo; October 27-31, 2012; San Francisco.
30. Horsch CH. Delft University of Technology. 2016. A virtual sleepcoach for people suffering from insomnia URL: https://repository.tudelft.nl/islandora/object/uuid%3A8da2aedf-ac26-4763-9271-66c6981007fe [accessed 2019-02-12] [WebCite Cache ID 767isv9p6]
31. de Vries R, Truong K, Evers V. Crowd-designed motivation: Combining personality and the Transtheoretical Model. In: Proceedings of the 11th International Conference on Persuasive Technology - Volume 9638. 2016 Presented at: PERSUASIVE 2016; April 5-7, 2016; Salzburg, Austria.
32. de Vries R, Truong K, Zaga C, Li J, Evers V. A word of advice: how to tailor motivational text messages based on behavior change theory to personality and gender. Pers Ubiquit Comput 2017;21:675-687. [doi: 10.1007/s00779-017-1025-1]
33. Busch M, Mattheiss E, Reisinger M, Tscheligi M. More than Sex: The Role of Femininity and Masculinity in the Design of Personalized Persuasive Games. In: Proceedings of the 11th International Conference on Persuasive Technology - Volume 9638. 2006 Presented at: PERSUASIVE 2016; April 5-7, 2016; Eindhoven, The Netherlands.
34. Maddux J, Rogers R. Protection motivation and self-efficacy: a revised theory of fear appeals and attitude change. J Exp Soc Psychol 1983 Sep;19(5):469-479. [doi: 10.1016/0022-1031(83)90023-9]
35. Glanz K, Rimer B, Viswanath K. Health Behavior and Health Education. San Francisco: Jossey-Bass; 2008.
36. Grigore E, Pereira A, Scassellati B. Modeling Motivational States in Adaptive Robot Companions. 2015 Presented at: AAAI Fall Symposium Series; November 12–14, 2015; Arlington, Virginia, US URL: https://pdfs.semanticscholar.org/e86a/fc21fc9dec15914544c8338a19e969651194.pdf
37. Sussenbach L, Riether N, Schneider S, Berger I, Kummert F, Lutkebohle I. A robot as fitness companion: towards an interactive action-based motivation model. 2014 Presented at: The 23rd IEEE International Symposium on Robot and Human Interactive Communication; August 25-29, 2014; Edinburgh, UK.
38. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
39. Weathers F, Litz B, Keane T, Palmieri P, Marx B, Schnurr P. The PTSD Checklist for DSM-5 (PCL-5). 2013. PTSD: National Center for PTSD URL: https://www.ptsd.va.gov/ [accessed 2019-02-12] [WebCite Cache ID 767kbEKEZ]
40. Keller S, Zoellner L, Feeny N. Understanding factors associated with early therapeutic alliance in PTSD treatment: adherence, childhood sexual abuse history, and social support. J Consult Clin Psychol 2010 Dec;78(6):974-979 [FREE Full text] [doi: 10.1037/a0020758] [Medline: 20873895]
41. Tielman ML, Neerincx MA, Brinkman WP. 4TU Centre for Research Data. 2017. Dataset belonging to the paper: Generating personalized motivational messages for a virtual agent that assists in post-traumatic stress disorder therapy URL: https://data.4tu.nl/repository/uuid:3396c380-6102-4043-9693-847a3e482e03 [WebCite Cache ID 760a3Szjs]
42. Brinkman WP. Current and Future Research Directions in Mental Health Computing. 2013 Presented at: International Conference on Computation and Communication Advancement (IC3A)- 2013; January 11-13, 2013; West Bengal, India.
43. Baker T, Gustafson D, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res 2014 Feb 19;16(2):e36 [FREE Full text] [doi: 10.2196/jmir.2925] [Medline: 24554442]
44. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI'11; May 7-12, 2011; Vancouver, BC, Canada.
45. Kapp K. The Gamification of Learning Instruction. San Francisco: Pfeiffer; 2012.
46. Zandbelt LC, Smets EM, Oort FJ, Godfried MH, de Haes HC. Satisfaction with the outpatient encounter: a comparison of patients' and physicians' views. J Gen Intern Med 2004 Nov;19(11):1088-1095 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.30420.x] [Medline: 15566437]
47. Hall M, Zheng B, Dugan E, Camacho F, Kidd K, Mishra A, et al. Measuring patients' trust in their primary care providers. Med Care Res Rev 2002 Sep;59(3):293-318. [doi: 10.1177/1077558702059003004] [Medline: 12205830]
48. Tielman ML, Brinkman WP, Neerincx MA. Design guidelines for a virtual coach for post-traumatic stress disorder patients. 2014 Presented at: International Conference on Intelligent Virtual Agents; August 26-29, 2014; Boston, US.
49. Wang W. TU Delft Repositories. 2017. Self-management Support System for Renal Transplant Patients URL: https://repository.tudelft.nl/islandora/object/uuid%3A24516d5d-94ae-4cc6-a644-ed1b0343b8d6 [accessed 2019-02-12] [WebCite Cache ID 767mIqibS]

XSL•FO
RenderX

50. Tielman ML, Pasman W, Neerincx MA, Brinkman WP. Generating Motivatinonal Messages for PTSD Therapy. 2018. TUDelft URL: https://tracinsy.ewi.tudelft.nl/pubtrac/HealthPsychology/wiki/MotivateDuringTherapy [accessed 2019-02-07] [WebCite Cache ID 760apmjL3]
51. Beun RJ, Brinkman WP, Fitrianie S, Griffioen-Both F, Horsch C, Lancee J. Improving Adherence in Automated e-Coaching : A Case from Insomnia Therapy. In: Proceedings of the 11th International Conference on Persuasive Technology - Volume 9638. 2016 Presented at: PERSUASIVE 2016; April 5-7, 2016; Salzburg, Austria p. 276-287.
52. Tielman ML, Neerincx MA, Brinkman WP. Generating Situation-based Motivational Feedback in a Post-Traumatic Stress Disorder E-health System. 2017 Presented at: International Conference on Intelligent Virtual Agents; August 27-30,2017; Stockholm, Sweden URL: http://www.iva2017.org/papers/104980440.pdf
53. Thayaparan AJ, Mahdi E. The Patient Satisfaction Questionnaire Short Form (PSQ-18) as an adaptable, reliable, and validated tool for use in various settings. Med Educ Online 2013 Jul 23;18:21747 [FREE Full text] [doi: 10.3402/meo.v18i0.21747] [Medline: 23883565]
54. 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]
55. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco: Jossey-Bass Publishers; 2002:99-120.
56. Fogg B. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: Persuasive'09; April 26-29, 2009; Claremont, California, USA.
57. Andrade AD, Anam R, Karanam C, Downey P, Ruiz JG. An overactive bladder online self-management program with embedded avatars: a randomized controlled trial of efficacy. Urology 2015 Mar;85(3):561-567. [doi: 10.1016/j.urology.2014.11.017] [Medline: 25586473]
58. Akhtar S, Zhu Y. How transofrmational leadership influences follower helping behavior: the role of trust and procsocial motivation. J Organ Behav 2013;35:373-392. [doi: 10.1002/job.1884]
59. Tielman ML, Neerincx MA, Brinkman WP. Dataset belonging to the paper: Generating situation-based motivational feedback in a post-traumatic stress disorder e-health system. categories & probabilities of statements for motivating ptsd patients. 2017. TU Centre for Research Data URL: https://data.4tu.nl/repository/uuid:97c0c9c1-00ab-45da-8bff-c4e5dd6328d1 [accessed 2019-02-12] [WebCite Cache ID 767ozbCZT]
60. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann; 2003.
61. Wendel S. Designing for Behavior Change: Applying Psychology and Behavioral Economics. Sebastopol, US: O'Reilly Media; 2013.
62. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Info Syst 2009;24:28 [FREE Full text] [doi: 10.17705/1CAIS.02428]
63. Carroll J. Making Use: Scenario-Based Design of Human-Computer Interactions. Cambridge, US: MIT Press; 2000.
64. Cooper A. The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity. Indianapolis, US: Sams; 2004.
65. Michie S, West R, Brown J, Gainforth H. ABC of Behaviour Change Theories. UK: Silverback Publishing; 2014.
66. Ajzen I. The theory of planned behavior. Organisational behavior and human decision processes 1991;50:179-211 [FREE Full text] [doi: 10.1016/0749-5978(91)90020-T]
67. Rogers RW, Cacioppo JT, Petty RE. Cognitive and physiological processes in fear appeals and attitude change: a revised theory of protection motivation. In: Social Psychophysiology: A Sourcebook. New York: Guilford Press; 1983.

Abbreviations

ANOVA: analysis of variance
BCSS: behavior change support systems
eHealth: electronic health
e-mental: electronic-mental
MI: motivational interviewing
PCL: post-traumatic stress disorder-check-list
PTSD: post-traumatic stress disorder
TTM: transtheoretical model

XSL•FO
RenderX

Edited by K Eddens; submitted 23.10.17; peer-reviewed by RJ Beun, R de Vries; comments to author 29.11.17; revised version received 24.04.18; accepted 05.12.18; published 27.03.19.

Please cite as:
Tielman ML, Neerincx MA, Brinkman WP
Design and Evaluation of Personalized Motivational Messages by a Virtual Agent that Assists in Post-Traumatic Stress Disorder Therapy
J Med Internet Res 2019;21(3):e9240
URL: http://www.jmir.org/2019/3/e9240/
doi:10.2196/jmir.9240
PMID:30916660

©Myrthe L Tielman, Mark A Neerincx, Willem-Paul Brinkman. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between the Use of eHealth and Out-of-Hours Services in People With Type 1 Diabetes: Cross-Sectional Study

Anne Helen Hansen[1,2], BSW, MD, PhD; Tor Claudi[3], MD; Eirik Årsand[4,5], MSci, PhD

[1]Centre for Quality Improvement and Development, University Hospital of North Norway, Tromsø, Norway

[2]Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

[3]Department of Medicine, Nordland Hospital, Bodø, Norway

[4]Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway

[5]Department of Clinical Medicine, UiT The Arctic University of Norway, Tromsø, Norway

Corresponding Author:
Anne Helen Hansen, BSW, MD, PhD
Centre for Quality Improvement and Development
University Hospital of North Norway
PO Box 35
Tromsø, 9038
Norway
Phone: 47 91619655
Email: anne.helen.hansen@unn.no

Abstract

Background: Despite the increasing prevalence of diabetes and the increasing use of eHealth, little is known about the association between provider-based health services and eHealth among people with diabetes. This is the second study in a project exploring the associations between the use of eHealth and the use of provider-based health services.

Objective: The objective of this study was to investigate which eHealth services are used among out-of-hours (OOH) visitors with type 1 diabetes (T1D), and whether the use of eHealth (eg, apps, search engines, video services, and social media) was associated with the use of OOH services. We also wanted to investigate associations between anxiety, reassurance, and change in doctor-seeking behavior because of health information acquired from the Internet, and the use of OOH services.

Methods: We used data from a 2018 email survey of members of the Norwegian Diabetes Association (18-89 years old). Respondents with T1D were eligible for analyses. Using descriptive statistics, we estimated the use of OOH services and eHealth. Using logistic regressions, we studied the associations between the use of OOH services and the use of eHealth, as well as associations between the use of OOH services and reported consequences of using Internet-based health information.

Results: In the sample of 523 people with T1D (mean age 47 years), 26.7% (129/484) visited OOH services once or more during the previous year. Among the OOH visitors, search engines were used for health purposes by 86.7% (111/128), apps (health apps in general) by 63.6% (82/129), social media by 45.3% (58/128), and video services by 28.4% (36/127). The use of OOH services was positively associated with self-reported anxiety/depression (odds ratio [OR] 4.53, 95% CI 1.43-14.32) and with the use of apps (OR 1.73, 95% CI 1.05-2.85), but not with other types of eHealth. Those who had felt anxious based on information from the Internet were more likely to visit OOH services compared with those who had not felt anxious (OR 2.38, 95% CI 1.50-3.78). People who had decided to consult a doctor based on information from the Internet were more likely to visit OOH services (OR 2.76, 95% CI 1.64-4.66), compared to those who had not made such an Internet-based decision.

Conclusions: People with T1D were frequent users of OOH services, and the OOH visitors were frequent users of eHealth. The use of OOH services was positively associated with the use of health apps, with self-reported anxiety/depression, and with feeling anxious based on information from the Internet. Likewise, deciding to consult a doctor based on information from the Internet was positively associated with OOH visits. The use of eHealth seems to have a significant impact on people with T1D.

(J Med Internet Res 2019;21(3):e13465) doi:10.2196/13465

KEYWORDS

eHealth; internet; health care utilization; out-of-hours services; cross-sectional study; diabetes mellitus, type 1; Norway

XSL•FO
RenderX

Introduction

Increasing Prevalence of Diabetes

The prevalence of diabetes is increasing worldwide. Global prevalence in adults is estimated at 8.8% [1] and the Norwegian prevalence at 4.7% [2]. Around 245,000 people are diagnosed with diabetes in Norway. Of these, around 28,000 have type 1 diabetes (T1D) [2]. Patients represent a large proportion of health care contacts, and diabetes is a considerable burden on patients regarding morbidity and mortality [3]. Most patients do not reach the combined national treatment targets for prevention of complications [4-6].

Consequences of Increasing Use of eHealth Services

Electronic health (eHealth) is defined as "the transfer of health resources and health care by electronic means" [7]. Most Norwegian households (98%) had Internet access and access to a smart phone in 2017 [8,9]. Use of the Internet for health purposes has increased rapidly in the past decades [10-12], and 75%-80% of Internet users in the United States and Europe report conducting health-related searches [10-13]. Our first study in this project found that 87% of Norwegians with T1D used eHealth in one or more forms, exceeding the use of about 78% in the general population [14,15]. Although T1D is a prevalent chronic and challenging disease, the consequences of patients seeking health information from the Internet have not been comprehensively explored. Medlock et al studied the consequences of health information seeking among seniors in the Netherlands, finding that 38% had felt anxiety and 56% had felt reassured. Some had changed their plans regarding doctor visits: 48% had decided to go to the doctor, and 24% had decided *not* to go based on health information from the Internet [16].

Norwegian Health Care and Use of Out-of-Hours Services

Norwegian health care is based on universal insurance. However, consultations for adults are co-paid through a small fee [14]. Primary health care, including a regular general practitioner (GP) and GP-based out-of-hours (OOH) services, is provided to all residents by the municipalities through the patient list system. This includes almost all Norwegian residents [17]. Specialist services are operated by regional and local health enterprises owned by the national government. Access to specialist care is usually through referral from GPs or OOH services (the gatekeeper role). People with T1D are recommended to visit specialist services at least annually. There are indications that many patients miss this check-up [14]. We know little about the extent to which people with T1D visit Norwegian OOH services, but we know that people with any type of diabetes and people with chronic disease constitute a significant proportion of visitors [18]. Regarding the general population, 27.4% of listed patients contacted OOH services in 2008 (all contact types) [19], and 17% of the population visited OOH services in 2017 [20]. Norway has a high OOH contact rate compared to other countries, and 75% of contacts have been classified as non-urgent [21].

Associations Between Use of eHealth and Use of Out-of-Hours Services

Research on the association between the use of eHealth and provider-based health care is scarce [22,23]. However, frequent users of health services seem more likely to seek health information on the Internet compared with non-users [12]. This study is a part of the DIAcare project, aiming to investigate associations between eHealth and different provider-based health care services. The project also aims to investigate the associations between socioeconomic status and the use of eHealth, whether eHealth information is discussed in the clinical encounter, and whether the use of eHealth might lead to (or prevent) doctor visits [24]. Our first study in this project revealed no association between the use of eHealth and GP visits for T1D patients, whereas we found a positive association between the use of search engines and somatic specialist services [14]. This second study focuses on the association between the use of eHealth and the use of OOH services in persons with T1D. Understanding the associations between eHealth and the use of provider-based health services is important for patients as well as for health care providers, policy makers, and society, in order to enable evidence-based planning for future health care services in a society where eHealth is increasingly used.

Aim

The aim of this study was to investigate which eHealth services were used among OOH visitors with T1D, and the associations between the use of eHealth (ie, health apps in general, search engines, video services, and social media) and the use of OOH services. Furthermore, we aimed to investigate whether anxiety, reassurance, and change of doctor-seeking behavior because of the use of eHealth were associated with the use of OOH services.

Methods

Data

This cross-sectional study used data obtained in 2018 from members of the Norwegian Diabetes Association (NDA). As of December 31, 2017, the organization had 33,908 members, of which about 30% have T1D [25]. The Norwegian Social Science Data Service (NSD) Web survey distributed the invitations to a randomly selected sample of 5971 individuals who had their email addresses recorded by NDA. Initially, we planned to use data from the seventh Tromsø Study, conducted in 2015/2016 [24]. However, the Tromsø Study could not give us access, and we developed a tailored questionnaire based on the specific objectives of our study, using relevant questions from other published surveys [26,27].

Information about the study was posted together with the invitation. The questionnaire (Multimedia Appendix 1) included questions about health status including specific questions about disease duration, severity and treatment of diabetes, use of and experiences with eHealth and health care services, as well as demographic and socioeconomic information. We reviewed and tested the questionnaire several times before distribution to the informants. Non-respondents were given one reminder, submitted by email 15 days after the first request.

XSL•FO
RenderX

Participants

The respondents could not fill in the questionnaire more than once. Starting from 1250 participants, we first excluded the 66 individuals who did not suffer from diabetes themselves (eg, family members, health personnel, and others). We also excluded those who failed to respond to most of the questions (n=5), those who did not give information about gender (n=93), and participants with diabetes types other than T1D (n=563). The analyzed sample consisted of 523 respondents with T1D. For a flowchart of the study population, we refer to our previous study [14].

Variables

The dependent variable in all analyses was the use of OOH services once or more during the previous 12 months. The use of eHealth was subdivided into 4 variables: apps for mobile phone or tablet computers (health apps in general, not necessarily diabetes self-care apps), search engines (eg, Google), social media (eg, Facebook), and video services (eg, YouTube). We dichotomized these variables by merging the original four response options into "never or once" and "sometimes or often," in line with previous research [16].

The following questions were also asked: "Based on information from the Internet, have you felt anxious/felt reassured/decided to consult a doctor when you would otherwise *not* have consulted one/decided *not* to consult a doctor when you would otherwise have consulted one?" The answering options "no, once, sometimes, or often" were merged into "no" and "once, sometimes or often" for an easier interpretation of logistic regressions. Since we were interested in whether changes in decisions had ever taken place, we placed "no" in a separate category.

Age was grouped in 20-year age groups. The education categories were labeled low (primary/part of secondary school), middle (high school), high (college/university <4 years), and highest (college/university 4 years or more). Response options for self-rated health were excellent, good, fair, bad, and very bad. We merged the bad and very bad categories due to low numbers in the very bad category (4 respondents). Response options for self-reported degree of anxiety/depression were none, slight, moderate, severe, and extreme. We merged the severe and extreme categories due to only 1 respondent in the extreme category.

Since the use of health apps in general turned out to be an issue in this study, we wanted to inform the readers about the answers obtained from the following question in the questionnaire: "During the past 12 months, have you used apps for smartphone or tablet computer for follow-up of your own diabetes?" The answering options were "never, less than once a month, once a month, once a week, and every day." According to the study protocol, this variable was not included in the regression analyses.

Due to the relatively low participation rate, we constructed a dichotomous response time variable in order to compare late respondents with early respondents, assuming that the late respondents were more similar to non-respondents [28]. All participants who responded initially were placed in one group, and those who responded after the reminder were placed in the other group.

Analyses

Data were analyzed by means of descriptive statistics and logistic regressions. Correlations were tested with Spearman correlation test.

Use of OOH services was the dependent variable in all the analyses. In the first multivariate regression, the independent variables were the use of apps (general health apps), search engines, social media, and video services, gender, age, education, self-rated health, and self-reported degree of anxiety/depression. For the second set of analyses, we performed four univariate and four multivariate regression analyses. The independent variables in the four univariate analyses were "felt anxious," "felt reassured," "decided to consult a doctor when you would otherwise *not* have consulted one," and "decided *not* to consult a doctor when you would otherwise have consulted one." These were key independent variables in the multivariate analyses, which we adjusted for gender, age, education, and self-rated health. All the independent variables were introduced collectively into the multivariate models.

We compared those who did not respond to our initial inquiry, but eventually consented, with the early respondents. This was done by subsequently introducing the response time variable into the regression models. We used 95% confidence intervals (CI) throughout the study. All analyses were accomplished using Stata, version 14.2.

Ethics

The Regional Committee for Medical and Health Research Ethics (REK) found that an application for this project was not required according to the Norwegian Health Research Act (ref 2015/1779/REK nord). The data protection officer (Personvernombudet) at the University Hospital of North-Norway approved the study (ref. 2017/6579). The NSD data bureau received no information about the participants other than the email addresses.

Results

Participation

In total, 1250 persons aged 18-89 years answered the questionnaire, constituting a minimum response rate of 20.9%. We assume the real response rate to be higher since we had more than 400 bounce backs from email servers unable to deliver the invitation, and we do not know how many actually received the survey email. Eligible for analysis were the 523 persons who reported to have T1D.

Characteristics of Users of Out-of-Hours Services

Among the users of OOH services, the largest groups were men (66/129, 51.2%), people aged 40-59 years (53/129, 41.1%), married/cohabitants (87/99, 87.9%), people employed full-time or part-time (81/129, 62.8%), people with the highest education (41/128, 32.0%), the highest household income (63/126, 50.0%), good self-rated health (59/128, 46.1%), good self-rated regulation of diabetes (66/127, 51.9%), no anxiety/depression

XSL•FO
RenderX

(71/128, 55.5%), and who had lived with diabetes for 30 years or more (43/129, 33.3%) (Table 1).

Based on information from the Internet, 44.5% (57/128) of OOH users reported that they had felt anxious, 55.9% (71/127) had felt reassured, 28.7% (37/129) had decided to consult a doctor when they would otherwise not have consulted one, whereas 21.9% (28/128) had decided *not* to consult a doctor when they would otherwise have consulted one (Table 1).

Mean age among the users of OOH services was 47.2 years, 44.3 years for women, and 50.0 years for men. Median age was 48 years. Mean disease duration was 21.9 years (median 20 years).

Use of Out-of-Hours Services and eHealth

During the previous year, 26.7% (129/484) visited OOH services once or more. Men visited slightly more than women (29.1% versus 24.5%), and people aged 60 years and over visited slightly more than younger age groups (Table 2).

Among the OOH users, 63.0% (80/127) visited once and 26.8% (34/127) visited twice during the previous year. Only 2.4% (3/127) visited 5 times or more.

In the total sample as well as among the OOH users, search engines were the most widely used form of eHealth, followed by the use of health apps (all kinds). Among the OOH users, search engines were used by 86.7% (111/128), health apps (all kinds) by 63.6% (82/129), social media by 45.3% (58/128), and video services by 28.4% (36/127). Users of OOH services used apps slightly more than the total sample (Table 3).

Among the OOH users who reported having used health apps (all kinds) sometimes or often, the largest group (32/82, 39%) had never used apps for self-care of their diabetes, and the second largest group (28/82, 34%) had used apps for self-care less than once a month. Only 11% (9/82) had used apps for diabetes self-care every day (Table 4).

Associations Between Use of Out-of-Hours Services and eHealth

Visits to OOH services were positively associated with the use of health apps (odds ratio [OR] 1.73, 95% CI 1.05-2.85), but not associated with the use of any other type of eHealth. OOH services visits were positively associated with self-reported moderate or severe anxiety/depression, compared with no anxiety/depression (OR 2.18, CI 1.04-4.54, and OR 4.53, CI 1.43-14.32, respectively). Gender, age, education, and self-rated health were not associated with the use of OOH services (Table 5).

XSL•FO
RenderX

Table 1. Characteristics of total T1D sample and sample using out-of-hours services once or more during the previous 12 months.

Characteristics	Total sample % (n/N)	Users of OOH services % (n/N)
Gender		
Female	53.7 (281/523)	48.8 (63/129)
Male	46.3 (242/523)	51.2 (66/129)
Age		
18-39 years	34.0 (178/523)	32.5 (42/129)
40-59 years	42.6 (223/523)	41.1 (53/129)
60+ years	23.4 (122/523)	26.4 (34/129)
Marital status		
Single	11.0 (42/380)	12.1 (12/9)
Married/cohabitant	89.0 (338/380)	87.9 (87/99)
Main daily activity		
Working[a]	64.0 (308/481)	62.8 (81/129)
Pensioner, old age	13.5 (65/481)	14.0 (18/129)
Pensioner, disability	11.0 (53/481)	10.1 (13/129)
Pupil/student	7.3 (35/481)	7.0 (9/129)
Other	4.2 (20/481)	6.1 (8/129)
Education[b]		
Low	8.1 (39/480)	11.0 (14/128)
Middle	29.0 (139/480)	28.9 (37/128)
High	31.7 (152/480)	28.1 (36/128)
Highest	31.2 (150/480)	32.0 (41/128)
Household income[c]		
Low	14.1 (66/467)	16.7 (21/126)
Middle	34.9 (163/467)	33.3 (42/126)
High	51.0 (238/467)	50.0 (63/126)
Duration of diabetes		
<10 years	24.3 (127/522)	27.9 (36/129)
10-19 years	20.5 (107/522)	20.2 (26/129)
20-29 years	19.4 (101/522)	18.6 (24/129)
30 years and over	35.8 (187/522)	33.3 (43/129)
Self-rated regulation of diabetes		
Excellent	19.4 (101/520)	20.5 (26/127)
Good	56.2 (292/520)	51.9 (66/127)
Fair	19.8 (103/520)	20.5 (26/127)
Bad/very bad	4.6 (24/520)	7.1 (9/127)
Self-rated health		
Excellent	17.9 (93/521)	14.1 (18/128)
Good	51.6 (269/521)	46.1 (59/128)
Fair	21.7 (113/521)	28.9 (37/128)
Bad/very bad	8.8 (46/521)	10.9 (14/128)

XSL•FO
RenderX

Characteristics	Total sample % (n/N)	Users of OOH services % (n/N)
Degree of anxiety/depression		
None	65.0 (334/514)	55.5 (71/128)
Slight	22.2 (114/514)	21.9 (28/128)
Moderate	9.5 (49/514)	16.4 (21/128)
Severe	3.3 (17/514)	6.2 (8/128)
Based on information from the Internet, have you:		
Felt anxious (yes, once/sometimes/often)	31.8 (155/487)	44.5 (57/128)
Felt reassured (yes, once/sometimes/often)	54.6 (263/482)	55.9 (71/127)
Decided to consult a doctor when you would otherwise not have consulted one (yes, once/sometimes/often)	17.5 (87/497)	28.7 (37/129)
Decided not to consult a doctor when you would otherwise have consulted one (yes, once/sometimes/often)	18.3 (90/493)	21.9 (28/128)

[a]Full-time or part-time.

[b]Low (primary/part of secondary school), middle (high school), high (college/university <4 years), highest (college/university ≥4 years).

[c]Low (350,000 NOK or less), Middle (351,000-750,000 NOK), High (751,000 NOK or more).

Table 2. Proportion using out-of-hours services once or more during the previous 12 months, according to gender and age.

Characteristics	n/N	%	95% CI
Total sample	129/484	26.7	22.7-30.6
By gender			
Female	63/257	24.5	19.2-29.8
Male	66/227	29.1	23.5-35.3
By age, years			
18-39	42/156	26.9	19.9-33.9
40-59	53/211	25.1	19.2-31.0
60+	34/117	29.1	20.8-37.3

Table 3. Proportion using different kinds of eHealth sometimes or often during the previous 12 months.

Variables	Total sample			Users of out-of-hours services		
	n/N	%	95% CI	n/N	%	95% CI
Health apps (all kinds)	285/514	55.5	51.1-59.7	82/129	63.6	54.8-71.5
Search engines	431/513	84.0	80.6-86.9	111/128	86.7	79.6-91.6
Social media	232/513	45.2	40.9-49.6	58/128	45.3	36.8-54.1
Video services	118/506	23.3	19.8-27.2	36/127	28.4	21.1-36.9

Table 4. Proportion using apps for self-care of diabetes during the previous 12 months.

Frequency	Total sample (N=509)			Users of out-of-hours services (N=127)			Users of out-of-hours services and apps (N=82)		
	n	%	95% CI	n	%	95% CI	n	%	95% CI
Never	282	55.4	51.0-59.7	62	48.8	40.1-57.6	32	39.0	28.9-50.2
<1x per month	135	26.5	22.9-30.5	36	28.4	21.1-36.9	28	34.1	24.5-45.2
1x per month	41	8.1	6.0-10.8	10	7.9	4.3-14.1	9	11.0	5.7-20.0
1x per week	21	4.1	2.7-6.3	7	5.5	2.6-11.2	4	4.9	1.8-12.5
Every day	30	5.9	4.1-8.3	12	9.4	5.4-16.0	9	11.0	5.7-20.0

XSL•FO
RenderX

Table 5. Probability of using out-of-hours services once or more during the previous 12 months in a population with diabetes type 1 (multivariate logistic regressions).

Variables	Use of out-of-hours services (yes/no)		
	OR	*P*	95% CI
Health apps (all kinds)			
Never/once[a]	1.00		
Sometimes/often	*1.73* [b]	*.03* [b]	*1.05-2.85* [b]
Search engines			
Never/once[a]	1.00		
Sometimes/often	1.17	.67	0.57-2.40
Social media			
Never/once[a]	1.00		
Sometimes/often	0.76	.29	0.46-1.26
Video services			
Never/once[a]	1.00		
Sometimes/often	1.23	.45	0.72-2.10
Gender			
Female[a]	1.00		
Male	1.26	.31	0.81-1.96
Age			
18-39[a]	1.00		
40-59	0.88	.63	0.53-1.47
60+	1.32	.36	0.73-2.36
Education[b]			
Low[a]	1.00		
Middle	0.71	.44	0.31-1.65
High	0.65	.32	0.28-1.51
Highest	0.82	.65	0.35-1.91
Self-rated health			
Excellent[a]	1.00		
Good	1.15	.66	0.62-2.15
Fair	1.53	.25	0.75-3.15
Bad/very bad	1.26	.63	0.50-3.16
Self-reported anxiety/depression			
None	1.00		
Slight	1.13	.65	0.66-1.96
Moderate	2.18	.04	1.04-4.54
Severe	*4.53* [b]	*.01* [b]	*1.43-14.32* [b]

[a]Reference groups.

[b]Low (primary/part of secondary school), Middle (high school), High (college/university < 4 years), Highest (college/university 4 years or more.

XSL•FO
RenderX

Table 6. Probability of using out-of-hours services once or more during the previous 12 months according to reported effects of using the Internet for health information (univariate and multivariate logistic regressions).

Variables	Use of out-of-hours services (univariate logistic regressions)			Use of out-of-hours services (multivariate logistic regressions)[a]		
	OR	*P*	95% CI	OR	*P*	95% CI
Based on information from the Internet, have you:						
Felt anxious[b]	*2.13* [c]	*<.001* [c]	*1.40-3.25* [c]	*2.38* [c]	*<.001* [c]	*1.50-3.78* [c]
Felt reassured[b]	1.04	.83	0.70-1.58	1.21	.39	0.79-1.86
Decided to consult a doctor, when you would otherwise not have consulted one[b]	*2.7 0*[c]	*<.001* [c]	*1.65-4.42* [c]	*2.76* [c]	*<.001* [c]	*1.64-4.66* [c]
Decided not to consult a doctor, when you would otherwise have consulted one[b]	1.34	.25	0.81-2.21	1.33	.29	0.79-2.27

Statistically significant findings are marked in italics.

[a]adjusted for gender, age, education, and self-rated health.

[b]0=no (reference), 1=once, sometimes, or often.

Associations Between Use of Out-of-Hours Services and Reported Effects of Using the Internet for Health Information

Those who had felt anxious based on information from the Internet were more than twice as likely to visit OOH services compared with those who had not felt anxious, both in univariate (OR 2.13, CI 1.40-3.25) and in multivariate regression models (OR 2.38, CI 1.50-3.78). People who had decided to consult a doctor based on Internet information (when they would otherwise *not* have consulted one) were almost three times as likely to visit OOH services (OR 2.70, CI 1.65-4.42, and OR 2.76, CI 1.64-4.66, for univariate and multivariate regressions, respectively), compared with those who had not changed their decision to consult a doctor.

Feeling reassured based on information from the Internet was not associated with the use of OOH services. Likewise, we found no association between deciding *not* to consult a doctor (when they would otherwise have consulted one) and the use of OOH services (Table 6).

All findings presented in Tables 5 and 6 persisted after including the response time variable in the regression analyses. There were no strong correlations (defined as Spearman rho >.5) between the independent variables in any of the models.

Discussion

Principal Findings

We found that 26.7% (129/484) of people with T1D visited OOH services once or more during the previous year. Search engines were used sometimes or often by 86.7% (111/128) of the OOH visitors, apps by 63.6% (82/129), social media by 45.3% (58/128), and video services by 28.4% (36/127). Visits to OOH services were positively associated with self-reported anxiety/depression and with the use of health apps, but not with other forms of eHealth. Among those who had used OOH services as well as health apps during the previous year, 39% (32/82) had never used apps for self-care of their diabetes. Those who had felt anxious based on information from the Internet were more than twice as likely to visit OOH services compared with those who had not felt anxious. People who had decided to consult a doctor based on information from the Internet (when they would otherwise not have consulted one) were almost three times as likely to visit OOH services compared with those who had not made a decision to consult a doctor based on information from the Internet.

Extensive Use of Out-of-Hours Services by People With T1D

Among people with T1D, 26.7% reported one or more visit to OOH services during the previous year. Sandvik et al reported 27.4 contacts with OOH services per 100 list patients in Norway (general population, 2008) [19]. Unlike our rate, this rate includes more than just visits (ie, house calls, telephone contacts, and simple contacts), indicating a higher visit rate among people with T1D compared with the general population. The National Centre for Emergency Primary Health Care reported 1,332,024 visits to OOH services in Norway in 2017 [29], indicating a rate of 25.3% in a population of 5,258,317 inhabitants (January 1, 2017). Children, adolescents, and tourists were included, as well as the number of visits for those who visited more than once, meaning that our rate solely for visits is higher. Statistics Norway found that 17.0% of the population (all ages, children included) consulted OOH services in 2017 [20]. Since children and adolescents are frequent OOH users [29], the adult general population rate is probably lower. This is in line with data collected by the Tromsø Study, suggesting a preliminary general population OOH visit rate of 13.5% (2740/20,294) among people aged 40 years and over [27]. Despite constraints regarding comparison with these figures, we conclude that our rate for OOH visits among people with T1D is higher than these general population rates. Our findings contribute to underpinning previous documentation that people with diabetes are frequent users of OOH services in Norway [18,29]. Research from other countries confirms that this is an international phenomenon. Diabetes was the most frequent chronic somatic disorder among frequent attenders in the Netherlands in 2007 [30], and patients with diabetes accounted for almost 1 out of 10 emergency department visits in the United States (2010) [31].

XSL•FO
RenderX

Health Apps Widely Used by Attenders of Out-of-Hours Services

In the first paper in this project, we documented that people with T1D are frequent users of all the investigated four forms of eHealth [14]. OOH attenders with T1D used all forms of eHealth equally or more than the total T1D sample, and the largest difference was found for apps (Table 3). This extensive use is in line with the illness behavior model stating that people in poorer health more likely seek Web-based disease-related information [32]. It is also in line with a German study reporting that app users were more likely to report chronic conditions, such as diabetes [33].

Low Daily Use of Diabetes Self-Care Apps

Most of the combined OOH and app users had used apps for self-care of their diabetes; however, only 11.0% of them used diabetes self-care apps every day (Table 4). Only 5.9% of the total T1D sample used diabetes self-care apps every day, which is a low percentage compared to 24% in adults with T1D reported in a recent Australian study [34]. They found that the use of self-care apps was positively associated with female gender and lower age. More female participants and a lower mean age in the Australian study may explain some of the differences from our rate. However, the study methodology and question wording vary, and the figures are not directly comparable.

Positive Associations Between Out-of-Hours Services Visits and Use of Health Apps and Self-Reported Anxiety or Depression

The possibility of visiting OOH services was higher for persons with T1D who used health apps (all kinds), compared with those who did not (Table 5). We found no other studies investigating this relation.

Moderate and severe self-reported anxiety/depression was more prevalent among the users of OOH services, compared with the total sample (Table 1). Severe anxiety/depression increased the possibility of visiting OOH services 4-5 times, compared with those who reported no anxiety or depression (Table 5). Extensive use of OOH services among people with mental problems and psychiatric disease is well documented in previous research [18,30,35,36].

In many countries, Norway included, there is concern about unnecessary use of emergency medical services [19]. Patients, as well as GPs, consider worry an important reason for contact [37,38]. From a medical perspective, most OOH use is non-urgent, and motives for contacting OOH services are primarily patient-related [21,38]. Self-reported anxiety/depression and the use of health apps might be such patient-related factors. A common feature of apps, OOH services, and anxiety/depression is that they operate independently of opening hours. We consider this a possible explanation of the association between the use of health apps and the use of OOH services, as well as the association between anxiety/depression and the use of OOH services. Unsatisfactory design and functionality of many apps might also partly explain this finding [39]. Furthermore, a lack of GP availability or accessibility during daytime might contribute in some regions [19,38,40].

The extensive use of OOH services among people with T1D, and particularly among those reporting anxiety/depression, indicates that some patients do visit OOH services rather than waiting for a consultation with the regular physician. This might indicate that policy makers and GPs should consider increasing the regular GPs' capacity and/or extending opening hours. In this regard, eHealth consultations could be a valuable supplement to face-to face consultations.

Many studies regarding mobile apps and worries, anxiety, or emotional distress focus on apps as possible tools in coping with these problems [41]. However, some have studied the potential of digital devices to create or intensify worries and anxiety in susceptible individuals. Recent research found that the Internet has the potential to reduce as well as exacerbate health anxiety [42-44] and that individuals with moderate to high levels of health anxiety experienced more anxiety during and after online symptom checking, whereas those with low illness anxiety experienced relief [44]. Thus, the use of eHealth might as a side effect intensify health anxiety and, used in the wrong context, cause harm instead of benefit to health. If this is the case, it might contribute to explaining the positive association between the use of health apps and the use of OOH services.

Many apps used for monitoring of disease may meet the definition of a medical device [45]. Still, most apps have not been evaluated by authorization authorities and are not CE-marked (Conformité Européenne, European Conformity) according to European Union directives, or regulated by the US Food and Drug Administration [45,46]. Possible risks related to the app itself have been focused [39]; however, little is known about the use of apps outside tightly controlled research settings [34]. A wider focus on effectiveness related to the use by different individuals in different settings has not been addressed thoroughly [42,43,47]. In our opinion, such possible side effects should be evaluated for all equipment considered as medical devices, apps included.

Anxiousness and Change of Health Care Seeking Behavior Based on Internet Information

We found that 31.8% (115/487) of people with T1D and 44.5% (57/128) of the OOH users had felt anxious based on information from the Internet (Table 1). According to data collected by the Tromsø Study (general population), 25.8% had felt anxious [27], whereas 38.0% among elderly people in the Netherlands had felt anxious based on information from the Internet [16]. Our figures for people with T1D are placed between them, whereas figures for the OOH users exceeds the other rates.

The differences between the T1D sample and the OOH users among them were greater for feeling anxious (31.8% vs 44.5%) than for feeling reassured (54.6% vs 55.9%), which might indicate that using the Internet for health purposes could have more negative effects for the OOH users, compared with the T1D sample. Even if many people reported anxiety, more people had felt reassured based on Internet information in our study,

XSL•FO
RenderX

as well as in the Tromsø Study and in the Dutch study (47.7% and 56.0%, respectively) [16,27].

We found that 17.5% (87/497) of the total T1D sample and 28.7% (37/129) of the OOH users among them decided to consult a doctor based on Internet information—when they would otherwise not have consulted one (Table 1). The similar rate from the Tromsø Study data was 23.6% [27]. Again, our rate for the OOH users exceeds the rates for the total T1D sample, as well as the general population rate obtained from the Tromsø Study. This might indicate that information from the Internet has a larger influence on the OOH users in this regard. Among elderly people in the Netherlands, almost half of the participants (48%) had decided to consult a doctor based on Internet information, which might be explained by cultural diversities, differences in the use and in trusting information from the Internet, as well as methodological differences [16].

Those who had felt anxious based on information from the Internet were 2-3 times more likely to visit OOH services than those who had not felt anxious. People who had decided to consult a doctor based on information from the Internet—when they would otherwise not have consulted one—were almost three times as likely to visit OOH services compared with those who had not changed their decision regarding consulting a doctor. A recent systematic review and meta-analyses found a positive correlation between health anxiety and online health information seeking [48]. As health anxiety levels increase, the relationship between health information seeking online and visiting a doctor based on information found online also increases [42,49]. Other studies asked whether use of the Internet had changed the frequency of doctor visits. Around 90% in these general population studies reported that use of the Internet for health purposes did not change their health care seeking behavior (United States 94%, Japan 88.9%, France 88.6%) [50-52]. Consequently, around 10% did change their health care seeking behavior, as they made either more visits or fewer visits due to information from the Internet. However, other available studies in different populations, most of them general populations, cannot be directly compared with the current study of people with T1D. Our findings may suggest that when decisions to consult a doctor are based on Internet information, people tend to consult as soon as possible. Since OOH services are available at any time, they might be a natural choice.

Strengths and Limitations

Strengths and limitations have been explored in detail in our first study in this project [14]. The strengths of this study are similar to the strengths discussed in the first paper in this project [14]. The most important strength is the focus on a scarcely investigated research field, which might contribute to evidence-based planning for future health care services in a society where eHealth is increasingly used. Other strengths are the detailed questionnaire specifically tailored to people with diabetes, the recruitment of participants from all of Norway, the inclusion of a wide age span of participants, and that we were able to analyze the data shortly after they were collected. Finally, the collection of data in cooperation with NDA enabled us to develop excellent user participation with a large and important group of health care users.

A limitation of this study is the low estimated participation rate. However, response rate must not be confused with response quality [28]. We found that older people dominated among the late respondents compared with the early respondents [14]. As late respondents might be more similar to non-respondents, younger individuals might be overrepresented.

Distribution of the questionnaire by email is another limitation, which excluded those who do not use the Internet or do not have an email address. Since 97% of Norwegian households have Internet access, we do not think that this affected our results significantly [14]. It is well known that women, healthier persons, higher socioeconomic groups, and middle-aged people are more likely to participate in surveys [14]. This suggests that women, people around 40 to 80 years, people in better health, and higher socioeconomic groups might be overrepresented in our study, thus tending to level out a possible skewness in the opposite directions.

Other relevant limitations were recall bias, the validity of self-reported data, and the cross-sectional study design, as reported and discussed in detail in the first study in the DIAcare project [14]. It is not possible to judge the magnitude or direction of a possible non-response bias, since different factors might pull the tendency in different directions or level each other out. The low response rate is in itself not an indication of low representativeness, as non-response bias may be a problem even if response rates are high [53]. We suggested that non-response bias posed a limited threat to our study's validity; however, generalization must be made with caution.

Increasing travel distance is associated with reduced use of OOH services in Norway [54]. The lack of information about travel distance is thus a study limitation. However, we have no reasons to believe that the use of apps or other of the independent variables are influenced by travel distance and do not consider this a confounding threat to our results.

It should also be mentioned that worries/anxiety/depression/emotional distress are not defined according to diagnostic manuals in this study and rely solely on self-report. We consider self-report to be interesting as such in this field and do not think that this has disturbed the validity of our results.

Future Research

This study investigated the use of health apps in general and associations with the use of OOH services. For future research, the more specific use of diabetes self-care apps and associations with the use of different health care services would be beneficial. Nor did this study investigate the reasons for visiting OOH services, which would be interesting as well: were visits directly related to the use of eHealth for diabetes self-care, or were they due to other health care needs? Furthermore, the finding that information acquired from the Internet is associated with making decisions to attend OOH services merits further investigations.

Conclusions

We found that people with T1D were extensive users of OOH services and that OOH service users were extensive users of eHealth. There was a positive association between the use of OOH services and the use of health apps, as well as between

XSL•FO
RenderX

the use of OOH services and self-reported anxiety or depression. Feeling anxious based on information from the Internet was positively associated with visiting OOH services. Likewise, there was a positive association between deciding to consult a doctor based on information from the Internet and the probability of visiting OOH services. The use of eHealth seems to have a large impact on people with T1D. This study investigated the use of health apps in general, and we think that the more specific use of diabetes self-care apps and associations with the use of different health care services would be of interest for future research. Decision making regarding doctor visits based on information from the Internet also merits further investigations, along with specific investigations regarding reasons for visiting.

Acknowledgments

We thank the Norwegian Diabetes Association for their cooperation in the performance of this study. Without their generous contribution, this study could not have been realized. We also thank the Northern Norway Health Authorities who provided funding for this research and approved the changes regarding data collection that had to be made in order to complete the study.

Authors' Contributions

All authors contributed to the design and conduct of the study. AHH drafted the manuscript. All authors contributed with improvements and critical revisions and approved the final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire.

[DOC File, 48KB - jmir_v21i3e13465_app1.doc]

References

1. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017 Jun;128:40-50. [doi: 10.1016/j.diabres.2017.03.024] [Medline: 28437734]
2. Norwegian Institute of Public Health. URL: https://www.fhi.no/globalassets/dokumenterfiler/rapporter/2018/helsetilstanden-i-norge-2018.pdf [accessed 2018-06-16] [WebCite Cache ID 70DeT070T]
3. Sandahl K, Nielsen LB, Svensson J, Johannesen J, Pociot F, Mortensen HB, et al. Increased mortality in a Danish cohort of young people with Type 1 diabetes mellitus followed for 24 years. Diabet Med 2017 Dec;34(3):380-386. [doi: 10.1111/dme.13124] [Medline: 27027777]
4. Jenssen TG, Tonstad S, Claudi T, Midthjell K, Cooper J. The gap between guidelines and practice in the treatment of type 2 diabetes A nationwide survey in Norway. Diabetes Res Clin Pract 2008 May;80(2):314-320. [doi: 10.1016/j.diabres.2007.12.025] [Medline: 18279994]
5. Mouland G. [Diabetes in general practice--were treatment goals reached?]. Tidsskr Nor Laegeforen 2014 Jan 28;134(2):168-172 [FREE Full text] [doi: 10.4045/tidsskr.13.0375] [Medline: 24477150]
6. Cooper JG, Claudi T, Thordarson HB, Løvaas KF, Carlsen S, Sandberg S, et al. Treatment of type 1 diabetes in the specialist health service--data from the Norwegian Diabetes Register for Adults. Tidsskr Nor Laegeforen 2013 Nov 12;133(21):2257-2262 [FREE Full text] [doi: 10.4045/tidsskr.13.0153] [Medline: 24226332]
7. World Health Organization. URL: http://www.who.int/ehealth/en/ [accessed 2018-06-16] [WebCite Cache ID 70DfcHSuE]
8. Statistics Norway. Technology and Innovation. URL: https://www.ssb.no/teknologi-og-innovasjon/faktaside [accessed 2019-01-23] [WebCite Cache ID 75dRo1irv]
9. Statistics Norway. Internet and Mobile Phones. URL: https://www.ssb.no/teknologi-og-innovasjon/faktaside/internett-og-mobil [accessed 2019-01-23] [WebCite Cache ID 75dRub2Wq]
10. Wangberg SC, Andreassen HK, Prokosch H, Santana SMV, Sørensen T, Chronaki CE. Relations between Internet use, socio-economic status (SES), social support and subjective health. Health Promot Int 2008 Mar;23(1):70-77. [doi: 10.1093/heapro/dam039] [Medline: 18083686]
11. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. BMC Public Health 2007 Apr 10;7:53 [FREE Full text] [doi: 10.1186/1471-2458-7-53] [Medline: 17425798]
12. Bujnowska-Fedak MM. Trends in the use of the Internet for health purposes in Poland. BMC Public Health 2015 Feb 27;15:194 [FREE Full text] [doi: 10.1186/s12889-015-1473-3] [Medline: 25886280]

XSL•FO
RenderX

13. Fox S, Duggan M. Health Online. Washington, DC: Pew Research Center; 2013. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2018-06-16] [WebCite Cache ID 70Dh4ruWf]
14. Hansen AH, Broz J, Claudi T, Årsand E. Relations Between the Use of Electronic Health and the Use of General Practitioner and Somatic Specialist Visits in Patients With Type 1 Diabetes: Cross-Sectional Study. J Med Internet Res 2018 Nov 07;20(11):e11322 [FREE Full text] [doi: 10.2196/11322] [Medline: 30404766]
15. Sørensen T, Andreassen HK, Wangberg SC. Prosjektrapport e-helse i Norge 2013. 2014. URL: https://ehealthresearch.no/files/documents/Prosjektrapporter/NST-rapport_2014-02_e-helse_i_Norge_2013.pdf [accessed 2018-06-16] [WebCite Cache ID 70DhrW5OT]
16. Medlock S, Eslami S, Askari M, Sent D, de Rooij SE, Abu-Hanna A. The consequences of seniors seeking health information using the internet and other sources. Stud Health Technol Inform 2013;192:457-460. [Medline: 23920596]
17. Hansen AH, Kristoffersen AE, Lian OS, Halvorsen PA. Continuity of GP care is associated with lower use of complementary and alternative medical providers: a population-based cross-sectional survey. BMC Health Serv Res 2014 Dec 10;14:629 [FREE Full text] [doi: 10.1186/s12913-014-0629-7] [Medline: 25491638]
18. Sandvik H, Hunskaar S. Frequent attenders at primary care out-of-hours services: a registry-based observational study in Norway. BMC Health Serv Res 2018 Dec 25;18(1):492 [FREE Full text] [doi: 10.1186/s12913-018-3310-8] [Medline: 29940934]
19. Sandvik H, Hunskår S, Diaz E. Use of emergency medical services by patients encompassed by the Regular GP scheme. Tidsskr Nor Laegeforen 2012 Oct 30;132(20):2272-2276 [FREE Full text] [doi: 10.4045/tidsskr.12.0090] [Medline: 23736193]
20. Statistics Norway. General practitoners services. 2018. URL: http://intranett.unn.no/?lang=no_NO [accessed 2019-01-22] [WebCite Cache ID 75cKt2uI0]
21. Thoresen CK, Sandvik H, Hunskaar S. Cancer patients' use of primary care out-of-hours services: a cross-sectional study in Norway. Scand J Prim Health Care 2016 Sep;34(3):232-239 [FREE Full text] [doi: 10.1080/02813432.2016.1207140] [Medline: 27406005]
22. Catan G, Espanha R, Mendes RV, Toren O, Chinitz D. Health information technology implementation - impacts and policy considerations: a comparison between Israel and Portugal. Isr J Health Policy Res 2015;4:41 [FREE Full text] [doi: 10.1186/s13584-015-0040-9] [Medline: 26269740]
23. 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. [Medline: 11780707]
24. Hansen AH, Bradway M, Broz J, Claudi T, Henriksen O, Wangberg SC, et al. The Use of eHealth and Provider-Based Health Services by Patients with Diabetes Mellitus: Protocol for a Cross-Sectional Study. JMIR Res Protoc 2016 Oct 31;5(4):e207 [FREE Full text] [doi: 10.2196/resprot.6529] [Medline: 27799136]
25. The Norwegian Diabetes Association: Årsberetning 2017. 2018. URL: https://issuu.com/knutjarle/docs/a_rsberetning_2017_ [accessed 2018-06-16] [WebCite Cache ID 70DkKIFTA]
26. 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 08;17(1):e10 [FREE Full text] [doi: 10.2196/jmir.3749] [Medline: 25574815]
27. The Tromsø Study. URL: http://tromsoundersokelsen.uit.no/tromso/ [accessed 2018-06-16] [WebCite Cache ID 70Dkah0hU]
28. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. JAMA 2012 May 2;307(17):1805-1806. [doi: 10.1001/jama.2012.3532] [Medline: 22550194]
29. Sandvik H, Hunskår S, Blinkenberg J. Årsstatistikk fra legevakt 2017. Report #2-2018.: National Centre for Emergency Primary Health Care, Uni Research; 2017. URL: https://bora.uib.no/bitstream/handle/1956/17544/%c3%85rsstatistikk_fra_legevakt_2017.pdf?sequence=1&isAllowed=y [WebCite Cache ID 75dSUx5tC]
30. den Boer-Wolters D, Knol MJ, Smulders K, de Wit NJ. Frequent attendance of primary care out-of-hours services in the Netherlands: characteristics of patients and presented morbidity. Fam Pract 2010 Apr;27(2):129-134. [doi: 10.1093/fampra/cmp103] [Medline: 20032165]
31. Washington R, Andrews R, Mutter R. Emergency department visits for adults with diabetes, 2010 statistical brief #167. URL: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb167.jsp [WebCite Cache ID 75dSx2pAT]
32. Li J, Theng Y, Foo S. Predictors of online health information seeking behavior: Changes between 2002 and 2012. Health Informatics J 2016 Dec;22(4):804-814. [doi: 10.1177/1460458215595851] [Medline: 26261218]
33. Ernsting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmey A, et al. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. J Med Internet Res 2017 Dec 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
34. Trawley S, Baptista S, Browne JL, Pouwer F, Speight J. The Use of Mobile Applications Among Adults with Type 1 and Type 2 Diabetes: Results from the Second MILES-Australia (MILES-2) Study. Diabetes Technol Ther 2017 Dec;19(12):730-738. [doi: 10.1089/dia.2017.0235] [Medline: 29028442]
35. Dent A, Hunter G, Webster AP. The impact of frequent attenders on a UK emergency department. Eur J Emerg Med 2010 Dec;17(6):332-336. [doi: 10.1097/MEJ.0b013e328335623d] [Medline: 20038842]

XSL•FO
RenderX

36. Flarup L, Moth G, Christensen MB, Vestergaard M, Olesen F, Vedsted P. Chronic-disease patients and their use of out-of-hours primary health care: a cross-sectional study. BMC Fam Pract 2014 Jun 09;15:114 [FREE Full text] [doi: 10.1186/1471-2296-15-114] [Medline: 24912378]
37. Keizer E, Maassen I, Smits M, Wensing M, Giesen P. Reducing the use of out-of-hours primary care services: A survey among Dutch general practitioners. Eur J Gen Pract 2016 Sep;22(3):189-195. [doi: 10.1080/13814788.2016.1178718] [Medline: 27248713]
38. Keizer E, Smits M, Peters Y, Huibers L, Giesen P, Wensing M. Contacts with out-of-hours primary care for nonurgent problems: patients' beliefs or deficiencies in healthcare? BMC Fam Pract 2015 Oct 28;16:157 [FREE Full text] [doi: 10.1186/s12875-015-0376-9] [Medline: 26510620]
39. Brzan PP, Rotman E, Pajnkihar M, Klanjsek P. Mobile Applications for Control and Self Management of Diabetes: A Systematic Review. J Med Syst 2016 Sep;40(9):210. [doi: 10.1007/s10916-016-0564-8] [Medline: 27520615]
40. Zhou Y, Abel G, Warren F, Roland M, Campbell J, Lyratzopoulos G. Do difficulties in accessing in-hours primary care predict higher use of out-of-hours GP services? Evidence from an English National Patient Survey. Emerg Med J 2015 May;32(5):373-378 [FREE Full text] [doi: 10.1136/emermed-2013-203451] [Medline: 24850778]
41. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
42. Singh K, Brown RJ. Health-related internet habits and health anxiety in university students. Anxiety Stress Coping 2014;27(5):542-554. [doi: 10.1080/10615806.2014.888061] [Medline: 24467278]
43. Singh K, Fox J, Brown RJ. Health anxiety and Internet use: A thematic analysis. Cyberpsychology: Journal of Psychosocial Research on Cyberspace 2016;10(2):- [FREE Full text] [doi: 10.5817/CP2016-2-4]
44. Doherty-Torstrick ER, Walton KE, Fallon BA. Cyberchondria: Parsing Health Anxiety From Online Behavior. Psychosomatics 2016;57(4):390-400 [FREE Full text] [doi: 10.1016/j.psym.2016.02.002] [Medline: 27044514]
45. Holubova A, Bradway M, Årsand E, Hallgren D, Hartvigsen G. Do mobile medical apps need to follow European and US regulations or not: decisions exemplified by diabetes management app. Linköping: Linköping University Electronic Press; 2015 Jun 15 Presented at: 13th Scandinavian Conference on Health Informatics; June 15-17, 2015; Tromsø, Norway p. 55.
46. Krieger WH. When are medical apps medical? Off-label use and the Food and Drug Administration. Digit Health 2016;2:2055207616662782 [FREE Full text] [doi: 10.1177/2055207616662782] [Medline: 29942563]
47. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: A systematic review. J Telemed Telecare 2018 Jan;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
48. 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]
49. Eastin MS, Guinsler NM. Worried and wired: effects of health anxiety on information-seeking and health care utilization behaviors. Cyberpsychol Behav 2006 Aug;9(4):494-498. [doi: 10.1089/cpb.2006.9.494] [Medline: 16901253]
50. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
51. Takahashi Y, Ohura T, Ishizaki T, Okamoto S, Miki K, Naito M, et al. Internet use for health-related information via personal computers and cell phones in Japan: a cross-sectional population-based survey. J Med Internet Res 2011 Dec 14;13(4):e110 [FREE Full text] [doi: 10.2196/jmir.1796] [Medline: 22169526]
52. Beck F, Richard J, Nguyen-Thanh V, Montagni I, Parizot I, Renahy E. Use of the internet as a health information resource among French young adults: results from a nationally representative survey. J Med Internet Res 2014 May 13;16(5):e128 [FREE Full text] [doi: 10.2196/jmir.2934] [Medline: 24824164]
53. Halbesleben JRB, Whitman MV. Evaluating survey quality in health services research: a decision framework for assessing nonresponse bias. Health Serv Res 2013 Jun;48(3):913-930 [FREE Full text] [doi: 10.1111/1475-6773.12002] [Medline: 23046097]
54. Raknes G, Morken T, Hunskår S. [Travel distance and the utilisation of out-of-hours services]. Tidsskr Nor Laegeforen 2014 Nov;134(22):2151-2155 [FREE Full text] [doi: 10.4045/tidsskr.14.0443] [Medline: 25423980]

Abbreviations

CI: confidence interval
GP: general practitioner
NDA: Norwegian Diabetes Association
NSD: Norwegian Social Science Data Service
REK: Regional Committee for Medical and Health Research Ethics
OOH: out-of-hours
OR: odds ratio
T1D: type 1 diabetes

XSL•FO
RenderX

Edited by G Eysenbach; submitted 23.01.19; peer-reviewed by A Albanese-O'Neill, K Fitzner, K McMillan, L Hu; comments to author 14.02.19; revised version received 20.02.19; accepted 07.03.19; published 21.03.19.

Please cite as:
Hansen AH, Claudi T, Årsand E
Associations Between the Use of eHealth and Out-of-Hours Services in People With Type 1 Diabetes: Cross-Sectional Study
J Med Internet Res 2019;21(3):e13465
URL: http://www.jmir.org/2019/3/e13465/
doi:10.2196/13465
PMID:30896437

©Anne Helen Hansen, Tor Claudi, Eirik Årsand. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Sustainable Adoption of Digital Health Innovations: Perspectives From a Stakeholder Workshop

Michelle Helena Van Velthoven[1], BSc, MSc, PhD; Carlos Cordon[2], PhD

[1]Department of Paediatrics, University of Oxford, Oxford, United Kingdom

[2]International Institute for Management Development Business School, Lausanne, Switzerland

Corresponding Author:
Michelle Helena Van Velthoven, BSc, MSc, PhD
Department of Paediatrics
University of Oxford
John Radcliffe Hospital
Oxford, OX3 9DU
United Kingdom
Phone: 44 7425167828
Email: michelle.van.velthoven@gmail.com

Abstract

Background: There are various complex reasons that influence sustainable adoption of innovations in health care systems. Low adoption can be caused by a lack of support from one or more stakeholders because their needs and expectations are not always considered or aligned.

Objective: This study aimed to identify stakeholders' perceptions of barriers and facilitators toward the sustainable adoption of digital health innovations.

Methods: A stakeholder workshop was attended by 12 participants with a range of backgrounds on August 25, 2017, including people representing the views from patients, carers, local hospitals, pharmacy retailers, health insurers, health services researchers, engineers, and technology and pharmaceutical companies in Switzerland. On the basis of adoption of innovation frameworks, we asked participants to interview each other about 3 factors influencing the adoption of digitally delivered health interventions: (1) Facilitators and barriers in the external system, (2) Needs and expectations of stakeholders, and (3) Safety, quality, and usability of innovations. The worksheets and videos generated from the workshop were qualitatively analyzed and summarized.

Results: Facilitators for adoption mentioned were high levels of income and education, and digital health is a high priority to stakeholders. Main common interests of different stakeholders were patient satisfaction and job protection. Health care spending was a misaligned interest: although some stakeholders were keen on spending more to obtain or provide the highest quality of care, others were focused on reducing health care spending to provide cost-effective services. Switzerland's diversity and complexity, in terms of its organization with 26 cantons (administrative divisions), were barriers as these made it harder to ensure interoperability of interventions. A culture of innovation was considered a push factor, but adoption was inhibited by persistent paper-based systems, a fear of change, and unwillingness to share data. The sustainability of interventions can be promoted by making them patient-centered, meaning that patients should be involved throughout their development.

Conclusions: Promoting sustainable adoption of digital health remains challenging despite various push factors being in place. Barriers related to fragmentation, patient-centeredness, data security, privacy, trust, and job security need to be addressed. A strength is that people from a wide range of backgrounds attended the workshop. A limitation is that the findings are focused on the macro level. In-depth case studies of specific issues need to be conducted in different settings.

(J Med Internet Res 2019;21(3):e11922) doi:10.2196/11922

KEYWORDS

adoption; digital health; stakeholders; telemedicine

XSL•FO
RenderX

Introduction

Background

There is an ongoing discussion of how digital innovation can be used to improve health systems (the organization of people, institutions, and resources that deliver health care services to meet the health needs of people) around the world. Some countries have successfully adopted digital health innovation, for example, patient portals for Web access to clinicians are widespread in the United States [1]. However, health systems are notorious for slow and unsuccessful adoption of digital health innovation [2]. Currently, relatively few digital health innovations have been efficiently used despite various actors having expressed enthusiasm for digital health, and large investments are being made [3]. There are promising digital health innovations that are not widely adopted, for example, in the United Kingdom, general practice emails [4] or outpatient video consultations [5] are still not commonly used, or they are abandoned when people fail to scale them up or sustain use over time at an organization or system level. For example, a personal electronic health (eHealth) records program implemented in the United Kingdom between 2007 and 2010 was discontinued because very few people opened an account. Those who did, found it not easy or useful to use, and their expectations regarding self-management were not met [6].

Digital health innovations, also known as digitally delivered interventions, can be categorized in ways in which digital and mobile technologies are being used to support health system needs, including interventions for the public and patients, health care workers, health system managers, and data services [7]. Digital health innovations involve interactions within the health system, as well as the wider social, legal, political, and economic context [8]. There are numerous frameworks, theories, and systematic reviews of studies on the adoption of information and technology in health care [6,9-13]. This research has shown that not enough is known about the social, organizational, and cultural elements of successful implementation and adoption of digital health. Recently, a comprehensive framework on the diffusion of innovation in service organizations researched the spread, and sustainability of innovation in health services was expanded to research issues beyond adoption [14]. Factors influencing nonadoption and abandonment are complex and include the health condition, technology, value proposition, adopter system (comprising professional staff, patient, and lay caregivers), organizations, wider context, and interaction and mutual adaptation among all these factors over time. For example, more specific mentioned reasons for nonuse are complexities related to regulation [15,16], dealing with changes in workflow [17], the trust of the population regarding the privacy and security of data, and the conflicting interest of different stakeholders in the health care system [18,19].

Objectives

One of the reasons for low adoption of innovations is a lack of support from 1 or more health care stakeholders, such as hospitals, health insurers, pharmaceutical companies, retailers, regulators, and patients [2]. The needs and expectations of different stakeholders are not always considered or aligned, and in some cases, there are conflicts among stakeholders. The aim of the workshop was to identify stakeholders' perceptions as facilitators and barriers toward the successful adoption of digital health innovations, using Switzerland as a case study.

Methods

Setting

This stakeholder workshop took place on August 25, 2017 at the International Institute for Management Development (IMD) Business School in Lausanne, Switzerland. In this qualitative study, participants took part in small group exercises where examples from the Swiss health care system were discussed using relevant frameworks.

Participants

Participants were selected through personal connections by searching the internet for relevant contacts and snowballing (eg, asking participants whether they know another relevant person). A total of 12 participants with a wide range of backgrounds attended the workshop, including people representing the views from patients, carers, local hospitals, pharmacy retailers, health insurers, health services researchers, engineers, and technology and pharmaceutical companies (Table 1). All participants provided verbal permission for the information they gave during the day to be used for research purposes. They were neither asked to pay for the workshop nor did they receive payments. Refreshments and lunch were provided during the day.

Questions

On the basis of the nonadoption, abandonment, spread, scale-up, and sustainability framework for digital health innovations [14], we focused on 3 aspects from this framework: (1) Facilitators and barriers in the wider external system (political, regulatory, professional, and sociocultural), (2) Needs and expectations of stakeholders (eg, patients and the public, health professionals, health clinics/hospitals, pharmaceuticals, and insurers), and (3) development of safe, high-quality, and usable digital health innovations (technology). This framework and these aspects were chosen as they provide insights into the most recent and comprehensive thinking on the adoption of digital health innovations.

XSL•FO
RenderX

Table 1. Workshop participants' background.

Participant number	Stakeholder and position, and background
1	University hospital: medical director
2	University hospital: head of electronic health domain
3	Patient representative: medical writer background
4	Patient representative: insurance consultant background
5	Patient representative: caretaker background
6	Academic: health sciences researcher
7	Pharmaceutical company: consumer health director
8	Pharmaceutical/technological company: computer engineer lead
9	Private equity: health investor specialist
10	Health insurance company: senior project manager
11	Pharmaceutical retail: specialist
12	Business: digital transformation specialist

Exercise

We asked participants to take part in a "Round Robin exercise" in which they interviewed each other on 3 questions (see Multimedia Appendix 1). This method was chosen because it facilitates participants who are less assertive to contribute to a discussion by incorporating the responses of all participants to questions. Furthermore, participants are asked to provide answers without having listened to other participants, which reduces participants being influenced by others. This helps to overcome the issue that 1 or a few of the most talkative and dominant participants set the tone for the discussion, which might result in overlooking the ideas from others. We provided participants with detailed instructions on the exercise. A total of 3 groups of 4 participants were formed and given separate rooms. We provided each group 1 of the 3 questions to discuss (participants in group 1, question 1, etc). To ensure that all participants understood the exercise, they were given 15 min to discuss their question and the process of the exercise in their group and feed their understanding back to the facilitators. A total of 2 participants of each group rotated among the groups. During these rounds, 1 participant interviewed a member of another group (interviewee). After a 10-min interview, the participants reversed roles, and the interviewee became an interviewer. After the interview rounds, participants gathered in their own group to discuss and summarize their findings. Afterwards, all participants returned to the large room where each group presented their findings, and in-depth discussions explored their views on facilitators and barriers toward the adoption of digital health innovations.

Data Collection and Analysis

We asked participants to write out their findings on large A4 worksheets. An assistant filmed the general discussion at the end of the workshop and took photos of the A4 sheets. The authors of this study were the moderators and observers of the workshop. They greeted the participants individually as they arrived and provided a welcome and introduction talk at the start of the workshop. They observed how the participants engaged in the exercise, answered questions, and listened to comments. Thematic analysis was conducted by the authors through separately watching videos in an active way (searching for meaning) to obtain an overview of the findings for each question and note their thoughts. Initial codes were given to findings (units of texts). The authors compared and discussed their coding, searched for themes, and sorted codes into potential themes. The authors carried out this process independently and discussed and compared their findings. Themes were related to each other to develop an explanation in relation to the research question. Close attention was paid to how the general discussion linked to the A4 sheets with findings. Data were summarized, and participants were sent the results and asked for feedback that was incorporated in the results.

Results

Overview

In the context of Switzerland, we describe 3 aspects of adoption of digital health innovations: (1) Facilitators and barriers in the wider external system (political, regulatory, professional, and sociocultural), (2) Interests of stakeholders (eg, patients and the public, health professionals, health clinics/hospitals, pharmaceuticals, and insurers), and (3) Development of safe, high-quality, and usable digital health innovations (technology).

Key Facilitators and Barriers for Digital Health Adoption

Facilitators for digital health adoption mentioned were Switzerland being a rich country with high levels of education and digital health being a high priority to different stakeholders (Table 2). Diversity and complexity, particularly in terms of Switzerland's organization with 26 cantons, were barriers as this was thought to make it harder to ensure interoperability of interventions. An innovation culture was seen as beneficial, but at the same time, it was thought that Swiss people feared change and were not willing to share data, and there is a tradition of paper-based systems in the country. A key success factor for digital health adoption involved clarification of the context around interventions, such as regulation for security and privacy.

XSL•FO
RenderX

Table 2. Key facilitators, barriers, and success factors for digital health adoption.

Facilitators	Barriers	Success factors
Priority: Digital health is a priority to improve the quality of health services in Switzerland; politicians are educated on the importance and issues related to digitization	Decentralization and lack of interoperability: Switzerland has 26 cantons and different health laws; there is no common interface for electronic patient record interoperability. Furthermore, there are 4 different languages used (French, German, Italian, and Romansh)	Context clarification: There is a need to put national regulation in place for security, privacy, and "hacking" of digital health innovations (medical devices)
High-income country: Large investments are being made in innovation; furthermore, most people are educated and have personal technologies such as smartphones	Lack of long-term planning: People can change health insurance every year, which limits long-term investment of insurance companies	Responsibilities clarification: The responsibilities of different stakeholders need to be clarified, for example, who pays which costs
Size: Switzerland is a small country where people know each other	Experience: There is relatively limited experience within Switzerland	Import: Switzerland could make better use of the considerable number of expats bringing in their expertise
Culture of innovation: Switzerland has a thriving start-up, and health companies' "valley" and local initiatives are trialed. Many initiatives over the past two decades have taken place, including the establishment of pharma companies	High health care costs: The overuse of health care is insufficiently limited. Patients pay high insurance and in return want high-quality services. There is a lack of willingness to pay for digital health	Value proposition clarification: The benefits and impact on costs for citizens, for example, fees for services, need to be clarified
Consensus: Swiss people like to solve problems together with all stakeholders and common agreement. Opportunities to share data are provided	Low agility: There is a general lack of willingness to share personal data. Rigid and slow adoption of innovation, for example, a paper billing system ("System de Tarification")	Change enablement: Change processes need to be better enabled to achieve an appropriate proportion of people using a digital health innovation

Furthermore, the benefits and costs for patients and the public would have to be better explained.

Aligned and Differing Interests of Stakeholders

The main common interest for different stakeholders (see Table 3) was increased efficiency, which could lead to a reduction in cost. It was mentioned that these savings could then be used, for example, for research and development of drugs for rare diseases. Conflicts were mentioned for health care spending, although some stakeholders such as patient groups and certain companies were keen on spending more to obtain the highest-quality care, and in the long-term, this could make health care spending unsustainable. Furthermore, there were conflicting interests mentioned for the better use of data, for example, although patients would be interested in contributing to research and development of drugs and devices, they were concerned about threats to their privacy.

Facilitators and Barriers Toward Developing Digital Health Interventions

To develop safe, high-quality, and usable innovations, it was mentioned that agile engineer approaches could be used to more rapidly develop innovations with measurable outcomes, but traditional health care companies such as pharmaceutical companies are more used to lengthy drug development processes (see Table 4). Patient-facing interventions should be patient-centered, meaning that patients should be involved in their development. However, there was said to be a lack of studies on patient involvement in digital health studies. Data derived from digital health tools could be used for individualization of health care if personal data were to be obtained. Although more jobs are needed to enable digitization in health care, there were concerns about the loss of jobs caused by digitization.

XSL•FO
RenderX

Table 3. Aligned and conflicting interests of stakeholders.

Stakeholders, aligned interests	Conflicting interests
Pharmaceutical companies	
The possibility to access more targeted data to be used in the development of drugs	The possibility of being left behind when data are not shared
Potentially increased customer satisfaction	Patients could become more demanding, which could result in the need to change culture and processes
Increased efficiency in the use of resources, for example, by increased compliance of patients taking drugs	A potential reduction in drug sales, for example, by paying per pill versus paying per box of pills, which improves the accuracy of drugs dispensing
Increase in rare disease research and development on the basis of gains from other disease areas	Costs could be escalating by investment in digital health with uncertain returns
Hospitals and health workers	
Saving staff time and lowering costs	Uncertainty about who pays for what time
Better collaboration with colleagues, for example, by opportunities for information sharing	Health workers could be losing some of their autonomy, for example, by more traceable work by tracking what a staff member is doing
Patients and the public	
The opportunity to contribute feedback to research and development, for example, postmarket feedback on adverse effects of drugs	Threats to privacy, for example, it being uncertain what happens to personal data
Insurers	
An increased efficiency in treatment, for example, reduced time in the hospital	Challenges for balancing expensive treatment versus the cost to society. It might be the right decision to pay for expensive treatment if the patient can return to society sooner
Expert high-quality care for patients	Private versus mandatory insurance, for example, risk selection on the basis of available personal data. Even though legally this is not possible, it is happening unofficially

Table 4. Facilitators and barriers toward developing safe, high-quality, and usability innovations.

Facilitators	Barriers
Agile approaches can be used to develop digital health innovations	Pharmaceutical companies are not used to using agile approaches
A sufficient mass of secure data, for example, use of data for individualization of treatment	People do not want to share data and are concerned about data ownership
The government is seen as credible and could use this to, for example, define what a medical device is and support the establishment of standards for safety	Companies developing digital health innovations are not always seen as credible
Patient-facing interventions should be focused on patients, which can be achieved by, for example, involving them throughout the development process	Lack of understanding what patient-centeredness really means through studies on patient involvement
Clinicians would like to be involved	Clinicians' role for involvement needs to be defined
Interventions need to have shared benefits and measurable outcomes	There are concerns about how to measure cost-effectiveness
There is a need for more jobs in digital health	People are concerned about job losses
Insurers need to be transparent about their willingness to pay for digital health innovations	People are able to change their health care insurance

Discussion

Principal Findings

This paper reports on findings from a workshop with the aim to identify stakeholders' perceptions of barriers and facilitators toward the successful adoption of digital health innovations. The workshop led to 3 main interesting insights. First, there was a lack of understanding of how the different stakeholders in the ecosystem work and how they are incentivized. For example, it was assumed that 1 of the key roles of health insurance is to control health care costs. However, the reality in Switzerland and other countries with mandatory health insurance is that there is neither incentive nor a role for insurers to control or reduce health care–related costs [20]. Second, there is a lack of patient involvement in the development of digital health initiatives. Despite the push for patient-centeredness [21], 2 representatives from patient organizations in our workshop explained that they were never consulted by any of the other stakeholders. Third, stakeholders launch initiatives with a lack

XSL•FO
RenderX

of understanding of the basics of digital solutions. Usability was frequently mentioned as the main problem in digital solutions in health care even though there is a considerable amount of literature on developing user-centered and engaging interventions [22-24].

Topics that were expected to be discussed included issues around data security, as they are frequently discussed in popular media [25]. According to the World Health Organization (WHO), Switzerland adopted a national eHealth policy in 2007 and a policy on multilingualism in eHealth in 2010 [26]. Switzerland does not have a national policy or strategy on the use of social media by government organizations; health care organizations promote health messages as part of health promotion campaigns, and individuals use social media to learn about health issues. Furthermore, there is no policy or strategy for governing the use of data in the health sector or by private companies [26]. Putting these policies in place could help to clarify the context in which digital health innovations can be safely and securely used in Switzerland. Participants mentioned that Swiss people do not like to share data and that Switzerland is a "paper country." The culture of change was not discussed even though there is a lot of academic literature on this topic [17,27], because participants assumed that in our connected world, people are open to digital innovation adoption. It was not expected that the risk of losing employment was an important issue in Switzerland given its high employment rate and difficulty with hiring specialized people.

Strengths and Limitations

Strengths of this workshop include that we invited people from a wide range of backgrounds and we purposively selected a small number of people to allow for in-depth discussions. Limitations include that the workshop was only held for 1 day and that the findings are focused on the macrolevel and do not provide meso or microlevel insights. The findings are contextual to Switzerland and not necessarily generalizable to other settings.

Comparison With Previous Work

Our findings have implications for policy makers, practitioners, and companies who want to develop digital health innovations. First, companies should keep in mind that countries such as Switzerland, despite being small, can be highly fragmented. Integration of care and data and scale-up of innovations are challenging with tiered governance and 26 cantons with different health laws. As a result, Switzerland does not have a national eHealth record system [26]. Furthermore, this means that many telemedicine services (eg, teleradiology, teledermatology, telepathology, and remote patient monitoring services) and mobile health programs (eg, toll-free emergency, health call centers, and community mobilization for accessing/providing health information) are regional, local, or in an information state [26]. Second, companies should consider the needs of users of digital health innovations and involve them throughout the design and development. Third, the incentives for digital health–related jobs should be better organized. Although Switzerland has a high employment rate and difficulty with hiring specialized people, job losses are a concern for Swiss people that limit implementation-readiness of providers. Although some jobs may be replaced by digital health innovations, this also creates new jobs. According to the WHO, less than 25% of health sciences students receive preservice training in eHealth, and more than 75% of health professionals receive in-service training in eHealth [26]. Educating and (re)training the workforce in digital health will be important to reduce job loss concerns.

Conclusions

In conclusion, countries such as Switzerland with an advanced infrastructure for information and communication technology and a high quality of care make an attractive place for companies to develop digital health innovations. However, barriers related to fragmentation, patient-centeredness, trust, and job security need to be addressed.

Acknowledgments

This study was supported by the Research Value Center Value Chain 4.0 at the International Institute for Management Development Business School in Lausanne, Switzerland.

Authors' Contributions

MHV wrote this study. CC revised the study and provided important intellectual comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Digital health ecosystem Round Robin exercise.

[PDF File (Adobe PDF File), 35KB - jmir_v21i3e11922_app1.pdf]

References

1. Pearl R. Kaiser Permanente Northern California: current experiences with internet, mobile, and video technologies. Health Aff (Millwood) 2014 Feb;33(2):251-257. [doi: 10.1377/hlthaff.2013.1005] [Medline: 24493768]

XSL•FO
RenderX

2. Lennon MR, Bouamrane M, Devlin AM, O'Connor S, O'Donnell C, Chetty U, et al. Readiness for delivering digital health at scale: lessons from a longitudinal qualitative evaluation of a national digital health innovation program in the United Kingdom. J Med Internet Res 2017 Feb 16;19(2):e42. [doi: 10.2196/jmir.6900] [Medline: 28209558]
3. Muoio D. MobiHealthNews. Digital health funding in Q1 2018 URL: https://www.mobihealthnews.com/content/digital-health-funding-q1-2018 [accessed 2019-02-05] [WebCite Cache ID 75xlyMXT1]
4. Brant H, Atherton H, Ziebland S, McKinstry B, Campbell JL, Salisbury C. Using alternatives to face-to-face consultations: a survey of prevalence and attitudes in general practice. Br J Gen Pract 2016 Jul;66(648):e460-e466 [FREE Full text] [doi: 10.3399/bjgp16X685597] [Medline: 27215571]
5. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018 Dec 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
6. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. Br Med J 2010 Nov 16;341:c5814. [doi: 10.1136/bmj.c5814]
7. World Health Organization. Classification of Digital Health Interventions v 1 URL: http://www.who.int/sorry/ [accessed 2019-02-05] [WebCite Cache ID 75xm47ft3]
8. 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. Br Med J 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
9. Gagnon M, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. J Med Syst 2012 Feb;36(1):241-277 [FREE Full text] [doi: 10.1007/s10916-010-9473-4] [Medline: 20703721]
10. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006;6:3 [FREE Full text] [doi: 10.1186/1472-6947-6-3] [Medline: 16401336]
11. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: toward an integrative view. Inform Manage 2006 Apr;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
12. Robert G, Greenhalgh T, MacFarlane F, Peacock R. Adopting and assimilating new non-pharmaceutical technologies into health care: a systematic review. J Health Serv Res Policy 2010 Oct;15(4):243-250. [doi: 10.1258/jhsrp.2010.009137] [Medline: 20592046]
13. Wisdom JP, Chor KH, Hoagwood KE, Horwitz SM. Innovation adoption: a review of theories and constructs. Adm Policy Ment Health 2014 Jul;41(4):480-502 [FREE Full text] [doi: 10.1007/s10488-013-0486-4] [Medline: 23549911]
14. 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 1;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
15. Garell C, Svedberg P, Nygren JM. A legal framework to support development and assessment of digital health services. JMIR Med Inform 2016 May 25;4(2):e17 [FREE Full text] [doi: 10.2196/medinform.5401] [Medline: 27226391]
16. Parimbelli E, Bottalico B, Losiouk E, Tomasi M, Santosuosso A, Lanzola G, et al. Trusting telemedicine: a discussion on risks, safety, legal implications and liability of involved stakeholders. Int J Med Inform 2018 Apr;112:90-98. [doi: 10.1016/j.ijmedinf.2018.01.012] [Medline: 29500027]
17. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of ehealth interventions: systematic review of the literature. J Med Internet Res 2018 May 1;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
18. Devlin AM, McGee-Lennon M, O'Donnell CA, Bouamrane M, Agbakoba R, O'Connor S, "dallas" evaluation team. Delivering digital health and well-being at scale: lessons learned during the implementation of the dallas program in the United Kingdom. J Am Med Inform Assoc 2016 Jan;23(1):48-59 [FREE Full text] [doi: 10.1093/jamia/ocv097] [Medline: 26254480]
19. Panyard DJ, Ramly E, Dean SM, Bartels CM. Bridging clinical researcher perceptions and health IT realities: a case study of stakeholder creep. Int J Med Inform 2018 Feb;110:19-24. [doi: 10.1016/j.ijmedinf.2017.11.014] [Medline: 29331251]
20. Sturny I. International Health Care System Profiles. The Swiss Health Care System URL: https://international.commonwealthfund.org/countries/switzerland/ [accessed 2019-02-05] [WebCite Cache ID 75xoDYlk4]
21. Bokhour BG, Fix GM, Mueller NM, Barker AM, Lavela SL, Hill JN, et al. How can healthcare organizations implement patient-centered care? Examining a large-scale cultural transformation. BMC Health Serv Res 2018 Dec 7;18(1):168 [FREE Full text] [doi: 10.1186/s12913-018-2949-5] [Medline: 29514631]
22. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
23. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]

XSL•FO
RenderX

24. 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]
25. BBC News. Data Protection URL: https://www.bbc.co.uk/news/topics/cwz4lvzgq9gt/data-protection [accessed 2019-02-05] [WebCite Cache ID 75xohguHu]
26. World Health Organization. 2015. Atlas of eHealth country profiles 2015: the use of eHealth in support of universal health coverage: Based on the findings of the 2015 global survey on eHealth URL: https://www.who.int/goe/publications/atlas_2015/en/ [accessed 2019-02-05] [WebCite Cache ID 75xp1DJwp]
27. Willis CD, Saul J, Bevan H, Scheirer MA, Best A, Greenhalgh T, et al. Sustaining organizational culture change in health systems. J Health Organ Manag 2016;30(1):2-30. [doi: 10.1108/JHOM-07-2014-0117] [Medline: 26964847]

Abbreviations

eHealth: electronic health
IMD: International Institute for Management Development
WHO: World Health Organization

Edited by G Eysenbach; submitted 11.08.18; peer-reviewed by S Zheng, M Bellio, W Zhang; comments to author 12.10.18; revised version received 06.12.18; accepted 12.12.18; published 25.03.19.

Please cite as:
Van Velthoven MH, Cordon C
Sustainable Adoption of Digital Health Innovations: Perspectives From a Stakeholder Workshop
J Med Internet Res 2019;21(3):e11922
URL: http://www.jmir.org/2019/3/e11922/
doi:10.2196/11922
PMID:30907734

©Michelle Helena Van Velthoven, Carlos Cordon. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Improving the Quality of Consumer Health Information on Wikipedia: Case Series

Shira Schecter Weiner[1*], PT, PhD; Jill Horbacewicz[1*], PT, PhD; Lane Rasberry[2*], BS; Yocheved Bensinger-Brody[1*], PT, PhD

[1]Doctor of Physical Therapy Program, School of Health Science, Touro College, New York, NY, United States

[2]Data Science Institute, University of Virginia, Charlottesville, VA, United States

[*]all authors contributed equally

Corresponding Author:
Shira Schecter Weiner, PT, PhD
Doctor of Physical Therapy Program
School of Health Science
Touro College
320 W 31st Street
Physical Therapy Department
New York, NY,
United States
Phone: 1 212 463 0400 ext 55256
Email: shira.weiner4@touro.edu

Abstract

Background: Wikipedia is one of the most consulted health resources in the world. Since the public is using health information from Wikipedia to make health care decisions, improving the quality of that health information is in the public interest. The open editable content design of Wikipedia and quality control processes in place provide an opportunity to add high-value, evidence-based information and take an active role in improving the health care information infrastructure.

Objective: The aim of this project was to enhance Wikipedia health pages using high-quality, current research findings and track the persistence of those edits and number of page views after the changes to assess the reach of this initiative.

Methods: We conducted Wikipedia Editathons with 3 different cohorts of Physical Therapy (PT) students to add high-quality health information to existing Wikipedia pages. Students synthesized best evidence information and updated and/or corrected existing Wikipedia entries on specific health pages. To evaluate the impact of these contributions, we examined two factors: (1) response to our contributions from the Wikipedia editing community, including number and type of subsequent edits as well as persistence of the student contributions and (2) number of page views by the public from the time of the page edits.

Results: A total of 98 PT students in 3 different cohorts engaged in Editathons, editing 24 health pages. Of the 24 edits, 22 persisted at the end of the observation period (from time of entry to May 31, 2018) and received nearly 8 million page views. Each health page had an average of 354,724 page views.

Conclusions: The Wikipedia Editathon is an effective way to continuously enhance the quality of health information available on Wikipedia. It is also an excellent way of bridging health technology with best-evidence medical facts and disseminating accurate, useful information to the public.

(J Med Internet Res 2019;21(3):e12450) doi:10.2196/12450

KEYWORDS

consumer health information; health literacy; Wikipedia; public health; physical therapy education

Introduction

Wikipedia is one of the most, if not the most, consulted health resources in the world [1]. Therefore, improving the quality of its content ensures the dissemination of high-level information to all who visit its pages. The open editable content design of Wikipedia and the quality control processes in place provide an opportunity to add high-value, evidence-based information and take an active role in improving the health care information infrastructure.

XSL•FO
RenderX

The internet is a major source for the dissemination of health information. Research has shown that 50% of American adults seek health information from online resources [2]. This information can potentially influence health beliefs and health behaviors. The US government set the following goal in its Healthy People 2020 initiative: "use health communication strategies and health information technology to improve population health outcomes and health care quality, and to achieve health equity." The internet's reach and potential for improving the health of individuals as well as populations should be recognized by health care professionals as another tool in their toolbox to improve health care quality and equity.

When accessing information on the internet, users begin with a search engine and then select from the results. Wikipedia, which launched in 2001 as a free online encyclopedia, is frequently in one of the top five search result options. Traffic to Wikipedia's English-language medical content places it in the lead among other health internet sites, including WebMD and Mayo Clinic [1]. The English-language Wikipedia held approximately 30,000 articles on health-related topics in 2017 [3]. Across Wikipedia encyclopedias, in all languages, there were 155,000 health articles, which collectively received 4.8 billion page views in 2013. Those articles contained 950,000 citations [1]. The page view counts continue to increase with monthly averages estimated at 170 million views between July 2009 and April 2018, approximating more than 16 billion views during this period (see Multimedia Appendix 1 for spreadsheet data on page views).

Despite the fact that Wikipedia is a primary source for public health information, many health care professionals reject Wikipedia as a reliable source of information because of the open editable content. A more detailed understanding of Wikipedia's inner workings highlights the quality control processes in place to ensure inclusion of worthwhile information and detect misinformation. Because the articles are wiki documents, they are continuously editable and never considered final. Volunteer editors from all over the globe edit Wikipedia entries, and a history of all changes is tracked within the page and available for review. This nontraditional editorial model allows anyone with an internet connection to contribute and make changes. Typically, subject-matter experts, such as health care professionals and medical students, take on this role [4]. Therefore, as evidence-supported information from reliable and verifiable sources is added and misinformation or outdated content is removed, the quality of the information presented is continually enhanced. This is something that cannot be said for traditional publications. Using this crowdsourcing model facilitates disseminating a great deal of high-level information to large segments of the population.

Researchers have begun delving deeper into the quality of Wikipedia health content, utility of citations, and the editorial process. The Wikipedia standards demand the inclusion of citations, while most internet health sources do not impose this same standard. Hunter et al [5] compared both content and the relevant use of supporting references on Lexicomp (an online drug information source for clinicians) and Wikipedia. The results showed that although Wikipedia contained less content overall, all Wikipedia content was supported by peer-reviewed citations, while the same was true for only 63% of the Lexicomp information [5]. It has also been noted that the citations in Wikipedia's health pages become a "gateway to biomedical research," as clinicians use these links to launch further investigation into a particular topic [6].

A report by the Intercontinental Marketing Services Institute for Healthcare Informatics revealed that as many as 96% of all edits made in Wikipedia health pages are geared toward the patient [7]. It stands to reason that if the public uses the health information from Wikipedia to make health care decisions, improving the quality of that health information is in the public interest. While much has been written about the open editable content, the impact of the edits has not been assessed in terms of their reach. The purpose of editing the pages is to improve the health information retrieved by the public. Therefore, examining the page views since the edits is a necessary part of that process.

The aim of this project was to enhance Wikipedia health pages using high-quality, current research findings and track the persistence of those edits and the number of page views after the changes. Page views are reflective of public interest in a topic, with more page views demonstrating greater public interest. The Touro College Doctor of Physical Therapy (PT) program initiated Wikipedia Editathons supported by a grant from Consumer Reports magazine to add high-quality health information to existing Wikipedia pages. By adding accurate, current, best evidence to specific Wikipedia pages in layman's terms, enhanced information can be widely disseminated. Consumers benefit by using information that can guide evidence-based health decisions. Clinicians benefit by engaging with patients who are well informed, prepared to partner in their health care process and aware of self-care strategies that may promote well-being. It is important to note that increasing page readership was not a goal of this initiative, as attracting readers to specific pages was beyond the scope of this project.

Methods

The Editathon is a 3-hour event that is embedded within an existing course in the PT curriculum. The goal was for each student to add, at minimum, one claim backed with one citation to one Wikipedia page.

Preparing for the Editathon

The preparation for the Editathon involved several steps prior to the day of the actual event. First, we educated the students on how to search the literature and how to assess quality evidence. All students received 6 hours of training from two experienced physical therapy researchers on where to search, how to search, and how to assess the quality of sources within their physical therapy coursework. The next step was for each student to create a personal Wikipedia account and username. Simultaneously, we set up a program page under the "programs and events" dashboard in Wikipedia (meta.wikimedia.org/wiki/Programs_&_Events_Dashboard) listing all of the editors in the cohort.

Working closely with a resident Wikipedian (Wikipedia expert), we then used Wikipedia's tracking and algorithms to identify

XSL•FO
RenderX

the pages with the greatest need for updates, defined as pages that were frequently visited, with high page views, but with a low number of supporting citations. The focus on frequently visited pages suggests that these topics are of great interest to the public, and the lack of citations demonstrates the need for editorial update. While the number of citations tracked by Wikipedia does not infer quality of those citations on those health pages, our directive was to add high-quality findings from recent (within 5 years when possible) systematic reviews. Those health pages that could be enhanced by physical therapy information were specifically identified for inclusion and added to the program's event dashboard. It should be noted that some topics that are listed on our dashboard are closely related. This is because some common conditions are referred to by different terms and therefore have distinct Wikipedia pages. One such example is back pain and low back pain. The presence of both pages may foster greater access to the public based on preferred user terminology.

We established small working student groups and each group chose a topic of interest. Groups began by reading the chosen Wikipedia page looking for any PT-related information or lack thereof as well as the accuracy of the information and the recency and quality of the citations. While reading the chosen page, students asked themselves the following as they relate to PT: (1) What would I have liked to know more about when reading the Wikipedia entry? (2) What more might others who seek out this Wikipedia page like to know about this topic? (3) Is there more current evidence that should be added or more recent references that might better support the topic? We instructed students to locate 3 to 5 systematic reviews (Cochrane Reviews when available) that were less than 5 years from publication and related to the topic of interest. Of primary importance was that the reference could provide information that may complement, update, or correct the existing Wikipedia entry and better inform the reader of this topic. The preparatory process concluded with the reading, synthesis, and extraction of information to be added to the Wikipedia pages. Edits were reviewed by faculty and finalized prior to the Editathon.

Executing the Editathon

During the first half of the Editathon, a Wikipedia expert educated students on the scope, infrastructure, and metrics maintained by Wikipedia. The students were instructed on the editing process for adding information to a Wikipedia page and adding citations for these edits. The goal of the day was for each student to contribute a minimum of one edit ranging from a sentence to a paragraph along with corresponding references. The Wikipedian and faculty agreed that no deletions were to be made of existing Wikipedia entries during this Editathon. If students identified information on a page that contradicted current evidence, they were advised to use a format that recognized the previous information while providing updated accurate information along with supporting citations (for example, "While previous conventional wisdom suggested [erroneous information], more recent evidence suggests [current, accurate information]"). During the second half of the Editathon, students entered their edits and citations on their pages. As this process was finalized, students monitored their entries for responses in real time from the Wikipedia community at large by monitoring the edit link of their topic. At the conclusion of the Editathon, the facilitator shared suggestions for participants to follow-up on their changes, such as returning to Wikipedia to receive feedback on their contributions from editors-at-large or view audience traffic for the edited articles. The facilitator also encouraged all workshop participants to continue to use their account and edit live Wikipedia articles independent of the Editathon.

Analysis

To evaluate the impact of our contributions on the dissemination of health information, we examined two specific factors: number of page views by the public since the time of the page edits and response to our contributions from the Wikipedia editing community including number and type of subsequent edits and persistence of the student contributions. Page views prior to the Editathon were used to identify pages with considerable traffic but were not considered in the outcome, as the primary purpose was to monitor the rate of dissemination from the time of the student contributions. Page views were recorded for 4 months for all pages following each Editathon. Page views were only tracked if the edit persisted. In addition, we monitored total page views from the inception of each of the 3 cohort Editathons through May 31, 2018. While time frames vary for follow-up as the Editathons took place in 3 different years, total views are relevant since our target was health information dissemination.

As no standard methodology exists to describe the impact of a new Wikipedia contribution, we developed our own systematic approach, which is easily reproducible. We ran the Editathon in the same 1-day format annually for 3 years. Our observation period began on the day the students submitted their edits, March 29, 2016, April 4, 2017, or February 5, 2018, and ended on May 31, 2018.

Using the history tab on each edited Wikipedia page, we identified the log of the student edits by student username and timestamp of the edit. We first insured that only those contributions that met the criteria of adding new health-related content with an acceptable citation to an academic source were included in our metrics. We then used the "compare selected revisions" link (also known as "track changes") located within the history tab to compare the text of a Wikipedia article as it existed at 2 different points in time—at the time it was added and on May 31, 2018, our end date.

All edits that we tracked were categorized as those that persisted (contribution remained on the page in original or modified format) or those that did not (contribution removed). For all edits that persisted, we further subdivided the responses into 3 categories based on how subsequent Wikipedia editors altered the content: (1) additions (student information was enhanced), (2) partial deletions, or (3) copy edits. We also tracked the number of subsequent edits and the number of editors who made those edits. To assess our impact on dissemination of health information for all edits that persisted, we used the "page view statistics" link to access the count of Web traffic on each Wikipedia page. We began our page view counts from the date of the Editathon through a 4-month follow-up period and monitored page counts for all Editathons collectively through May 31, 2018, the end of the observation period.

XSL•FO
RenderX

Results

As of the end of the observation period, 98 PT students in 3 different cohorts engaged in Editathons. Students edited 26 Wikipedia articles, with 7 articles in the March 2016 cohort, 8 articles in the April 2017 cohort, and 11 articles in the February 2018 cohort. Of these 26 articles, 24 articles met the criteria of having a student add health information and a relevant citation. The other 2 articles involved student copy editing without making a claim or adding a citation. Of the 24 edits, 22 persisted at the end of the observation period.

Wikipedia's in-platform communication system solicits a diverse base of editors to check recent editorial changes. From the date of edit through May 31, 2018, the 24 articles received between 2 and 556 edits each. Edits include those made in direct response to the student additions and other changes to the content between the time of student entries through the completion of the observation period. On average, each article received 90 subsequent edits, with a median of 36 edits per target article. While time frames vary for follow-up, as the Editathons took place in 3 different years, our target was sustainability of student entries and thus total edits are relevant. Subsequent edits may be influenced by newly published or conflicting research findings. In 17 of the 24 cases, the Wikipedia community editorial response was entered within 1 day, which is typical for new content submissions. This demonstrates the attentiveness of the Wikipedia editors at large, who monitor for changes and quality control. For the remaining 7 articles, the response time ranged from 11 to 323 days, with a median response time of 23 days. The 24 articles had 16 different reviewing editors inserting a wide range of edits (Table 1).

Two examples that illustrate the outcome of the editorial process are highlighted. In the first example, a student added 2 sentences related to the evidence regarding surgery and bracing for carpal tunnel syndrome to the Carpal Tunnel Wikipedia page and cited a systematic review from 2015 during the March 29, 2016, Editathon (Figure 1). The original contribution is on the left and the version of that page as of May 31, 2018, is on the right.

In this example, a Wikipedia community reviewer changed the student's use of "an orthosis" to "a brace" to meet Wikipedia's Manual of Style for Medicine, which targets a lay audience. Users operating semiautomated Wikipedia copyediting tools made subsequent edits to add the date of publication and a PubMed identifier (PMID) to the citation and convert a hyphen to an en dash. For our statistics on this contribution, we judged that the student edits persisted and noted all of the changes described above as copy edits. There were no sentence or citation deletions noted for this entry.

Table 1. Editorial activity on Wikipedia within 24 hours of student contributions.

Distinct Wikipedia editors	Instances when editor responded within 24 hours
Editor 1	10
Editor 2	5
Editor 3	2
Editor 4	2
Editors 5-16	12 (1 per editor)

Figure 1. Edits to the Wikipedia page titled Carpal Tunnel Syndrome at the editing event on March 29, 2016.

Original student edit showing new material added on March 29, 2016	Version at the end of the observation period, May 31, 2018
Surgery is more beneficial in the short term to alleviate symptoms (up to six months) than wearing an orthosis for a minimum of 6 weeks. However, surgery and wearing an orthosis resulted in similar symptom relief in the long term (12-18 month outcomes).	Surgery is more beneficial in the short term to alleviate symptoms (up to six months) than wearing an orthosis for a minimum of 6 weeks. However, surgery and wearing **a brace** resulted in similar symptom relief in the long term (12–18 month outcomes).
D'Angelo, Kevin; Sutton, Deborah; Côté, Pierre; Dion, Sarah; Wong, Jessica J.; Yu, Hainan; Randhawa, Kristi; Southerst, Danielle; Varatharajan, Sharanya. "The Effectiveness of Passive Physical Modalities for the Management of Soft Tissue Injuries and Neuropathies of the Wrist and Hand: A Systematic Review by the Ontario Protocol for Traffic Injury Management (OPTIMa) Collaboration." Journal of Manipulative and Physiological Therapeutics. 38 (7): 493–506. doi:10.1016/j.jmpt.2015.06.006.	D'Angelo, Kevin; Sutton, Deborah; Côté, Pierre; Dion, Sarah; Wong, Jessica J.; Yu, Hainan; Randhawa, Kristi; Southerst, Danielle; Varatharajan, Sharanya (2015). "The Effectiveness of Passive Physical Modalities for the Management of Soft Tissue Injuries and Neuropathies of the Wrist and Hand: A Systematic Review by the Ontario Protocol for Traffic Injury Management (OPTIMa) Collaboration." Journal of Manipulative and Physiological Therapeutics. 38 (7): 493–506. doi:10.1016/j.jmpt.2015.06.006. PMID 26303967.

XSL•FO
RenderX

The second example (see Figure 2) describes an edit added to the Fall Prevention health page from the Editathon held on February 5, 2018. The student summarized information from a 2017 systematic review demonstrating that resistance training leads to enhanced functional mobility such as improvement in balance and reduction in falls. The student contributed to the page with 5 separate additions of new information that would be useful to someone seeking information on how to prevent falls and added the systematic review as the reference for each edit. The next day, a Wikipedia editor noted that the systematic review was listed as a new source each time and posted as citations 10, 11, 12, 13, and 14. The editor combined the repeated citations but did not make any changes to the actual content of the edits. At the most recent review of the page on May 31, 2018, all the edits persisted and the number of the citations, originally 10, was now 33, indicating that 23 more references had been added to the page above the section edited by the student.

For our statistics on this contribution, we judged this contribution as follows: the student edits persisted and the contribution underwent copy edits. The total number of edits by the Wikipedia community to the 22 persisting contributions are described in Table 2.

Figure 2. Edits to the Wikipedia page titled Fall Prevention at the editing event on February 5, 2018.

Original student edit showing new material added on Feb. 5, 2018	Version at the end of the observation period, May 31, 2018
After the age of 50, adults experience a decrease in muscle mass, a condition known as sarcopenia, by approximately 2% every year [10].	After the age of 50, adults experience a decrease in muscle mass, a condition known as sarcopenia, by approximately 2% every year [33].
A systematic review concluded that resistance training can slow down the rate of loss in muscle mass and strength. It has been recommended that older adults participate in resistance training 2-3 times a week to weaken the effects of sarcopenia [11].	A systematic review concluded that resistance training can slow down the rate of loss in muscle mass and strength. It has been recommended that older adults participate in resistance training 2-3 times a week to weaken the effects of sarcopenia [33].
Exercise as a single intervention has been shown to prevent falls in community dwelling older adults. A systematic reviews suggests that having an exercise regimen that includes challenging balance workouts for three or more hours per week results in a lesser chance of falling [12].	Exercise as a single intervention has been shown to prevent falls in community dwelling older adults. A systematic review suggests that having an exercise regimen that includes challenging balance workouts for three or more hours per week results in a lesser chance of falling [33].
Resistance training has been shown to be beneficial beyond fall prevention, as it also helps improve functional mobility and activities of daily living such as walking endurance, gait speed, and stair climbing [13].	Resistance training has been shown to be beneficial beyond fall prevention, as it also helps improve functional mobility and activities of daily living such as walking endurance, gait speed, and stair climbing [33].
Research explains that this significant increase in performance can be accomplished after the age of 90. In order for older adults to gain confidence in resistance training, which may ultimately lead to fall prevention effects, they need to obtain the recommended amount of daily activity [14].	Research explains that this significant increase in performance can be accomplished after the age of 90. In order for older adults to gain confidence in resistance training, which may ultimately lead to fall prevention effects, they need to obtain the recommended amount of daily activity [33].

Table 2. Summary of edits to 22 Wikipedia pages with persisting contributions.

Change type	Amount (n)	Description of change
Additions	10	Enhanced student prose or replaced a student citation with an alternative relevant citation
Partial deletions	6	Deleted at least one sentence of prose
	3	Shortened student explanation
	5	Kept at least one citation and deleted at least one citation
Copy edits	21	Changes to software code, punctuation, or citation format

XSL•FO
RenderX

Table 3. Page views from time of contribution through May 31, 2018, and within 4 months after the edit (for years 2016 and 2017).

PT[a]-relevant Wikipedia page	Date of initial edit	Page views through May 31, 2018	Page views at 4 months from entry
Back pain	March 29, 2016	700,575	98,613
Carpal tunnel syndrome	March 29, 2016	2,374,443	357,717
Developmental coordination disorder	March 29, 2016	684,635	120,323
Patient education	March 29, 2016	43,405	5871
Physical exercise	March 29, 2016	167,147	175,413
Urinary incontinence	March 29, 2016	774,051	102,126
Sports injury	March 29, 2016	150,767	19,645
Sports injury[b]	April 4, 2017	—	19,903
Chronic pain	April 4, 2017	165,432	49,267
Joint dislocation	April 4, 2017	89,410	23,080
Physical therapy	April 4, 2017	666,181	191,090
Physical therapy education	April 4, 2017	51,875	17,894
Spinal cord injury	April 4, 2017	307,377	83,308
Spinal disease	April 4, 2017	37,006	9453
Telerehabilitation	April 4, 2017	10,182	2849
Angina	February 5, 2018	218,465	218,465
Cerebral palsy	February 5, 2018	539,542	539,542
Chronic obstructive pulmonary disease	February 5, 2018	484,912	484,912
Disabled sports	February 5, 2018	19,324	19,324
Fall prevention	February 5, 2018	5135	5135
Low back pain	February 5, 2018	162,590	162,590
Plagiocephaly	February 5, 2018	50,988	50,988
Prosthesis	February 5, 2018	100,478	100,478

[a]PT: physical therapy.

[b]The same topic was edited during two Editathons—page views represents sum total from time of initial edit.

Last, to measure impact using the Wikipedia Pageviews Analysis tool, we found that the 22 articles with persisting student edits received 7,803,920 page views (Table 3). The number of page views from date of the student entries through the first 4 months after the entries and through May 31, 2018, are shown. Four months was chosen to compare activity on each edited page for the same time frame across the different years. At the end of the observation period, each edited Wikipedia health page had an average of 354,724 page views while the median number of page views was 164,011. It is interesting to note the student-edited Wikipedia pages with the highest (Carpal Tunnel Syndrome and Urinary Incontinence) and lowest (Telerehabilitation and Fall Prevention) page views.

Discussion

Principal Findings

The objective of the Editathons within the PT program at Touro College was to disseminate best-evidence PT-relevant health information to the public using principles of health literacy. These principles include assuring that health information for the public is presented in an easy-to-understand format. Based on the persistence of student contributions and total page views, this initiative was successful in accomplishing that goal. Although the doctoral PT student editors had never done this before and may have anticipated editorial criticism or deletions [4], most student contributions persisted. This persistence demonstrates that the Editathon is an effective method for improving the quality of Wikipedia's health pages. It is important to emphasize that high-quality, timely information is presented in a format that is accessible to typical Wikipedia users. Wikipedia's history logs show that the Wikipedia community responded to project contributions in all 24 cases and usually within 1 day of the submission. In all cases, the edits enhanced the student contributions and added more information for the public to access. Examining the editorial response to these 24 health pages for three Editathons may not be sufficient to generalize how all reviewing happens in Wikipedia. However, these findings demonstrate that Wikipedia's internal communication systems facilitate the editorial process by matching reviewers to incoming edits and providing guidelines to assist reviewers in the editorial process [8,9].

XSL•FO
RenderX

An added benefit of the student edits is that these changes evoked interest in the health page by the editorial community. This in turn led to additional edits and inclusion of citations that further enhanced the content quality, improving the delivery of timely information to the reader. This is evident in Figure 2, where it is noted that 23 additional citations were added between February 5, 2018 (the date of the Editathon), and May 31, 2018 (the end of the observation period).

As deficiencies in the translation of research into practice are widely recognized [10-12], the Wikipedia model is one way to close this gap, bringing current, relevant information to the masses. Taking on the role of Wikipedia editor compelled students to gather and synthesize current, best-evidence information, promoting learning [13]. Living reviews, where current research findings are added to existing best-evidence documents, have the potential to expedite the dissemination of current research findings and translation of knowledge to practice. Ultimately, the public benefits from these collective efforts. Despite this systematic approach, there are those who continue to reject Wikipedia as a useful health resource. While one may argue that editors cannot capture all misinformation, and organizers are aware of this limitation [14], Wikipedia remains one of the most accessed health information sites [1].

Limitations

There are several limitations to this study that must be considered. One limitation is the fact that we chose the same end date for tracking page views, and thus the length of follow-up time for each Editathon varied. Edits that existed longer than others could potentially have had more views. However, to address this limitation we also tracked the number of page views during a 4-month period following each Editathon for a more uniform comparison. As our goal was to report on the reach/dissemination of the added information, the final page view count is relevant. In addition, since there was no established methodology for monitoring the persistence of edits, we devised a systematic approach that has not yet been reproduced. Finally, it is not possible to ascertain if the viewers read the part of the page with the added health information or if the information was useful to the reader. Our data only tells us that the page was viewed.

Conclusion

Physical therapists and most health care professionals value the importance of patient education for optimizing outcomes. Patient education is key to promoting self-management, knowing when and where to seek care and which treatments are helpful or harmful, and facilitating compliance with lifestyle and medical regimens that promote optimal health. Through three Editathons, 98 students added best-evidence information to 24 Wikipedia pages that were viewed nearly 8 million times with an average of approximately 350,000 views per page. The Wikipedia Editathon is an excellent way of bridging health technology with best-evidence approaches to care and bringing accurate, useful information to the public. Understanding the editing infrastructure built into the Wikipedia framework allows clinicians to effectively use this vast resource. The Editathon is inexpensive and has far-reaching and lasting impact. Future efforts may include interprofessional collaboration and assessment of the utility of health information in Wikipedia target articles.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Wikipedia article page views.

[PDF File (Adobe PDF File), 16KB - jmir_v21i3e12450_app1.pdf]

References

1. Heilman JM, West AG. Wikipedia and medicine: quantifying readership, editors, and the significance of natural language. J Med Internet Res 2015;17(3):e62 [FREE Full text] [doi: 10.2196/jmir.4069] [Medline: 25739399]
2. Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. Health Info Libr J 2017 Dec;34(4):268-283. [doi: 10.1111/hir.12192] [Medline: 29045011]
3. Shafee T, Masukume G, Kipersztok L, Das D, Häggström M, Heilman J. Evolution of Wikipedia's medical content: past, present and future. J Epidemiol Community Health 2017 Nov;71(11):1122-1129 [FREE Full text] [doi: 10.1136/jech-2016-208601] [Medline: 28847845]
4. Azzam A, Bresler D, Leon A, Maggio L, Whitaker E, Heilman J, et al. Why medical schools should embrace Wikipedia: final-year medical student contributions to Wikipedia articles for academic credit at one school. Acad Med 2017 Dec;92(2):194-200 [FREE Full text] [doi: 10.1097/ACM.0000000000001381] [Medline: 27627633]
5. Hunter JA, Lee T, Persaud N. A comparison of the content and primary literature support for online medication information provided by Lexicomp and Wikipedia. J Med Libr Assoc 2018 Jul;106(3):352-360 [FREE Full text] [doi: 10.5195/jmla.2018.256] [Medline: 29962913]
6. Maggio LA, Willinsky JM, Steinberg RM, Mietchen D, Wass JL, Dong T. Wikipedia as a gateway to biomedical research: the relative distribution and use of citations in the English Wikipedia. PLoS One 2017;12(12):e0190046 [FREE Full text] [doi: 10.1371/journal.pone.0190046] [Medline: 29267345]

XSL•FO
RenderX

7. Aitken M, Altmann T, Rosen D. Engaging patients through social media: Is healthcare ready for empowered and digitally demanding patients?. Parsippany, NJ: IMS Institute; 2014 Jan. URL: http://magazine.imshealth.it/wp-content/uploads/2015/09/IIHI_Social_Media_Report.pdf [WebCite Cache ID 73sYVQVGt]
8. WikiProject Medicine/Tools.: Wikipedia; 2017 Jun 03. URL: https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine/Tools [accessed 2018-11-12] [WebCite Cache ID 73seGjFTr]
9. Recent Changes Patrol.: Wikipedia; 2018 Aug 04. URL: https://en.wikipedia.org/wiki/Wikipedia:Recent_changes_patrol [accessed 2018-11-12] [WebCite Cache ID 73sduFqkr]
10. Hanney SR, Castle-Clarke S, Grant J, Guthrie S, Henshall C, Mestre-Ferrandiz J, et al. How long does biomedical research take? Studying the time taken between biomedical and health research and its translation into products, policy, and practice. Health Res Policy Syst 2015 Jan 01;13:1 [FREE Full text] [doi: 10.1186/1478-4505-13-1] [Medline: 25552353]
11. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011 Dec;104(12):510-520 [FREE Full text] [doi: 10.1258/jrsm.2011.110180] [Medline: 22179294]
12. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012 May 31;7:50 [FREE Full text] [doi: 10.1186/1748-5908-7-50] [Medline: 22651257]
13. Shane-Simpson C, Che E, Brooks P. Giving psychology away: implementation of Wikipedia editing in an introductory human development course. Psychol Learning Teaching 2016;15(3):268-293. [doi: 10.1177/1475725716653081]
14. Knowledge Integrity.: Wikpedia; 2018 Aug 04. URL: https://meta.wikimedia.org/wiki/Knowledge_Integrity [accessed 2018-11-12] [WebCite Cache ID 73sdU7yPv]

Abbreviations

PT: physical therapy
PMID: PubMed identifier

Edited by G Eysenbach; submitted 09.10.18; peer-reviewed by Z Rosen, K Friel, J Willinsky, B Arnoldussen; comments to author 31.10.18; revised version received 14.11.18; accepted 15.11.18; published 18.03.19.

Please cite as:
Weiner SS, Horbacewicz J, Rasberry L, Bensinger-Brody Y
Improving the Quality of Consumer Health Information on Wikipedia: Case Series
J Med Internet Res 2019;21(3):e12450
URL: https://www.jmir.org/2019/3/e12450/
doi:10.2196/12450
PMID:30882357

©Shira Schecter Weiner, Jill Horbacewicz, Lane Rasberry, Yocheved Bensinger-Brody. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Detecting the Interdisciplinary Nature and Topic Hotspots of Robotics in Surgery: Social Network Analysis and Bibliometric Study

Lining Shen[1,2,3*], PhD; Shimin Wang[1,2*], MS; Wei Dai[1*], MS; Zhiguo Zhang[1,3], PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science & Technology, Wuhan, China

[2]Institute of Smart Health, Huazhong University of Science & Technology, Wuhan, China

[3]Hubei Provincial Research Center for Health Technology Assessment, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Lining Shen, PhD
School of Medicine and Health Management
Tongji Medical College
Huazhong University of Science & Technology
No.13 Hangkong Road
Wuhan, 430030
China
Phone: 86 027 83692730
Fax: 86 027 83692727
Email: shenln@163.com

Abstract

Background: With the widespread application of a robot to surgery, growing literature related to robotics in surgery (RS) documents widespread concerns from scientific researchers worldwide. Although such application is helpful to considerably improve the accuracy of surgery, we still lack the understanding of the multidiscipline-crossing status and topic distribution related to RS.

Objective: The aim of this study was to detect the interdisciplinary nature and topic hotspots on RS by analyzing the current publication outputs related to RS.

Methods: The authors collected publications related to RS in the last 21 years, indexed by the Web of Science Core Collection. Various bibliometric methods and tools were used, including literature distribution analysis at the country and institution level and interdisciplinary collaboration analysis in the different periods of time. Co-word analysis was performed based on the keywords with high frequency. The temporal visualization bar presented the evolution of topics over time.

Results: A total of 7732 bibliographic records related to RS were identified. The United States plays a leading role in the publication output related to RS, followed by Italy and Germany. It should be noted that the Yonsei University in South Korea published the highest number of RS-related publications. Furthermore, the interdisciplinary collaboration is uneven; the number of disciplines involved in each paper dropped from the initial 1.60 to the current 1.31. Surgery; Engineering; Radiology, Nuclear Medicine, and Medical Imaging; and Neurosciences and Neurology are the 4 core disciplines in the field of RS, all of which have extensive cooperation with other disciplines. The distribution of topic hotspots is in imbalanced status, which can be categorized into 7 clusters. Moreover, 3 areas about the evolution of topic were identified, namely (1) the exploration of techniques that make RS implemented, (2) rapid development of robotic systems and related applications in surgery, and (3) application of a robot to excision of tissues or organs targeted at various specific diseases.

Conclusions: This study provided important insights into the interdisciplinary nature related to RS, which indicates that the researchers with different disciplinary backgrounds should strengthen cooperation to publish a high-quality output. The research topic hotspots related to RS are relatively scattered, which has begun to turn to the application of RS targeted at specific diseases. Our study is helpful to provide a potential guide to the direction of the field of RS for future research in the field of RS.

(J Med Internet Res 2019;21(3):e12625) doi:10.2196/12625

XSL•FO
RenderX

KEYWORDS

robotics; robotic surgery; interdisciplinary collaboration; topic hotspot; topic bursts; co-word analysis; bibliometric analysis; bibliometrics; social network analysis; robotic surgical procedures; laparoscopy

Introduction

Background

A robot is a mechatronic device that can be programmed to perform some tasks automatically, the emergence of which has significantly improved people's quality of life. Apart from revolutionizing the manufacturing sector, robots have now found their way out of the factory and into applications such as agriculture, aerospace, and education [1], with no exception of medicine. One of the advantages of surgical robots over traditional surgery is that they are smart and precise and can accomplish their purpose more accurately. Dexterous robots have emerged in the last decade in response to the need for fine motor control assistance in applications as diverse as surgery, undersea welding, and mechanical manipulation in space [2]. For example, robots have also been integrated into operating rooms around the world and have enabled or improved many new minimally invasive surgical procedures [3-5].

Growing literature related to robotics in surgery (RS) documents widespread concerns from scientific researchers worldwide. As described in the literature, robots were introduced to the medical industry in the last century, initially for auxiliary work such as nursing and image transmission, so that doctors can get better examination results [6]. For example, installing a camera on the robotic arm and then sending the captured results to the screen can help the doctor to perform a more accurate examination for patients. Some researchers pointed out that surgical robots have also been used to try to solve some intractable diseases and increase the accuracy and safety of surgery [7,8]. Furthermore, with the development of information technology, the application of surgical robots has not only been limited to examination functions. However, other researchers believed that long-term feasibility still needs further assessment, although the treatment effect is acceptable in the short term [9]. In addition, the greatest surgical innovation of the past 3 decades has been the emergence of minimally invasive surgery in which many surgical robots are currently used. Therefore, the effect of minimally invasive surgery will also be greatly improved by means of a magnified view and improved ergonomics and dexterity provided by robotic platforms [6], which can reduce patient discomfort, costs, and hospital time [1]. However, limited data are available regarding safety and efficacy [10]. Therefore, some researchers compared the difference between robotic surgery and traditional surgery in detail [11] and evaluated the clinical effectiveness of surgical robots through function and outcomes [12].

Obviously, scientific researchers mainly focused on the clinical applications of surgical robots. However, to the best of our knowledge, little is known about the research situation, interdisciplinary nature, and research hotspots related to RS from the perspective of bibliometrics, which can help us comprehensively understand the process of the evolution of the related disciplines and research themes involved in RS.

Objectives

This study analyzed RS-related publications from the perspective of bibliometrics to address the above limitations. Specifically, the purposes of this study were listed as follows:

1. How did the literature in the field of RS be distributed worldwide?
2. What was the interdisciplinary collaboration of RS in the last decades?
3. What were the topic hotspots and evolution process in the field of RS?

Methods

Sample and Data Collection

In this study, we chose publications indexed in the Web of Science Core Collection (WoSCC) database as the data source. As WoSCC adheres to a strict evaluation process and provides the most influential, relevant, and credible information, it is most suitable for subsequent bibliometric analysis in this study [13].

To fully retrieve RS-related publications, combining with the above literature review on RS, we constructed the following search strategy: #1 TS=((“robot* AND *assist*”) AND *surg*); #2 TS=((“robot* AND *guid*”) AND *surg*); #3 TI=(robot*AND *surg*); #4 #1 OR #2 OR #3. Moreover, the document type was limited to article and review; the time span of publications covered the period from 1986 to 2017.

Subsequently, a total of 10,087 bibliographic records were identified and downloaded on September 20, 2018. To perfect the research, the main inclusion and exclusion criteria were formulated after 2 researchers independently reviewed and evaluated the 1000 pilot bibliographic records. The inclusion criteria were as follows: (1) the content of the papers primarily focused on RS and (2) all study designs. The main exclusion criteria were as follows: (1) the record had no subject categorizations or keywords (eg, book review and notification), (2) the study merely mentioned robotic surgery as one of the surgical approaches [14], and (3) the content of the research did not focus on RS but patients with no robotic surgery [15]. Any discrepancies were discussed until consensus was reached in this process. Subsequently, 1 researcher reviewed the remaining records according to the above criteria. Finally, a total of 7732 bibliographic records were obtained for further bibliometric analysis, with 2355 inappropriate or irrelevant records removed, so as to elucidate the interdisciplinary nature and research topic hotspots in the field of RS internationally.

Design of Data Analysis Method

There are various indicators used in the study to better demonstrate distribution of the literature. The total local citation score (TLCS) and the total global citation score (TGCS) were calculated in this study, which have been the key indicators capable of evaluating the relevance of each research paper in

XSL•FO
RenderX

our sample [16]. TLCS refers to the number of times that a set of papers included in a collection has been cited by other papers within the same collection, whereas TGCS refers to the number of times that a set of papers included in a collection has been cited in the WoSCC [17]. In addition, the average global citation score (AGCS) is the mean value of TGCS. However, it should be noted that TLCS presents the important papers in a chosen research area, whereas TGCS mainly displays the effects of the papers related to a chosen research area on the papers in the WoSCC [18]. Distribution of the literature was presented using the HistCite tool, which is an analysis and visualization software that helps us to obtain information at the country and institution level [19]. Meanwhile, we divided the 21 years into 4 periods of time to exam the distribution at the country level.

Furthermore, interdisciplinarity and cross-disciplinarity have been buzzwords for the last few years, which are used to describe contributions from and collaborations among several or more disciplines. Interdisciplinary means that the content of research is not only a method or ability in a field but a field that involves more [20]. Through interdisciplinary research, we can more comprehensively understand the research content of a field. Interdisciplinary inevitably exists between disciplines, indicating that the scope involved in a certain field is constantly expanding [21]. Meanwhile, research areas constitute a subject categorization scheme that is shared by all Web of Science product databases. The literature indexed by WoSCC is assigned to at least 1 subject category, which is mapped to 1 research area [22]. VOSviewer—a software tool developed by Nees Jan van Eck and Ludo Waltman at Leiden University's Centre for Science and Technology Studies [23]—was employed to visualize the interdisciplinary collaboration on the basis of subject categorization of publication [24]. Each node represents a discipline, whereas the connection between nodes represents collaborations between disciplines. In addition, nodes with a close connection are assigned the same color to form their respective clusters. Furthermore, a co-occurrence matrix was generated by using the Bibliographic Item Co-occurrence Mining System (BICOMS) [25] to calculate the centrality, which includes degree centrality, closeness centrality, and betweenness centrality by using Ucinet6.6 [26]. Degree centrality is simply the number of tie of a given type that a node has; closeness is an inverse measure of centrality in the sense that large numbers indicate that a node is highly peripheral, whereas small numbers indicate that a node is more central; betweenness centrality is a measure of how often a given node falls along the shortest path between 2 other nodes [27]. Moreover, we analyzed the centrality in the different periods of time based on the top 5 centralities over the period from 1997 to 2017.

In addition, we used Cortext to visualize the evolution of individual disciplines and interdisciplinary clusters. The tubes layout represents the transformation of cluster of discipline over time [28-30]. The width of tubes represents the number of records in which they appear in the same cluster. Darker tubes mean more disciplines are shared between 2 consecutive time periods.

Finally, 3 stages were completed, as follows, regarding the analysis of research hotspots. First, BICOMS was employed to calculate the frequency of keywords. Subsequently, a total of 13,706 keywords were obtained and merged based on the following 4 criteria [31]: (1) merging some keywords into corresponding Medical Subject Headings terms using PubMed (eg, “gynaecology” and “lymphadenectomy” were merged into “gynecology” and “lymph node excision,” respectively); (2) unifying the uppercase and lowercase of some keywords (eg, “Laparoscopy” and “Bladder cancer” were changed to “laparoscopy” and “bladder cancer,” respectively); (3) standardizing the singular and plural of keywords (eg, “child” and “pediatric” were changed to “children” and “pediatrics,” respectively); and (4) merging some synonym keywords (eg, “minimal invasive surgery” and “MIS” were replaced by “minimally invasive surgery”). After merging, 90 keywords with frequencies not less than 40 were obtained.

Second, we used BICOMS to generate the 88×88 co-occurrence matrix of keywords with a frequency not less than 40. It is worth noting that we removed robotic surgery and surgical robot because they are our research object. Then, a social network map was drawn with respect to these 88 keywords by Ucinet6.6 and VOSviewer [26,32,33], which intuitively reflects the relationship between keywords of high frequency. The relative size of nodes is proportional to the frequency of keywords, whereas the relative width of lines is proportional to the correlation between keywords [34].

Third, we detected the burst strength of the cleaned keywords and drew a temporal bar graph for high-burst strength keywords. Burst strength depicts the intensity of the burst, that is, how great the change is in the word frequency that triggered the burst. Kleinberg burst detection algorithm [35] can recognize the sudden increase of word frequency over time and detect the burst of keyword popularity effectively. We chose Science of Science (Sci2) [36], which can implement this algorithm to find out the burst terms in the processed data and calculate the burst strength. Finally, 48 keywords with a burst strength of not less than 4 were obtained. However, these keywords may only be core keywords to a certain extent. Further screening by word frequency can improve the quality of core keywords. The higher the number of keyword frequency, the more likely it is to become a hot topic in future. Then we drew a temporal visualization map of 26 keywords with a frequency no less than 40 and burst strength more than 4 by Sci2 [37]. Each keyword has its own starting and ending time, and the area of each bar reflects its burst strength.

Results

Literature Distribution

A total of 22,470 authors were identified with 7732 papers, which are affiliated to 4721 institutions from 74 countries. All of these papers were published in 1030 journals with 105,835 citations. In total, 12 publication languages were included, of which, English ranks the first, followed by German and French. It can be found that the quantity of literature related to RS is growing rapidly over time, and the United States takes the top spot at every time slice (Figure 1). It should be noted that the developed countries had contributed to the majority of the publication, although China is among the top 10 (for the publication output of top 10 countries, see Multimedia Appendix

XSL•FO
RenderX

1). In addition, South Korea ranked the fifth in the publication output from 2008 to 2012 and 2013 to 2017. The distribution of institutions is shown in Table 1. Obviously, Yonsei University in South Korea takes the first place, with the highest TLCS and TGCS, followed by Cleveland Clinic and Mayo Clinic. Memorial Sloan-Kettering Cancer Center has the highest AGCS, with high academic influence and collaboration in RS-related research, followed by Stanford University.

Figure 1. Top 5 countries of robotics in surgery-related research in each period of time.

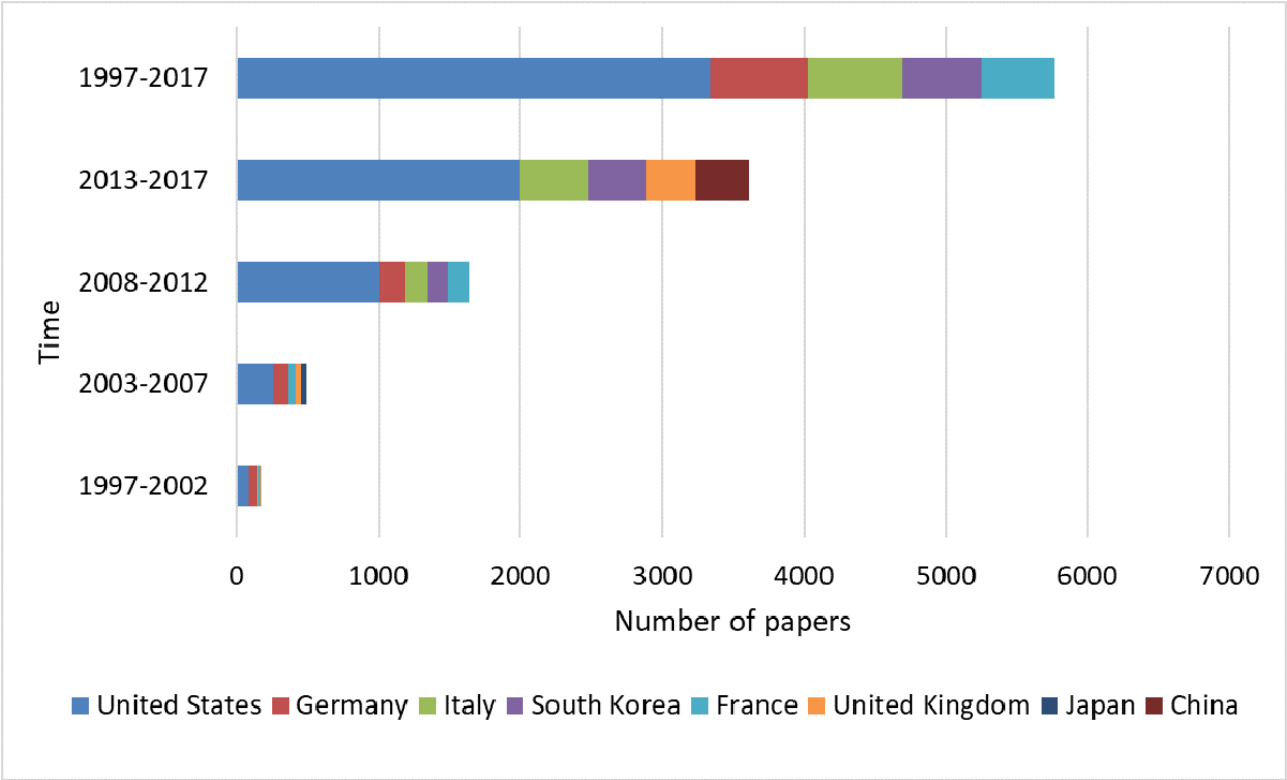


Table 1. Distribution of the top 10 institutions with robotics in surgery-related research.

Number	Institution	Recs[a]	Publication, %	Cumulative percentage	TLCS[b]	TGCS[c]	AGCS[d]
1	Yonsei University	197	2.55	2.55	1498	3579	18.17
2	Cleveland Clinic	167	2.16	4.71	990	3268	19.57
3	Mayo Clinic	111	1.44	6.15	978	2674	24.09
4	The Johns Hopkins University	104	1.35	7.50	701	3485	33.51
5	University of Pittsburgh	102	1.32	8.82	416	2108	20.67
6	Imperial College London	98	1.27	10.09	424	1799	18.36
7	Harvard University	96	1.24	11.33	714	2759	28.74
8	Stanford University	90	1.16	12.49	908	3327	36.97
9	Vanderbilt University	83	1.07	13.56	585	1986	23.93
10	Memorial Sloan-Kettering Cancer Center	78	1.01	14.57	867	3262	41.82

[a]Recs: number of published papers.

[b]TLCS: total local citation score.

[c]TGCS: total global citation score.

[d]AGCS: average global citation score.

XSL•FO
RenderX

Interdisciplinary Nature

Visualization of the Interdisciplinary Collaboration on Robotics in Surgery–Related Research

The number of disciplines is increasing over the 4 time slices; a total of 91 disciplines are involved as shown in Table 2. The disciplines covered by RS have expanded from 34 in the first 6 years to 85 in the last 5 years. A downward trend is presented for the average number of disciplines in each paper, from 1.60 to 1.31.

Figures 2 to 6 visualize the interdisciplinary collaboration for RS-related research for each period of time and overall (for detailed clusters, see Multimedia Appendix 2). The links between disciplines have been clearly shown in proportion, and the collaboration of disciplines within the cluster is significantly more than that between clusters. Although the number of clusters changed over time, there were still several clusters in each period of time, each of which has 1 or several major disciplines. For example, Surgery, Oncology, Engineering, and Urology and Nephrology appear most frequently, each of which leads a cluster in each of the maps. Thus, such a cluster is the main research direction in the field of RS. In general, the main disciplines from 1997 to 2017 were Surgery and Urology and Nephrology, with a frequency of 2802 and 1837, respectively, accounting for 45.06% (4639/10295) of the total frequency in the period of time.

Table 2. The overall distribution of disciplines and clusters.

Time period	Time span	Number of papers	Number of disciplines	Number of clusters	Number of discipline occurrences in papers	Mean disciplines in each paper
1997-2002	6	217	34	10	347	1.60
2003-2007	5	665	55	10	974	1.46
2008-2012	5	1985	64	15	2563	1.29
2013-2017	5	4865	85	12	6411	1.31
1997-2017	21	7732	91	13	10295	1.33

Figure 2. Interdisciplinary collaboration on robotics in surgery-related research from 1997 to 2002.

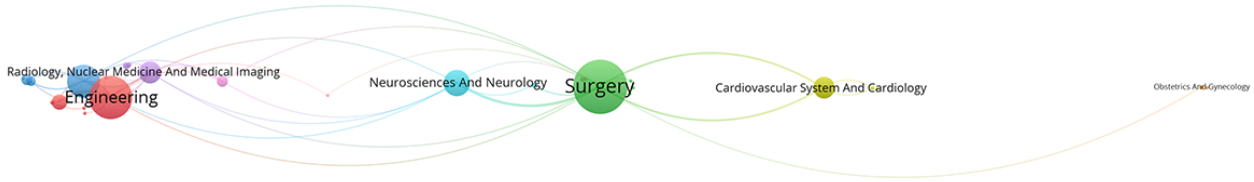
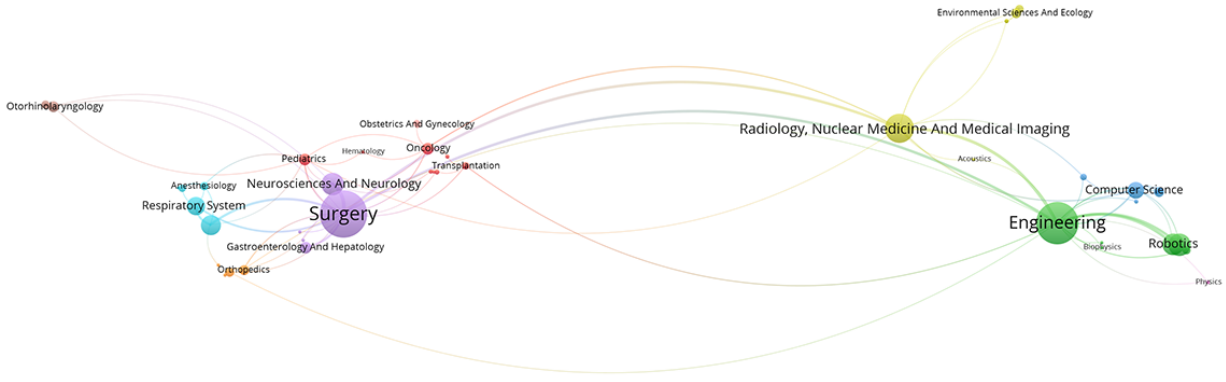


Figure 3. Interdisciplinary collaboration on robotics in surgery-related research from 2003 to 2007.



XSL•FO
RenderX

Figure 4. Interdisciplinary collaboration on robotics in surgery-related research from 2008 to 2012.

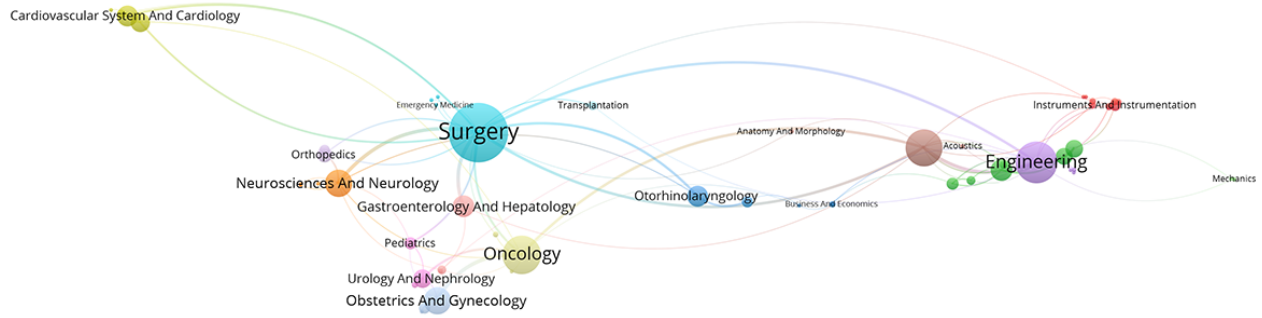


Figure 5. Interdisciplinary collaboration on robotics in surgery-related research from 2013 to 2017.

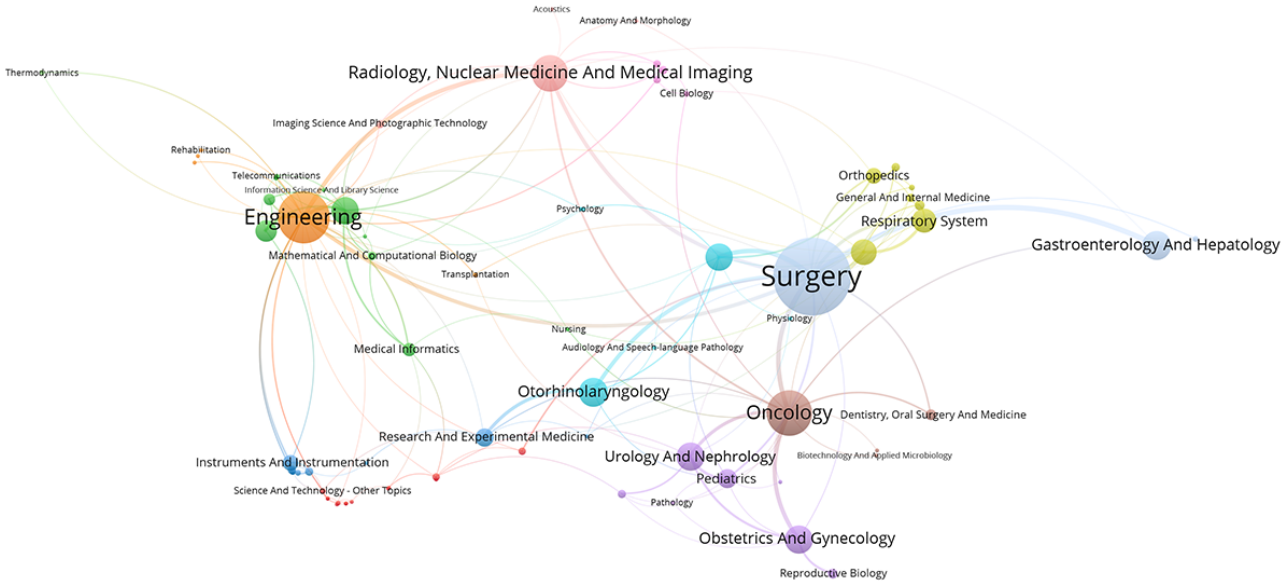
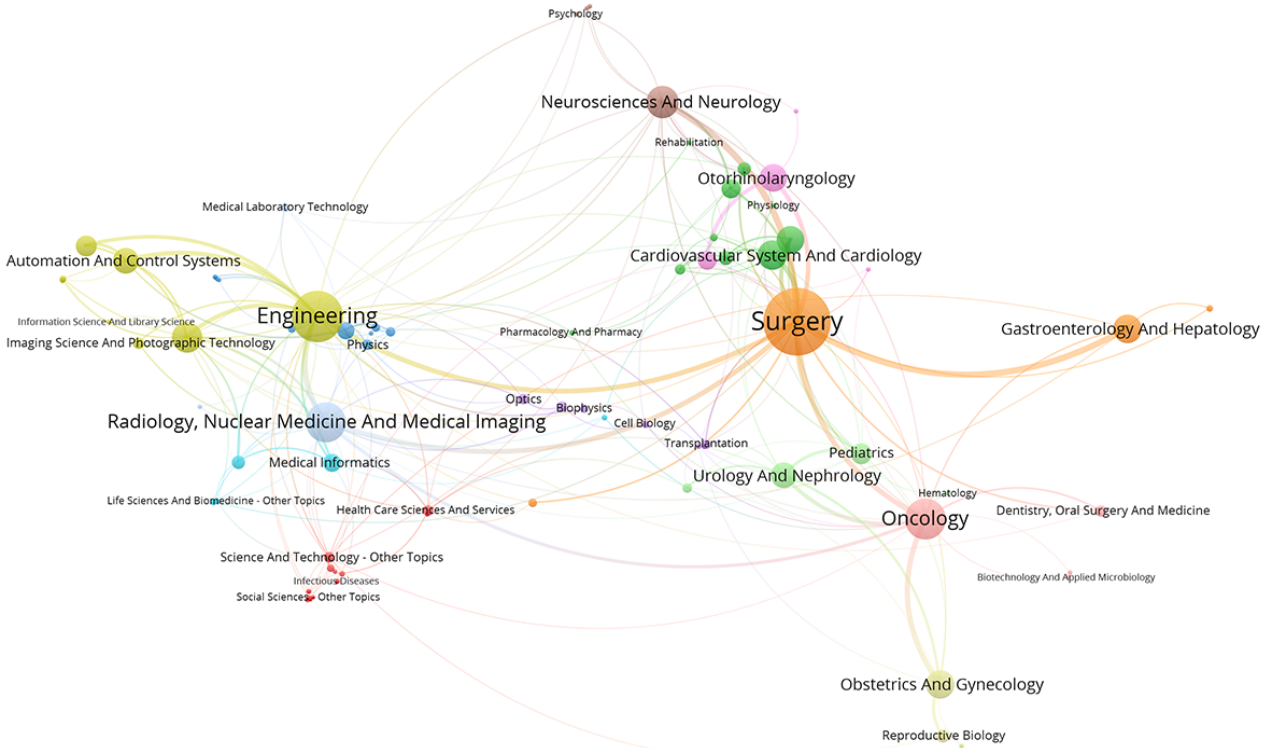


Figure 6. Interdisciplinary collaboration on robotics in surgery-related research from 1997 to 2017.



XSL•FO
RenderX

Network Analysis of Interdisciplinary Collaboration

Figures 7 to 9 show degree centrality, closeness centrality, and betweenness centrality for top 5 disciplines, which have held top 5 centralities over the years from 1997 to 2017 in different periods of time. The degree centrality of the top 5 disciplines is increasing continuously, whose trend is similar to the closeness centrality; although, there is a significant difference between them. The betweenness centrality of different disciplines intersected over time; there is no obviously upward or downward trend. From the perspective of degree centrality, the impact of Surgery on RS-related research is significantly higher than other disciplines. From the perspective of closeness centrality, Science and Technology played a more important role in RS-related research in the first 10 years. However, the evolution trend of the betweenness centrality is significantly different from the degree centrality and the closeness centrality, which shows that the position of top 5 disciplines, as an intermediary bridge, is not fixed.

Figure 7. Degree centrality of top 5 disciplines.

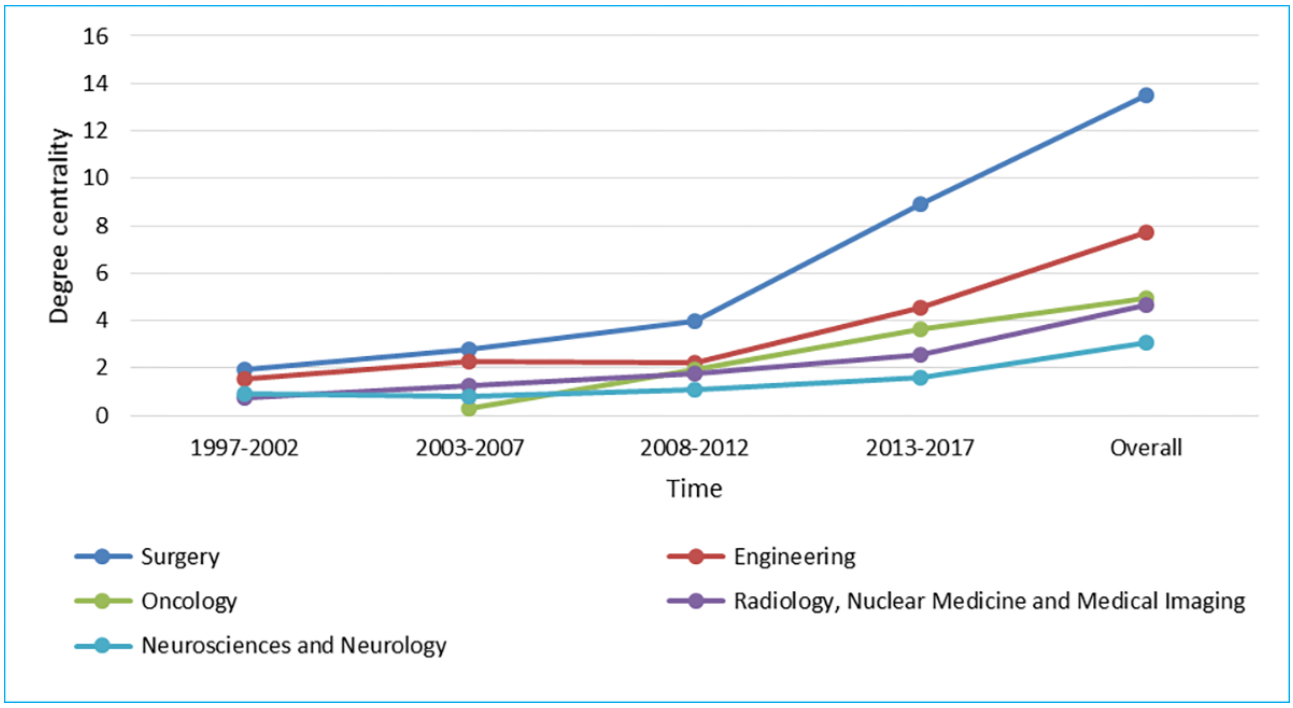
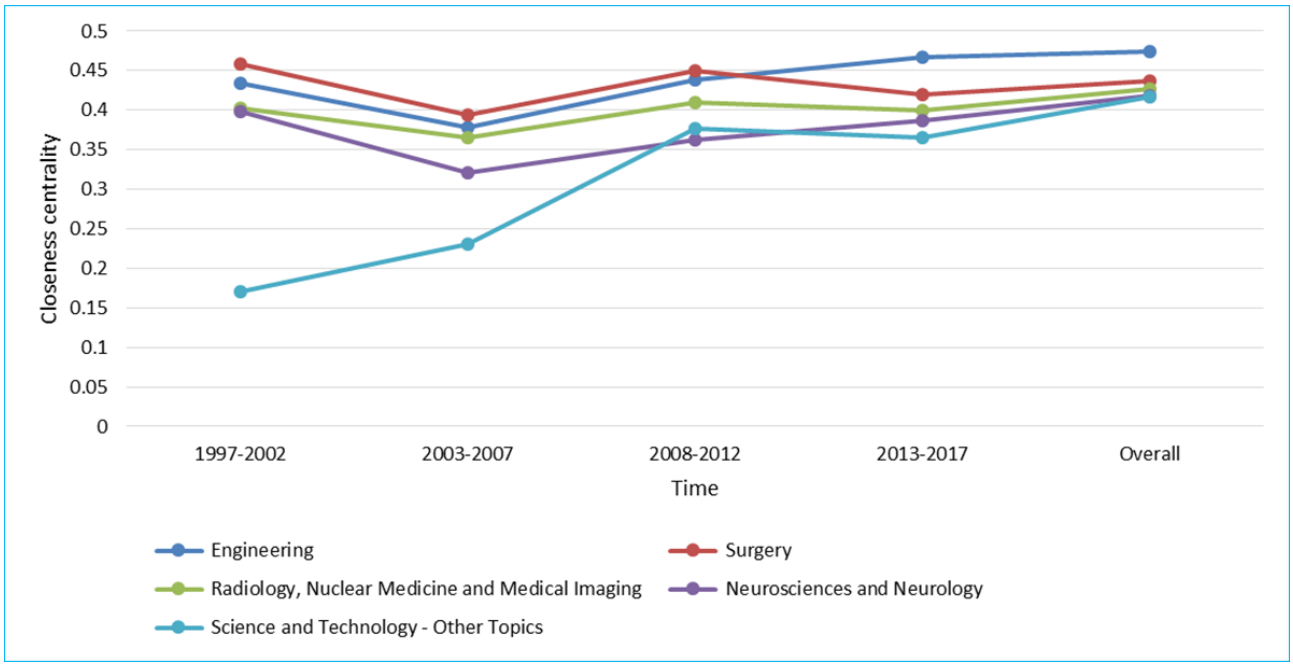


Figure 8. Closeness centrality of top 5 disciplines.



XSL•FO
RenderX

Figure 9. Betweenness centrality of top 5 disciplines.

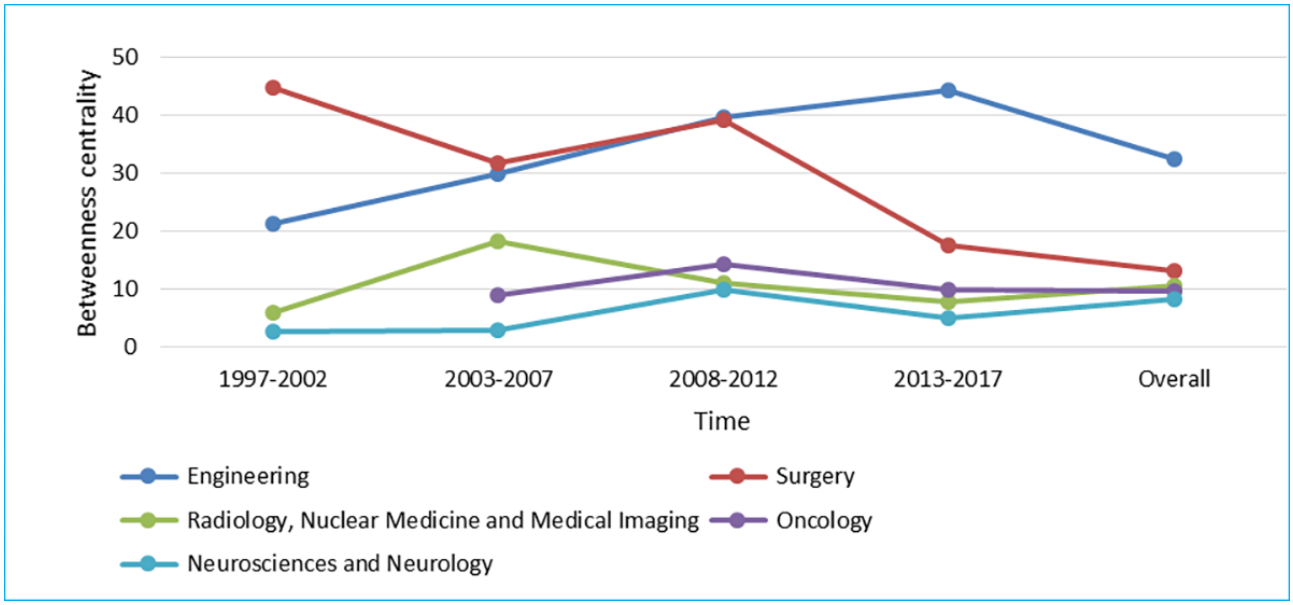
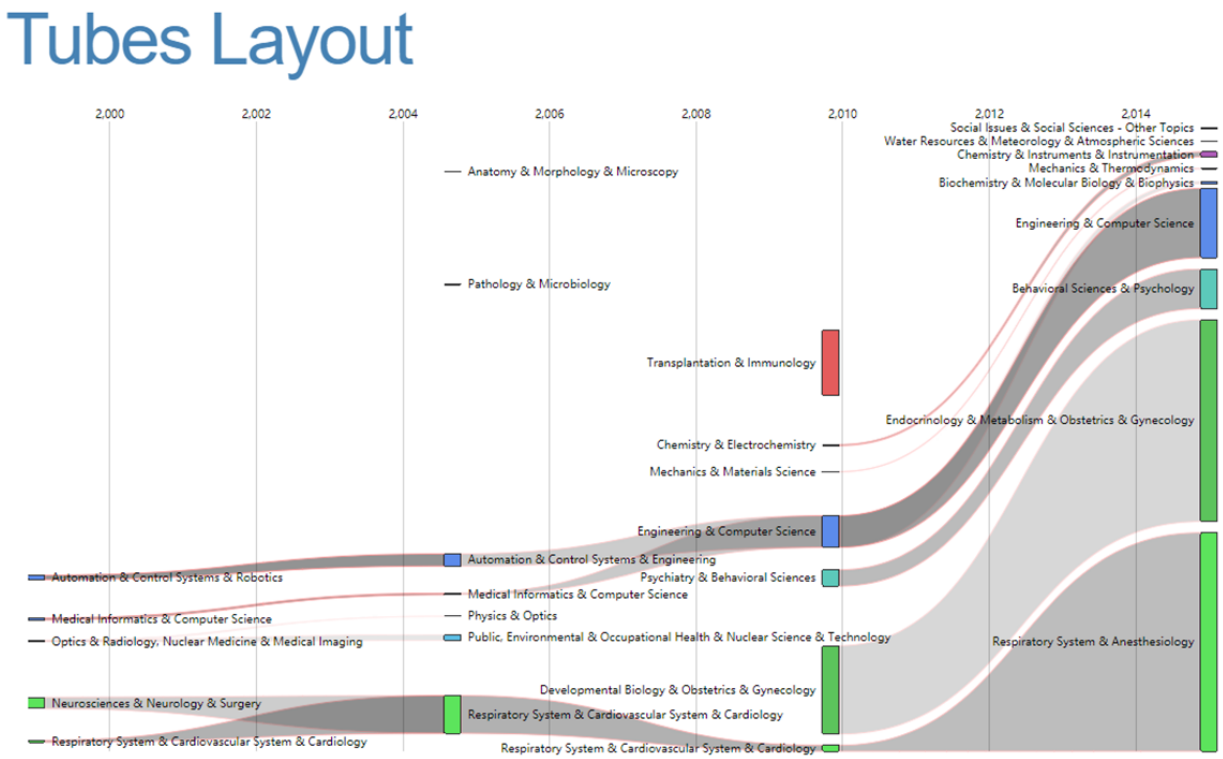


Figure 10. Evolution of discipline clusters over time.



Evolution of Discipline Clusters Over Time

Figure 10 shows the overall evolution of discipline clusters related to RS over time. The 4 columns represent the 4 periods of time above. For example, the fourth column represents the data from 2013 to 2017. The cluster related to Cardiovascular System and Cardiology was integrated with other disciplines to form a larger cluster during the period from 2003 to 2007, which was independent over the years from 1997 to 2002. Optics was separated from Radiology, Nuclear Medicine, and Medical Imaging from 2003 to 2007 to form an independent cluster with Physics. Obstetrics and Gynecology and other disciplines were merged into a larger cluster during the years from 2013 to 2017. Transplantation and other related disciplines constituted a relatively independent cluster in the time period from 2008 to 2012. Respiratory System was integrated with Anesthesiology to form a larger community with Anesthesiology from 2013 to 2017. In addition, the records in Psychology and Behavior Sciences are in an increasing trend from 2008 to 2017. It should be noted that Engineering and Computer Science were merged into a cluster in the period from 2008 to 2012, then gradually grew over time. In general, the results are consistent with the above analysis to some extent.

XSL•FO
RenderX

Research Hotspots

Distribution of Research Topic

There are 7 clusters for 88 high-frequency keywords in the field of RS. The name was refined according to the main keywords contained in each cluster, as shown in Table 3. Cluster 1 is mainly related to robotic-assisted laparoscopy and some applicable diseases, mainly including gynecological and bladder diseases; Cluster 2 refers to the related technologies involved in surgical robots such as microsurgery, image-guided surgery, navigation, and telesurgery; Cluster 3 focuses on the robotic-assisted laparoscopic surgery and some applicable disease, such as colorectal and gastric disease; Cluster 4 is about the da Vinci robot and transoral robotic surgery; Cluster 5 is associated with prostate diseases and corresponding surgery; Cluster 6 is related to the training of surgical robot; and Cluster 7 refers to kidney diseases and corresponding surgery.

All of these clusters are visualized in Figure 11. Each of the 7 colors represents a cluster. For example, the purple on the left side represents Cluster 5, of which, the keywords are mainly related to prostatectomy; the green, located in the right represents Cluster 2, which covers keywords related to techniques of computer-assisted surgery. These 7 clusters of keywords may be better identified in a density visualization (Figure 12), which immediately reveals the general structure. It can be seen that laparoscopy and minimally invasive surgery are the most important keywords in the RS-related research, in addition to robotic.

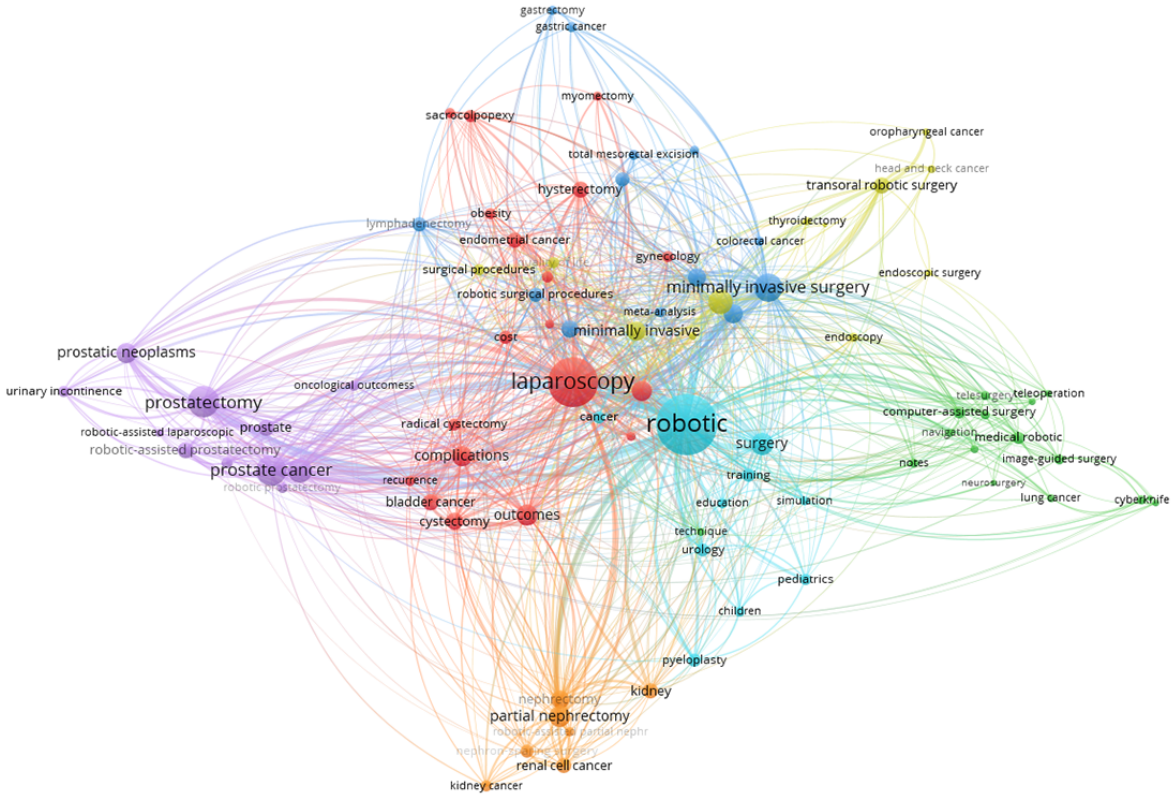
In general, the above 7 clusters of keywords on RS-related research can be combined into 3 themes according to the main content covered in each cluster: (1) various technologies, which include Cluster 2 and Cluster 6; (2) the robotic systems (ie, device and software) and related applications in surgery, which mostly contain Cluster 1, Cluster 3, and Cluster 4; and (3) prostate and kidney diseases and their corresponding operations, which include Cluster 5 and Cluster 7.

Table 3. 7 clusters of robotics in surgery-related research.

Cluster	Number of keywords	Cluster name	Keywords
1	19	Robotic-assisted Laparoscopy and some applicable diseases	Laparoscopy; robotic-assisted; complications; outcomes; endometrial cancer; hysterectomy; cystectomy; bladder cancer; cervical cancer; sacrocolpopexy; cost; radical cystectomy; pelvic organ prolapse; robotic-assisted laparoscopy; obesity; gynecology; myomectomy; robotic assisted; recurrence
2	16	Related technologies involved in surgical robots	Medical robotic; computer-assisted surgery; image-guided surgery; cyberknife; lung cancer; radiosurgery; navigation; telesurgery; microsurgery; augmented reality; notes; neurosurgery; haptics; teleoperation; technique; stereotactic radiosurgery
3	13	Robotic-assisted laparoscopic surgery and some applicable diseases	Minimally invasive surgery; robotic-assisted surgery; laparoscopic surgery; learning curve; rectal cancer; robotic surgical procedures; lymphadenectomy; gastric cancer; colorectal surgery; meta-analysis; colorectal cancer; gastrectomy; total mesorectal excision
4	12	da Vinci robot and transoral robotic surgery	da Vinci robot; transoral robotic surgery; minimally invasive; quality of life; endoscopy; head and neck cancer; endoscopic surgery; surgical procedures; surgical technique; oropharyngeal cancer; thyroidectomy; robotic thyroidectomy
5	11	Prostate diseases and corresponding surgery	Prostate cancer; prostatectomy; radical prostatectomy; robotic-assisted prostatectomy; prostatic neoplasms; prostate; robotic prostatectomy; robotic-assisted laparoscopic prostatectomy; urinary incontinence; oncological outcomes; continence
6	10	Training of surgical robot	Robotic; surgery; training; pyeloplasty; cancer; urology; pediatrics; simulation; children; education
7	7	Kidney diseases and corresponding surgery	Partial nephrectomy; nephrectomy; renal cell cancer; kidney; nephron-sparing surgery; robotic-assisted partial nephrectomy; kidney cancer

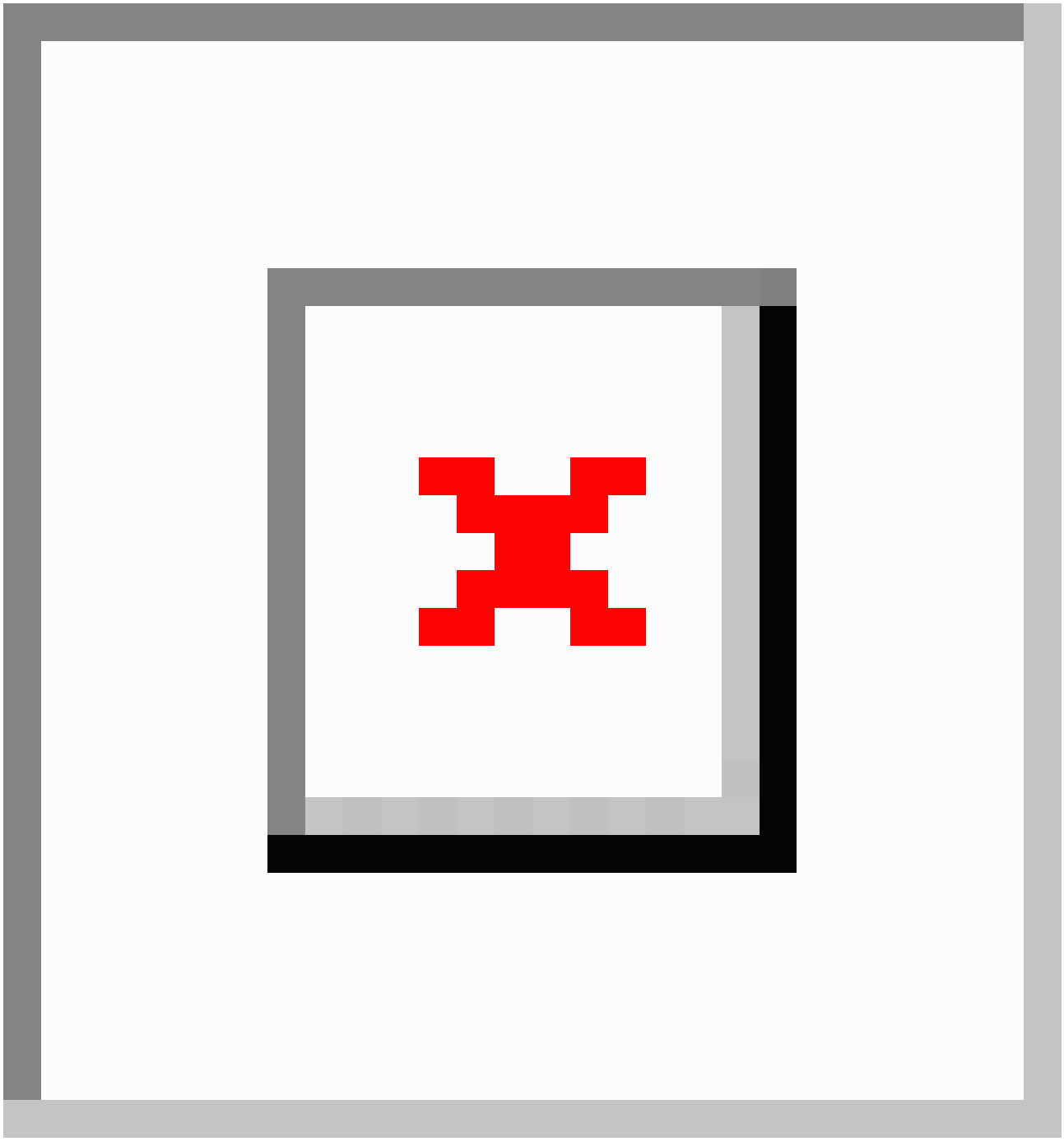
XSL•FO
RenderX

Figure 11. Visualization of the original 88×88 co-occurrence matrix.



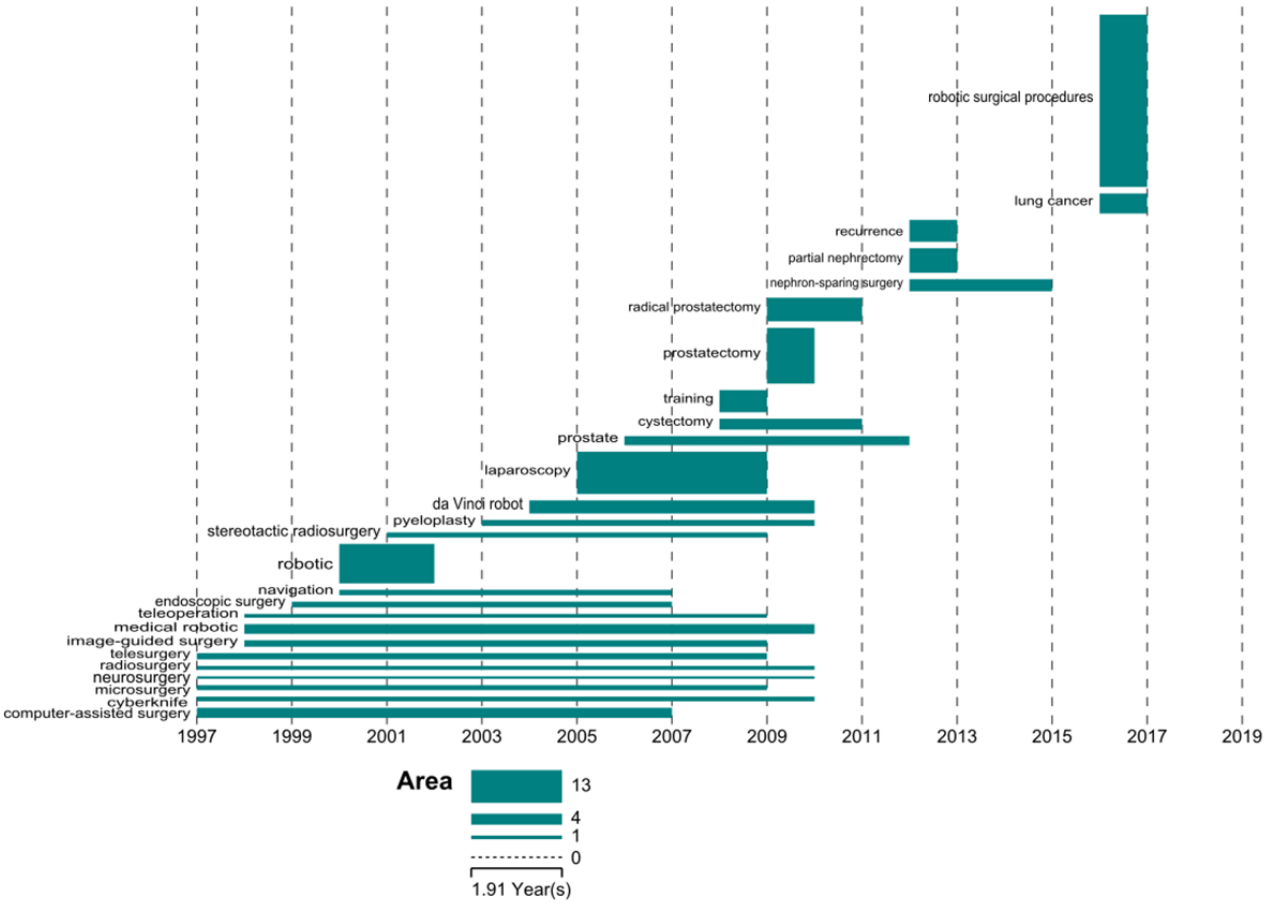
XSL•FO
RenderX

Figure 12. Co-occurrence density map.



XSL•FO
RenderX

Figure 13. Temporal bar graph for burst keywords.



Temporal Bar Graph for High Frequency and High Burst Keywords

There are 26 keywords whose burst strength is more than 4 and the frequency is not less than 40 (for details, see Multimedia Appendix 3). All 26 keywords represent the frontiers of research on RS to some extent. In addition, the frequencies of these 26 keywords are 5849, showing that 0.23% (26/11,303) of keywords account for 16.02% (5849/36,505) of the total 36,505 frequencies.

The temporal bar graph of 26 keywords visually demonstrates the evolution of the topic on RS-related research over time. As shown in Figure 13, it can be seen that early research mainly focused on computer-assisted surgery, microsurgery, image-guided surgery, and medical robotic during the period from 1997 to 2007, indicating that people began to pay attention to the application of computer technology in surgery. From 2000 to 2009, the representative burst keywords were robotic, da Vinci robot, laparoscopy, and prostate, suggesting that the various robotic systems began to appear and assisted the diagnosis and surgery of some diseases. The major burst keywords from 2009 to 2015 were prostatectomy, radical prostatectomy, and partial nephrectomy, showing that the robot was mainly used in the excision of various tissues and organs. In the past 2 years, the burst keywords were robotic surgical procedures and lung cancer. It is noteworthy that the term "robotic surgical procedures" was included in Medical Subject Headings terms in 2015, suggesting that the terms of RS have been standardized, and surgical robots have been applied to a wider field of treatment of tumor and cancer.

Discussion

Principal Findings

This study analyzed distribution of the literature related to RS, presented that the United States plays a leading role in the publication output, and presented that the Yonsei University in South Korea published the highest number of RS-related publications. Moreover, the disciplines covered by RS have expanded rapidly over the past years, the total number of which is more than 90. There exist some core disciplines in the field of RS, all of which have extensive cooperation with other disciplines. Obviously, these 4 disciplines make the majority of contributions to RS-related research. Furthermore, there are 7 clusters for the topic hotspots related to RS with imbalanced status, the evolution of which can be divided into 3 areas.

In this study, the distribution of RS-related research is imbalanced, although the RS-related research has caused global concerns. From the perspective of the country, the G7 (the United States, the United Kingdom, Germany, Canada, Italy, France, and Japan), which are headed by the United States, occupied a dominant position in this field. The research and development of surgical robots requires a large amount of human and financial resources. Apparently, the economic foundation plays an important role in the level of research and development

XSL•FO
RenderX

[38]. Therefore, countries with high economic levels are also relatively advanced in this field. From the perspective of the institution, 8 of the top 10 institutions belong to the United States, which further illustrates that the United States holds strongest research forces related to RS. Yonsei University in South Korea ranks the first with the most publication output because there are more than 10 hospitals affiliated to Yonsei Medical College [39], Severance Hospital, as one of which is the core institution for RS-related research in Asia.

This study showed that interdisciplinary collaboration of RS is widespread and has become more and more intensive in recent years. First, the number of clusters generated by visualization is gradually stable at around 13. The close connections among these disciplines aggregated into communities indicate how they support and supplement each other. However, the average number of disciplines in each paper has been reduced from 1.61 to 1.31, suggesting that the research content of each paper is more concentrated and more stable, although the disciplines of RS-related research are increasingly extensive [40]. This study on RS is mainly about how to cure a specific disease, such as various cancers, rather than a broad study of robotic surgery. Furthermore, several disciplines, such as Surgery, Oncology, Engineering, and Urology and Nephrology, are prominent in the cluster, and their collaborations with other disciplines are relatively close, showing that the main research related to RS is developed around these directions [41]. For example, Surgery has been in a dominant position for the last 21 years because it is the research topic in this study, but the relative position of Engineering disciplines that provide fundamentals and devices for RS is gradually declining, which may be due to the introduction of new disciplines with the deepening of RS-related research.

Obviously, cooperation between core disciplines needs to be further strengthened, as well as between other disciplines and core disciplines. The main disciplines in the field of RS are relatively stable, whereas other supporting disciplines are continuously changing at each period of time [42]. Surgery; Engineering; Radiology, Nuclear Medicine, and Medical Imaging; Neurosciences and Neurology have appeared in Figures 7 to 9, suggesting that these 4 disciplines are the core disciplines in the field of RS [43]. RS-related research is a relatively emerging interdisciplinary field, with a great potential impact on many areas of health care [44]. The exchange of ideas across disciplines promotes the progress of science. Medical robotics is fundamentally a team activity, involving academic researchers, clinicians, and industry. Each of these groups has unique expertise, and success comes from effective, highly interactive partnerships drawing upon this expertise [45]. Researchers with different disciplinary backgrounds have different professional knowledge, among which the cross-cooperation can promote the progress of a certain research subject. For example, in recent years, the treatment of tumor is a research hotspot in the field of RS, which needs not only doctors in Oncology but also researchers in other disciplines, such as Surgery and Engineering, who may come from different countries and institutions. Researchers with a background of different disciplines will provide different knowledge and skills to promote research on the subject of tumor treatment in the RS-related field. Therefore, more cooperation is needed in the field of RS, and the strengthening of cooperation can lead to the integration of knowledge, which means that RS is a more comprehensive research subject that includes technologies, devices, and the treatment of diseases.

Our study showed that the research focus on RS was relatively scattered and that each cluster has its own research emphasis but, in general, they can be merged into 3 main areas. Research hotspots clustering intuitively shows the relationship between 7 clusters of keywords, each of which represents a major research topic related to RS. However, some of these clusters have something in common. The clusters can be further divided into the following 3 parts: the first is the various technologies that make RS realized, such as computer-assisted surgery and image-guiding [46,47]; the second is various kinds of robotic systems and related applications in surgery, such as da Vinci robot, robotic-assisted laparoscopy, and robotic-assisted laparoscopic surgery [48-50]; the third is the application of RS in a variety of diseases and corresponding surgery, mainly tissue and organ excision, such as prostate and bladder [51].

Moreover, the research focus of several periods of time on RS reflected in the temporal bar graph is consistent with the results of the research hotspots clustering to some extent. It can be seen that the earliest research on RS is mainly about various technologies. RS is ultimately an application-driven research field. When technologies were relatively mature and measurable, the robotic systems and their related application began to emerge and became the research frontier, and surgeons began to accept and apply them in surgery. Then, the research focus began to turn to various diseases and their corresponding surgery. Apparently, it is undeniable that the robotic systems and surgery for various diseases are complementary to each other throughout the development of RS. Correspondingly, there were 3 aspects of researches in each period of time, but the focus was different. Generally, research on RS should be further promoted in these 3 areas to better strengthen the integration between surgery and robotics.

In addition, to clearly present the frontiers of RS-related research, we reduced the burst strength of keywords to 2.5 targeted at the dataset in the recent 3 years (2015 to 2017), which mainly contains lung cancer and lobectomy [52], rectal cancer and colorectal surgery [53,54], and esophageal cancer and esophagectomy [55]. It is obvious that RS-related research has begun to play an important role in the diagnosis and treatment of various specific cancers. Clearly, as the application of robots to surgery, targeted at specific diseases, has been rapidly expanded, better regulations and standards should be developed and implemented; and methods to assess safety should be adopted in future, which is conducive to exploit the full potential of robotics in medicine, especially in surgery, for the improved welfare of society everywhere [56]. In addition, other ethical concerns will emerge as robotic technologies become more intelligent with advances in cognitive software.

Limitations

There still exist several limitations in this study. First, it is difficult to visually observe several disciplines that are most closely related because the algorithm performed in VOSviewer

XSL•FO
RenderX

stipulates that the distance of disciplines in the visualization graph is closer if more collaborations between them exist. Second, the subject categorization indexed by WoSCC may be inaccurate, which may have a certain impact on the result of research and lead to some bias for the visualization of interdisciplinary collaboration. Finally, results of the topic hotspots analysis are affected by the keyword merging. This study merely merged the keyword with a frequency more than 5, that is, there are still some synonym keywords that should be merged. All of these may have some influence on the results of the topic clustering.

Conclusions

In this study, various bibliometric measures on RS-related research were performed using the corresponding visual tool. In all, on the base of the above study, some valuable results from RS-related research were obtained, including information on interdisciplinary collaboration and research hotspots, which offer a comprehensive understanding of RS-related research. Moreover, with the development of artificial intelligence and the further widespread application of robots to surgery, it should be reasonable to believe that the literature related to RS research will continue to grow in future. In addition, as the research develops, some new application of robots in surgery will form, which will give rise to new problems, such as the degree of interdisciplinary collaboration and its effect on research productivity, all of which are the future research focus. Furthermore, the research of application will be the strength to further improve the accuracy and safety and reduce cost, although the surgery remains the core discipline of RS-related research.

Acknowledgments

This study was supported by the Fundamental Research Funds for the Central Universities, Huazhong University of Science and Technology (No. 2016YXMS181). The authors would like to thank all anonymous reviewers for their valuable comments and input to this research.

Authors' Contributions

LS, the cofirst author and corresponding author, designed the study and conducted the research into literature distribution related to RS. SW, cofirst author, contributed to the data collection and research into research hotspots. WD, cofirst author, contributed to the research into the literature distribution and interdisciplinary nature. ZZ, the fourth author, reviewed the draft. All authors contributed to the manuscript’s preparation and approved the final accepted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Top 10 countries for robotics in surgery-related research.

[PDF File (Adobe PDF File), 30KB - jmir_v21i3e12625_app1.pdf]

Multimedia Appendix 2

Clusters for interdisciplinary collaboration related to robotics in surgery.

[PDF File (Adobe PDF File), 76KB - jmir_v21i3e12625_app2.pdf]

Multimedia Appendix 3

List of keywords with burst strength more than 4 and frequency not less than 40.

[PDF File (Adobe PDF File), 33KB - jmir_v21i3e12625_app3.pdf]

References

1. Burgner-Kahrs J, Rucker DC, Choset H. Continuum robots for medical applications: a survey. IEEE Trans Robot 2015 Dec;31(6):1261-1280. [doi: 10.1109/TRO.2015.2489500]
2. Liu H, Guo D, Sun F. Object recognition using tactile measurements: kernel sparse coding methods. IEEE T Instrum Meas 2016 Mar;65(3):656-665. [doi: 10.1109/TIM.2016.2514779]
3. Suprem A, Mahalik N, Kim K. A review on application of technology systems, standards and interfaces for agriculture and food sector. Comput Stand Interfaces 2013 Jun;35(4):355-364. [doi: 10.1016/j.csi.2012.09.002]
4. Hirzinger G. Robots in space - a survey. Adv Robotics 2012 Apr;9(6):625-651. [doi: 10.1163/156855395X00337]
5. Leite I, Martinho C, Paiva A. Social robots for long-term interaction: a survey. Int J of Soc Robotics 2013 Jan 25;5(2):291-308. [doi: 10.1007/s12369-013-0178-y]

XSL•FO
RenderX

6. Mutter D, Callari C, Diana M, Dallemagne B, Leroy J, Marescaux J. Single port laparoscopic cholecystectomy: which technique, which surgeon, for which patient? A study of the implementation in a teaching hospital. J Hepatobiliary Pancreat Sci 2011 May;18(3):453-457. [doi: 10.1007/s00534-010-0348-3] [Medline: 21153842]
7. Bae JK, Vavilin A, You JS, Kim H, Ryu SY, Jang JH, et al. Smartphone-based endoscope system for advanced point-of-care diagnostics: feasibility study. JMIR Mhealth Uhealth 2017 Jul 27;5(7):e99 [FREE Full text] [doi: 10.2196/mhealth.7232] [Medline: 28751302]
8. Nicolau S, Soler L, Mutter D, Marescaux J. Augmented reality in laparoscopic surgical oncology. Surg Oncol 2011 Sep;20(3):189-201. [doi: 10.1016/j.suronc.2011.07.002] [Medline: 21802281]
9. Hu JC, Gu X, Lipsitz SR, Barry MJ, D'Amico AV, Weinberg AC, et al. Comparative effectiveness of minimally invasive vs open radical prostatectomy. J Am Med Soc 2009 Oct 14;302(14):1557-1564 [FREE Full text] [doi: 10.1001/jama.2009.1451] [Medline: 19826025]
10. Jayne D, Pigazzi A, Marshall H, Croft J, Corrigan N, Copeland J, et al. Effect of robotic-assisted vs conventional laparoscopic surgery on risk of conversion to open laparotomy among patients undergoing resection for rectal cancer: the ROLARR randomized clinical trial. J Am Med Soc 2017 Dec 24;318(16):1569-1580 [FREE Full text] [doi: 10.1001/jama.2017.7219] [Medline: 29067426]
11. Yaxley JW, Coughlin GD, Chambers SK, Occhipinti S, Samaratunga H, Zajdlewicz L, et al. Robot-assisted laparoscopic prostatectomy versus open radical retropubic prostatectomy: early outcomes from a randomised controlled phase 3 study. Lancet 2016 Dec 10;388(10049):1057-1066 [FREE Full text] [doi: 10.1016/S0140-6736(16)30592-X] [Medline: 27474375]
12. Yang HX, Woo KM, Sima CS, Bains MS, Adusumilli PS, Huang J, et al. Long-term survival based on the surgical approach to lobectomy for clinical stage I nonsmall cell lung cancer: comparison of robotic, video-assisted thoracic surgery, and thoracotomy lobectomy. Ann Surg 2017 Feb;265(2):431-437 [FREE Full text] [doi: 10.1097/SLA.0000000000001708] [Medline: 28059973]
13. Web of Science. URL: http://login.webofknowledge.com/error/Error?Src=IP&Alias=WOK5&Error=IPError&Params=&PathInfo=%2F&RouterURL=http%3A%2F%2Fwww.webofknowledge.com%2F&Domain=.webofknowledge.com&locale=en_US [accessed 2018-10-24] [WebCite Cache ID 73PRPvyGH]
14. Bellangino M, Verrill C, Leslie T, Bell RW, Hamdy FC, Lamb AD. Systematic review of studies reporting positive surgical margins after bladder neck sparing radical prostatectomy. Curr Urol Rep 2017 Nov 7;18(12):99 [FREE Full text] [doi: 10.1007/s11934-017-0745-0] [Medline: 29116405]
15. Jeske DR, Linehan JA, Wilson TG, Kawachi MH, Wittig K, Lamparska K, et al. Two-stage classifiers that minimize PCA3 and the PSA proteolytic activity testing in the prediction of prostate cancer recurrence after radical prostatectomy. Can J Urol 2017 Dec;24(6):9089-9097. [Medline: 29260633]
16. Thulasi K, Arunachalam S. Mapping of cholera research in India using HistCite. Ann Libr Inform Stud 2010 Sep;57(3):310-326 [FREE Full text]
17. Garfield E, Pudovkin AI, Istomin VS. Why do we need algorithmic historiography? J Am Soc Inf Sci 2003 Mar;54(5):400-412. [doi: 10.1002/asi.10226]
18. Shen L, Xiong B, Li W, Lan F, Evans R, Zhang W. Visualizing collaboration characteristics and topic burst on international mobile health research: bibliometric analysis. JMIR Mhealth Uhealth 2018 Jun 5;6(5):e135 [FREE Full text] [doi: 10.2196/mhealth.9581] [Medline: 29871851]
19. Lucio-Arias D. A validation study of HistCite I: using the discoveries of fullerenes and nanotubes. 2007 Jun Presented at: 11th International Conference of the International Society for Scientometrics and Informetrics; June 25-27, 2007; Madrid, Spain p. 886-887.
20. Brown RR, Deletic A, Wong TH. Interdisciplinarity: how to catalyse collaboration. Nature 2015 Sep 17;525(7569):315-317 [FREE Full text] [doi: 10.1038/525315a] [Medline: 26381970]
21. Small H. Maps of science as interdisciplinary discourse: co-citation contexts and the role of analogy. Scientometrics 2010 Jun;83(3):835-849. [doi: 10.1007/s11192-009-0121-z]
22. Web of Science Core Collection Help. URL: http://images.webofknowledge.com/WOKRS530AR16/help/WOS/hp_research_areas_easca.html?tdsourcetag=s_pcqq_aiomsg [accessed 2018-10-28] [WebCite Cache ID 73V9T8SKE]
23. Perianes-Rodriguez A, Waltman L, van Eck N. Constructing bibliometric networks: a comparison between full and fractional counting. J Informetr 2016 Nov;10(4):1178-1195. [doi: 10.1016/j.joi.2016.10.006]
24. Hu JM, Zhang Y. Discovering the interdisciplinary nature of big data research through social network analysis and visualization. Scientometrics 2017 May 8;112(1):91-109. [doi: 10.1007/s11192-017-2383-1]
25. Cui L, Liu W, Yan L. Development of a text mining system based on the co-occurrence of bibliographic items in literature databases. New Tech of Libr and Inform Serv 2008 Aug;24(8):70-75. [doi: 10.11925/infotech.1003-3513.2008.08.12]
26. Lee WH. How to identify emerging research fields using scientometrics: an example in the field of information security. Scientometrics 2008 Sep;76(3):503-525. [doi: 10.1007/s11192-007-1898-2]
27. Borgatti SP, Everett MG, Johnson JC. Analyzing Social Networks First Edition. London: Sage Publications; 2013.

XSL•FO
RenderX

28. Rosvall M, Bergstrom CT. Mapping change in large networks. PLoS One 2010 Jan 27;5(1):e8694 [FREE Full text] [doi: 10.1371/journal.pone.0008694] [Medline: 20111700]
29. Leydesdorff L, Goldstone RL. Interdisciplinarity at the journal and specialty level: the changing knowledge bases of the journal. J Assn Inf Sci Tec 2013 Oct 9;65(1):164-177 [FREE Full text] [doi: 10.1002/asi.22953]
30. Cortext. URL: https://www.cortext.net/ [accessed 2018-10-24] [WebCite Cache ID 73Q3y5o43]
31. Shen L, Xiong B, Hu J. Research status, hotspots and trends for information behavior in China using bibliometric and co-word analysis. J Doc 2017 Jul 10;73(4):618-633. [doi: 10.1108/JD-10-2016-0125]
32. VOSviewer - Visualizing scientific landscapes. URL: http://www.vosviewer.com/ [accessed 2019-01-30] [WebCite Cache ID 75p8okq0L]
33. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010 Aug;84(2):523-538 [FREE Full text] [doi: 10.1007/s11192-009-0146-3] [Medline: 20585380]
34. Hu J, Zhang Y. Research patterns and trends of recommendation system in China using co-word analysis. Inform Process Manag 2015 Jul;51(4):329-339. [doi: 10.1016/j.ipm.2015.02.002]
35. Kleinberg J. Bursty and hierarchical structure in streams. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2002 Presented at: KDD'02; July 23-26, July; Edmonton, Alberta, Canada p. 91-101. [doi: 10.1145/775047.775061]
36. Sci2 Tool : A Tool for Science of Science Research and Practice. URL: https://sci2.cns.iu.edu/user/index.php [accessed 2018-10-24] [WebCite Cache ID 6vKq1XbM8]
37. Mane KK, Börner K. Mapping topics and topic bursts in PNAS. Proc Natl Acad Sci U S A 2004 Apr 6;101(Suppl 1):5287-5290 [FREE Full text] [doi: 10.1073/pnas.0307626100] [Medline: 14978278]
38. Lu HM, Li YJ, Chen M, Kim H, Serikawa S. Brain intelligence: go beyond artificial intelligence. Mobile Netw Appl 2018 Apr;23(2):368-375. [doi: 10.1007/s11036-017-0932-8]
39. Baidu Encyclopedia. Yonsei University URL: https://baike.baidu.com/item/%E5%BB%B6%E4%B8%96%E5%A4%A7%E5%AD%A6/3476477?fr=aladdin [accessed 2018-10-24] [WebCite Cache ID 73Q3mIuzU]
40. Hoffmann M, Marques HG, Arieta AH, Sumioka H, Lungarella M, Pfeifer R. Body schema in robotics: a review. IEEE T Auton Ment Dev 2010 Dec;2(4):304-324. [doi: 10.1109/TAMD.2010.2086454]
41. Marescaux J, Diana M. Next step in minimally invasive surgery: hybrid image-guided surgery. J Pediatr Surg 2015 Jan;50(1):30-36. [doi: 10.1016/j.jpedsurg.2014.10.022] [Medline: 25598089]
42. Liu M, Curet M. A review of training research and virtual reality simulators for the da Vinci surgical system. Teach Learn Med 2015;27(1):12-26. [doi: 10.1080/10401334.2014.979181] [Medline: 25584468]
43. Micali S, Pini G, Teber D, Sighinolfi MC, Stefani SD, Bianchi G, et al. New trends in minimally invasive urological surgery. What is beyond the robot? World J Urol 2013 Jun;31(3):505-513. [doi: 10.1007/s00345-010-0588-5] [Medline: 20812015]
44. Haidegger T, Sándor J, Benyó Z. Surgery in space: the future of robotic telesurgery. Surg Endosc 2011 Mar;25(3):681-690. [doi: 10.1007/s00464-010-1243-3] [Medline: 20652320]
45. Siciliano B, Khatib O. Springer Handbook Of Robotics Second Edition. Berlin: Springer International Publishing; 2016.
46. Taylor RH, Stoianovici D. Medical robotics in computer-integrated surgery. IEEE T Robotic Autom 2003 Oct;19(5):765-781. [doi: 10.1109/TRA.2003.817058]
47. DiGioia AM, Jaramaz B, Colgan BD. Computer assisted orthopaedic surgery. Image guided and robotic assistive technologies. Clin Orthop Relat Res 1998 Sep;354(354):8-16. [doi: 10.1097/00003086-199809000-00003] [Medline: 9755759]
48. Bodner J, Wykypiel H, Wetscher G, Schmid T. First experiences with the da Vinci operating robot in thoracic surgery. Eur J Cardiothorac Surg 2004 May;25(5):844-851. [doi: 10.1016/j.ejcts.2004.02.001] [Medline: 15082292]
49. Narducci F, Lambaudie E, Houvenaeghel G, Collinet P, Leblanc E. Early experience of robotic-assisted laparoscopy for extraperitoneal para-aortic lymphadenectomy up to the left renal vein. Gynecol Oncol 2009 Oct;115(1):172-174. [doi: 10.1016/j.ygyno.2009.04.032] [Medline: 19450870]
50. Nezhat C, Lavie O, Lemyre M, Unal E, Nezhat CH, Nezhat F. Robot-assisted laparoscopic surgery in gynecology: scientific dream or reality? Fertil Steril 2009 Jun;91(6):2620-2622 [FREE Full text] [doi: 10.1016/j.fertnstert.2008.03.070] [Medline: 18656185]
51. Ficarra V, Novara G, Artibani W, Cestari A, Galfano A, Graefen M, et al. Retropubic, laparoscopic, and robot-assisted radical prostatectomy: a systematic review and cumulative analysis of comparative studies. Eur Urol 2009 May;55(5):1037-1063. [doi: 10.1016/j.eururo.2009.01.036] [Medline: 19185977]
52. Veronesi G. Robotic lobectomy and segmentectomy for lung cancer: results and operating technique. J Thorac Dis 2015 Apr;7(Suppl 2):122-130 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2015.04.34] [Medline: 25984357]
53. Lee SH, Lim S, Kim JH, Lee KY. Robotic versus conventional laparoscopic surgery for rectal cancer: systematic review and meta-analysis. Ann Surg Treat Res 2015 Oct;89(4):190-201 [FREE Full text] [doi: 10.4174/astr.2015.89.4.190] [Medline: 26448918]
54. Kim D, Baek SJ, Kang DW, Roh YE, Lee JW, Kwak HD, et al. Robotic resection is a good prognostic factor in rectal cancer compared with laparoscopic resection: long-term survival analysis using propensity score matching. Dis Colon Rectum 2017 Mar;60(3):266-273. [doi: 10.1097/DCR.0000000000000770] [Medline: 28177988]

XSL•FO
RenderX

55. Suda K, Nakauchi M, Inaba K, Ishida Y, Uyama I. Robotic surgery for upper gastrointestinal cancer: current status and future perspectives. Dig Endosc 2016 Nov;28(7):701-713. [doi: 10.1111/den.12697] [Medline: 27403808]
56. Fei BW, Ng WS, Chauhan S, Kwoh CK. The safety issues of medical robotics. Reliab Eng Syst Saf 2001 Aug;73(2):183-192 [FREE Full text] [doi: 10.1016/S0951-8320(01)00037-0]

Abbreviations

AGCS: average global citation score
BICOMS: Bibliographic Item Co-Occurrence Mining System
RS: robotics in surgery
Sci2: Science of Science
TGCS: total global citation score
TLCS: total local citation score
WoSCC: Web of Science Core Collection

Edited by C Eichenberg; submitted 31.10.18; peer-reviewed by JA Sánchez-Margallo, L Moscote; comments to author 01.12.18; revised version received 31.12.18; accepted 02.01.19; published 26.03.19.

Please cite as:
Shen L, Wang S, Dai W, Zhang Z
Detecting the Interdisciplinary Nature and Topic Hotspots of Robotics in Surgery: Social Network Analysis and Bibliometric Study
J Med Internet Res 2019;21(3):e12625
URL: http://www.jmir.org/2019/3/e12625/
doi:10.2196/12625
PMID:30912752

©Lining Shen, Shimin Wang, Wei Dai, Zhiguo Zhang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Medical Doctors' Offline Computer-Assisted Digital Education: Systematic Review by the Digital Health Education Collaboration

Hayfaa Abdelmageed Wahabi[1], MSc, MD, PhD, FRCOG, FRCS (Ed); Samia Ahmed Esmaeil[1], MD, MBBS; Khawater Hassan Bahkali[1,2], MD, MPH; Maher Abdelraheim Titi[1,3], RN, MSN; Yasser Sami Amer[1,4], MD, MPed; Amel Ahmed Fayed[5,6], MD, PhD; Amr Jamal[1,2], MD, PhD; Nasriah Zakaria[7], PhD; Amna Rehana Siddiqui[2,8], MD, PhD; Monika Semwal[9], PhD; Lorainne Tudor Car[10], MD, PhD; Paul Posadzki[9], PhD; Josip Car[9], MSc, MD, PhD, FRCP, FFPH

[1]Research Chair of Evidence-Based Healthcare and Knowledge Translation, Deanship of Research, King Saud University, Riyadh, Saudi Arabia

[2]Department of Family and Community Medicine, College of Medicine, King Saud University, Riyadh, Saudi Arabia

[3]Patient Safety Unit, Quality Management Department, King Khalid University Hospital, King Saud Medical City, Riyadh, Saudi Arabia

[4]Clinical Practice Guidelines Unit, Quality Management Department, King Khalid University Hospital, King Saud University Medical City, Riyadh, Saudi Arabia

[5]College of Medicine, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

[6]High Institute of Public Health, Alexandria University, Alexandria, Egypt

[7]Medical Informatics and e-Learning Unit, Medical Education Department, College of Medicine, King Saud University, Riyadh, Saudi Arabia

[8]Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan

[9]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[10]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Hayfaa Abdelmageed Wahabi, MSc, MD, PhD, FRCOG, FRCS (Ed)
Research Chair of Evidence-Based Healthcare and Knowledge Translation
Deanship of Research
King Saud University
7373 Al Fawaz Streeat, Al Wadi District
Riyadh, 13313-4171
Saudi Arabia
Phone: 966 0553294955
Email: umlena@yahoo.com

Abstract

Background: The widening gap between innovations in the medical field and the dissemination of such information to doctors may affect the quality of care. Offline computer-based digital education (OCDE) may be a potential solution to overcoming the geographical, financial, and temporal obstacles faced by doctors.

Objective: The objectives of this systematic review were to evaluate the effectiveness of OCDE compared with face-to-face learning, no intervention, or other types of digital learning for improving medical doctors' knowledge, cognitive skills, and patient-related outcomes. Secondary objectives were to assess the cost-effectiveness (CE) of OCDE and any adverse effects.

Methods: We searched major bibliographic databases from 1990 to August 2017 to identify relevant articles and followed the Cochrane methodology for systematic reviews of intervention.

Results: Overall, 27 randomized controlled trials (RCTs), 1 cluster RCT (cRCT), and 1 quasi-RCT were included in this review. The total number of participants was 1690 in addition to the cRCT, which included 24 practices. Due to the heterogeneity of the participants, interventions, and outcomes, meta-analysis was not feasible, and the results were presented as narrative summary. Compared with face-to-face learning, the effect of OCDE on knowledge gain is uncertain (ratio of the means [RM] range 0.95-1.17; 8 studies, 495 participants; very low grade of evidence). From the same comparison, the effect of OCDE on cognitive skill gain is uncertain (RM range 0.1-0.9; 8 studies, 375 participants; very low grade of evidence). OCDE may have little or no effect on patients' outcome compared with face-to-face education (2 studies, 62 participants; low grade of evidence). Compared with no intervention, OCDE may improve knowledge gain (RM range 1.36-0.98; 4 studies, 401 participants; low grade of evidence). From the same comparison, the effect of OCDE on cognitive skill gain is uncertain (RM range 1.1-1.15; 4 trials, 495 participants; very low grade of evidence). One cRCT, involving 24 practices, investigated patients' outcome in this comparison and showed

XSL•FO
RenderX

no difference between the 2 groups with low-grade evidence. Compared with text-based learning, the effect of OCDE on cognitive skills gain is uncertain (RM range 0.91-1.46; 3 trials with 4 interventions; 68 participants; very low-grade evidence). No study in this comparison investigated knowledge gain or patients' outcomes. One study assessed the CE and showed that OCDE was cost-effective when compared with face-to-face learning if the cost is less than or equal to Can $200. No trial evaluated the adverse effect of OCDE.

Conclusions: The effect of OCDE compared with other methods of education on medical doctors' knowledge and cognitive skill gain is uncertain. OCDE may improve doctors' knowledge compared with no intervention but its effect on doctors' cognitive skills is uncertain. OCDE may have little or no effect in improving patients' outcome.

(J Med Internet Res 2019;21(3):e12998) doi:10.2196/12998

KEYWORDS

systematic review; medical education, digital education

Introduction

Background

Faced with the rapid innovations in medicine, structured postgraduate residency programs, in addition to continuing professional development (CPD) and continuing medical education, (CME), have been developed to advance and update the skills and knowledge of medical doctors and other health professions [1]. Evidence suggested that these programs are effective in improving the diagnostic and therapeutic competencies of health care professionals and patient-related outcomes [2-6]. Although face-to-face learning is the dominant method of teaching in these programs, it is being increasingly supplemented or replaced with digital learning using both Web-based and offline options [7]. Digital education involves the delivery of educational material through Information and Communication Technology using a wide variety of pedagogical designs and formats [8,9]. Digital learning is a plausible low-cost platform that provides convenient access to educational materials with flexibility in terms of pace, place, and time [10]. It has many advantages over traditional learning, such as outreach, flexibility, and adaptability, as it has the potential to reach a large number of learners regardless of the physical distance; in addition, it caters for the pace and time of the individual participant while reducing the overhead costs of the learning process [11-13]. These features of digital learning could prove it to be a plausible solution to the constraints faced by medical education in low- and middle-income countries (LMICs) where digital learning has been used effectively to improve existing health services [12] and the diagnostic and therapeutic competencies of the health care providers [2].

Despite the potential of digital learning to be the leading learning method in sciences, including medical education, some disadvantages have been reported for this method of learning. Lack of interaction with other learners and the instructor, because of flexibility in the timing of learning, may lead to social isolation and discouragement of team work [11,14].

Although digital learning caters for students' learning pace, it tends to deindividualize the instructor and fails to respond to the learners' individual needs [15]. Although digital learning is associated with reduced cost for the learner, the potentially large cost of designing some courses, such as those that involve virtual reality and simulation, cannot be ignored [15].

Unlike other learning methods, digitally-based courses must be designed carefully to meet the instructional objectives as the absence of an instructor for explanation makes such courses more susceptible to design flaws [15,16]. For effective digital learning in medical education, many essential design characteristics have been suggested, such as effective communication among the learners and validation and assessment of knowledge gained in addition to use of real-world scenarios [17].

Currently, there are many different technological platforms for digital learning, including Web-based learning activities, computer-based instruction, and mobile learning. The availability of many platforms paved the way for many learning opportunities in medical education, such as massive open Web-based courses, serious gaming and gamification, and virtual patient-based learning. The focus of this review is offline computer-based digital education (OCDE).

This type of digital education does not require internet or local area network connection, and the learning material is typically kept in either magnetic storage such as floppy discs or optical storage such as CD, digital versatile disc, flash memory, multimedia cards, and external hard discs, which facilitates the delivery of various educational material such as text, images, audio, and video material [18-20]. OCDE has many advantages, especially in settings where internet access is absent or limited. It has many of the advantages over other digital learning platforms as it provides a solution for those learners faced with geographical, financial, and temporal barriers to face-to-face education without compromising on the teaching and learning process or the intended outcomes [18,19,21].

Objectives

The objective of this systematic review was to evaluate the effectiveness of OCDE compared with face-to-face learning, no intervention, text-based learning, or other type of digital learning for improving medical doctors' knowledge, cognitive skills, and patient-related outcomes. The secondary objectives were to assess the cost and cost-effectiveness (CE) of OCDE and adverse effects of the interventions.

XSL•FO
RenderX

Methods

Eligibility Criteria

Types of Studies

We included randomized controlled trials (RCTs), cluster RCTs (cRCTs), and quasi-randomized trials that compared OCDE (personal computer or laptop) methods for medical doctors and dentists with face-to-face learning, written information, no intervention, or other OCDE. We excluded crossover trials because of the high likelihood of a carry-over effect [22]. We included trials reported in conference proceedings and abstracts when information could be obtained by contacting the authors. For a detailed description of the methodology, please refer to the study by Car et al [23].

Types of Participants

We included studies in which participants (learners) were medical doctors and dentists who were enrolled in postgraduate medical education programs. Studies with mixed participant groups, such as doctors and nurses, in which results for medical doctors could not be obtained separately, were excluded.

Types of Interventions

We included studies in which OCDE interventions were used to deliver educational content. Residency training programs and CME- and CPD-based programs that involved the use of OCDE interventions were included [24,25].

OCDE intervention refers to the use of personal computers or laptops that have assisted in the delivery of standalone multimedia materials without the need for internet or local area network connections [24,25].

We only considered studies that made the following intervention comparisons:

- Offline-based intervention versus traditional face-to-face learning.
- Offline-based intervention versus no intervention.
- Offline-based digital learning versus written text-based learning.
- Offline-based digital learning versus another method of digital learning.

Primary Outcomes

We included studies that reported at least one of the following primary or secondary outcomes

- Learner's knowledge: defined as the learners' factual or conceptual understanding.
- Learner's cognitive skills: defined as skills used in the process of acquiring knowledge, for example, skills learned for reading an x-ray film or learning the steps of performing a procedure.
- Patients' outcomes: defined as the direct observation of the application of knowledge and skills on the patients where the outcome of the trial is patients' physical, mental, and psychological condition, such as the clinical effect of optimizing medication regimens. In addition, we considered studies that aimed at improving health services, such as improving existing screening programs, as patients' outcomes.

Secondary Outcomes

- CE of digital learning
- Any adverse outcome

Types of Outcome Measures

These outcomes were assessed using any validated or non-validated instrument to measure the difference in pre and posttest scores. These assessments were either subjective (eg, self-reported) or objective (eg, questionnaire). When several posttest results were available, data were recorded as to when those tests were conducted and the difference between the pretest and the first posttest was used for the analysis.

Search Methods

We searched major bibliographic databases from 1990, when the virtual learning environments began and schools started delivering Web-based courses, till August 2017 to identify all relevant articles. We searched in English but included papers published in any language.

Electronic Searches

We searched the following databases: Medical Literature Analysis and Retrieval System Online (MEDLINE, via Ovid), EMBASE (via Ovid), Web of Science, Educational Resource Information Centre (via Ovid), Cochrane Central Register of Controlled Trials, Cochrane Library, PsycINFO (Ovid), Cumulative Index to Nursing and Allied Health Literature (via EBSCO), and ProQuest Dissertation and Theses Database. The MEDLINE search strategy was adapted to search other databases (Multimedia Appendix 1).

Searching Other Resources

We searched the reference lists of all included studies and relevant systematic reviews. We also searched the International Clinical Trials Registry Platform Search Portal and Current Controlled Trials metaRegister of Controlled Trials to identify the unpublished trials and contacted relevant investigators for further information.

Data Collection and Analysis

Study Selection

Two reviewers independently screened the titles and abstracts and identified studies potentially meeting the inclusion criteria. The full-text versions were retrieved and read in full. Finally, 2 review authors independently assessed the articles against the eligibility criteria. Any disagreements were resolved through discussion between the 2 authors. If no agreement was reached, a third author acted as an arbiter. Two reviewers verified the final list of included studies.

Data Extraction and Management

Two reviewers independently extracted and managed the data for each of the included studies and used a structured data recording form. In addition to the usual information on the study design and participants' demographics, we extracted data on relevant fields such as the country where the trial was conducted,

XSL•FO
RenderX

funding source, and duration of intervention. Disagreements between the review authors were resolved by discussion. A third review author acted as an arbiter in case disagreements were not resolved.

Dealing with Missing Data

Whenever possible, we attempted to obtain missing data from the original authors.

Assessment of Risk of Bias in Included Studies

Two reviewers independently assessed the risk of bias of each of the included studies using the Cochrane Collaboration's risk of bias tool [22]. Studies were assessed for the risk of bias in the following domains: random sequence generation, allocation concealment, blinding of participants or personnel, blinding of outcome assessors, completeness of outcome data (attrition bias), selective outcome reporting (reporting bias), and other sources of bias including baseline imbalance and contamination. For cRCTs, we assessed and reported the risk of bias associated with an additional domain: selective recruitment of cluster participants, baseline imbalance, attrition of clusters, and not accounting for cluster effect in analysis [26].

We judged the risk of bias for each study to be of 1 of 3 levels: high, low, or unclear risk of bias. We scored each study for risk of bias as follows: *low* if all key domains were scored as *low risk* or if 1 domain is scored as *unclear*. We scored the trial as *unclear* if 2 key domains were scored as *unclearrisk* and *high* if more than 2 key domains were scored *unclear risk* or 1 domain scored *highrisk*, adapted from the study by Davey et al [27].

Reporting bias was assessed qualitatively on the basis of the characteristics of included studies. Due to the heterogeneity of the trials (in terms of populations, interventions, comparator groups, and outcomes), data pooling was not feasible.

We used Evers checklist [28] to evaluate risk of bias in articles that examined the CE of OCDE.

Data Synthesis

We reported post intervention values for the outcomes of intervention and control groups and the effect size as reported by the authors (*P* value). In addition, we calculated the ratio of the means (RM) [4] whenever feasible. As the heterogeneity of populations, outcomes, and comparisons precluded meta-analysis, we provided a narrative summary of the results.

Summary of Findings Tables

For main comparisons, 2 authors used the *Grading of Recommendations Assessment, Development and Evaluation* (GRADE) criteria independently of one another to assess the quality of evidence [29]. We considered the following limitations: risk of bias, inconsistency of results, indirectness of the evidence, imprecision, or publication bias, and subsequently downgraded the quality of evidence where appropriate [30].

Results

Overview

The study selection process is shown in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Figure 1).

The initial search yielded 21,849 records. After screening the titles and abstracts, we obtained the full-text reports for 197 records and assessed them for inclusion in the review. Of these, we excluded 168 studies that did not meet the eligibility criteria (Multimedia Appendix 2). The remaining 29 articles were included in this review.

We contacted authors of abstracts for further information and data on their trials [31-37], we received a response from 2 authors, Rae et al [37], where the data were subsequently included in the review, and Ukabiala et al [36], which was subsequently excluded. We tried to contact the authors of the cRCT [38] for data and information to assess the risk of bias and to calculate the intracluster correlation coefficients (ICC); however, the contact information was void.

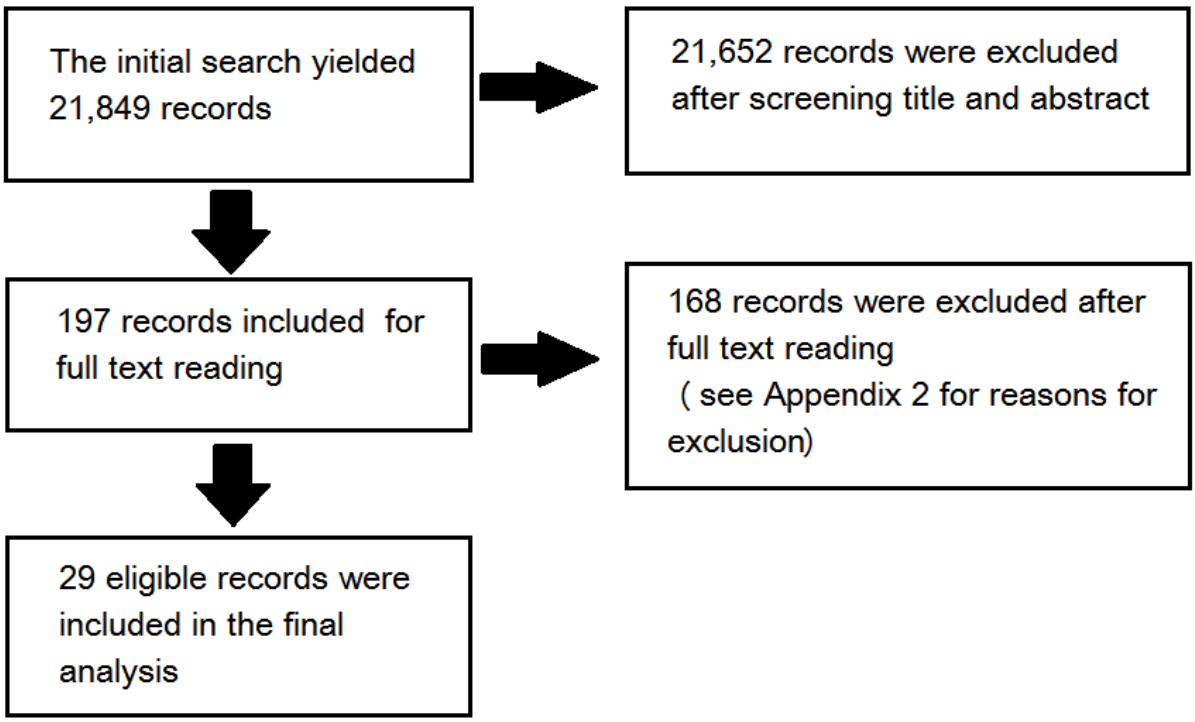
Included Studies

A total of 24 of the 29 included trials were parallel RCTs, each included 2 arms and 1 trial had 4 arms, 3 of which were included in this review, resulting in 2 interventions [37]. In addition, 1 trial had 3 arms [39], 1 was a factorial RCT [40], 1 was a cRCT [38], and 1 was a quasi-RCT [41]. All trials were published in peer-reviewed journals except for 1 trial where unpublished data were obtained from the authors [37]. Overall, 24 studies (83%, 24/29) were conducted in high-income countries [37-60], and the remaining 5 studies (17%, 5/29) were conducted in upper middle-income countries [61-65].

Two studies investigated OCDE in dentistry [39,45], whereas the rest investigated it in medicine, including 6 in surgery or anesthesia [37,47,48,50,51,54], 7 in internal medicine or family medicine [40,44,49,53,57,60,65], 5 in pediatrics [38,41,46,56,58], 3 in psychiatry [52,61,63], 1 in obstetrics and gynecology [42], and 1 in radiology [43]. In the 4 remaining trials, the subject of intervention was evidence-based medicine [55] or advanced life support [59,62,64]. We included 1 cRCT in this review [38]. We obtained data at the participants' level to estimate the outcome effect; however, the effect of clustering was not adjusted for and sufficient information was not available to perform reanalysis to account for ICC (Multimedia Appendix 3).

XSL•FO
RenderX

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



Participants' Characteristics

The total number of participants included across all trials was 1690 in addition to 1 cRCT [38], which was conducted in 24 medical practices, but the number of participants was not specified. The study with the largest number of participants included 88 dentists in the control group and 86 in the intervention group [45], and the smallest study included 5 participants in the control and 6 participants in the intervention group [58]. Most of the participants were trainees. In addition, 4 studies were conducted among 377 medical interns [55,62-64], 16 trials among a total of 686 postgraduate residents [37,42,44,46-50,52,54,58-61,65], and 1 study included 49 residents and faculty members [43], of the remaining studies, 7 were performed on 578 practicing doctors [39-41,45,53,56,57] (Multimedia Appendix 3).

Intervention Characteristics

Overall, 15 studies compared OCDE with classroom or face-to-face learning [41,43,47-51,53-55,60-64], 3 studies with 4 interventions [37,46,58] compared OCDE with text-based resources, and 9 studies [38-40,42,44,45,52,56,59] compared OCDE with no intervention. The other 2 trials compared OCDE with another OCDE [57,65]. In addition, 23 trials reported the duration of the exposure to the intervention [38,40-42,46-51,53-65], which ranged between 2.5 min [58] and 12 months [38,59] (Multimedia Appendix 3).

All the studies used OCDE that was delivered by either personal computers or laptops. In addition, 19 studies used software- or computer-based programs delivered via a variety of sources such as CD-ROM and stored in the computer [38-41,44,46,47,50,52,53,55-57,59,60,62-65], 6 studies used video recording [42,48,49,51,58,61], 3 studies [43,45,54] investigated the use of multimedia, and 1 study used both computer-based program and video recording [37].

Primary Outcomes

As meta-analysis was not feasible, we presented the results in a narrative summary format (Multimedia Appendix 4) and Summary of Findings (Multimedia Appendix 3).

Doctors' Knowledge

Among the 29 studies that compared OCDE with other interventions or to no intervention, knowledge was assessed in 13 studies (44%) [41-45,47,49,55,59,60,63-65]. Knowledge gain was assessed by multiple choice questions (MCQs) in 5 studies [42,47,59,63,64] and by test of true or false questions in 2 studies [44,60], none of the tests were validated. Either non-validated or low-internal validity open-ended questions were used in 3 studies to assess knowledge [43,49,65]: 1 study used a validated questionnaire comprising MCQs and structured questions [55], 1 used non-validated Likert scale questions [45], and 1 used invalidated anonymous scoring by the authors compared with a gold standard prepared by 3 neurologists [41] (Multimedia Appendix 3).

Doctors' Cognitive Skills

Overall, 16 RCTs with 17 interventions [37,39,43,45-48,50,52-54,56,58,61,62,65] assessed cognitive skills as an outcome. A total of 3 studies [48,50,54] used the validated anesthetist nontechnical skills scoring system to assess doctors' cognitive skills. One study [62] used the objective structured clinical examination (OSCE) tool for evaluation of skills. Four

XSL•FO
RenderX

studies [52,53,58,65] used thematic analysis to assess skills gain, including thinking process [58], task completion rate [65], number of empathetic statement responses to patients [53], and agreement between expert and participant in mental health capacity assessment [52].

Cognitive skills gain was assessed by many tools including calculating the accuracy of decision making within and between the study groups [39], questionnaires [45,56], and a 36-item checklist [37]. Multiple assessment tools were used in 2 studies to assess cognitive skills, Ottolini et al [46] used 2 tools, thematic analysis and a questionnaire, whereas Esfahani et al [61] used both the Jefferson empathy scale and OSCE. A questionnaire with open-ended questions was used in 1 study [43], and an MCQ test was used in another [47] (Multimedia Appendix 3).

Patients' Outcomes

A total of 4 studies [38,40,53,57] assessed patient outcomes. Bonevski et al [57] examined doctors' improved screening and detection of the patients' risk behaviors using a self-reported survey. Millard et al [40] evaluated the improvement in dementia diagnosis, following a computer-generated audit of the participants' practices. Tulsky et al [53] used a telephone survey to evaluate the patients' trust in their oncologists and the oncologists' perceived empathy and knowledge of the patients, following CD-ROM–based education. Lavigne et al [38] evaluated patients' outcomes by measuring the improvement in the children's attention-deficit/hyperactivity disorder (ADHD) symptoms using the Beck Anxiety Inventory, Swanson, Nolan, and Pehlam-IV Rating Scale, and the ADHD Rating Scales-IV completed by teachers and parents following the use of computer software education in medication dose titration (Multimedia Appendix 3).

Secondary Outcomes

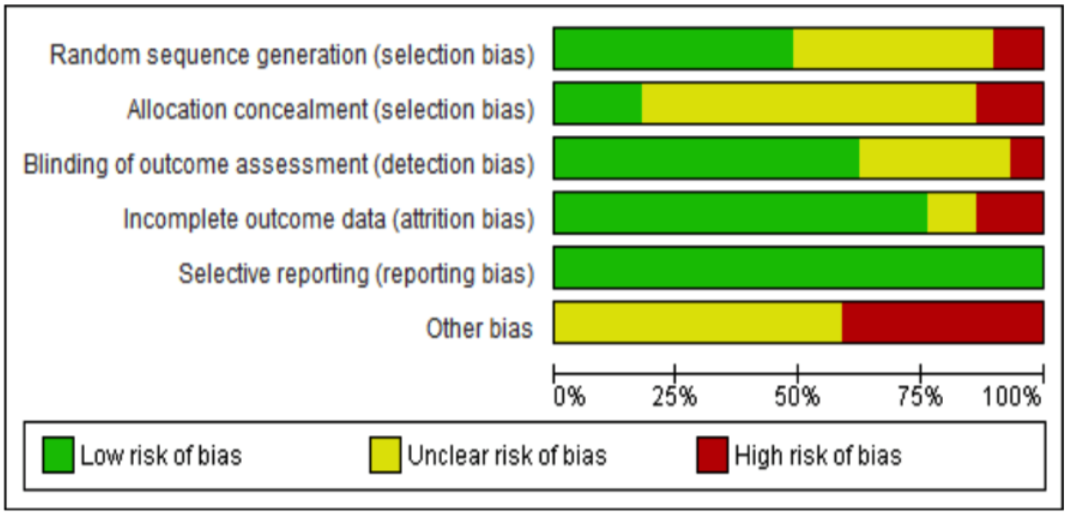
The CE of OCDE computer-based digital learning was examined in 1 study [51]. The study used data from an RCT included in this review [48]. It compared the cost and effectiveness of self-debriefing versus instructor debriefing using net benefit regression. The CE estimate was reported as the incremental net benefit, and the uncertainty was presented using a CE acceptability curve. The study concluded that digital learning was cost-effective if the intervention cost was less than or equal to Can $200 in the 2012 rate.

We did not find any RCT that compared the adverse effects of OCDE with other interventions.

Risk of Bias in Included Studies

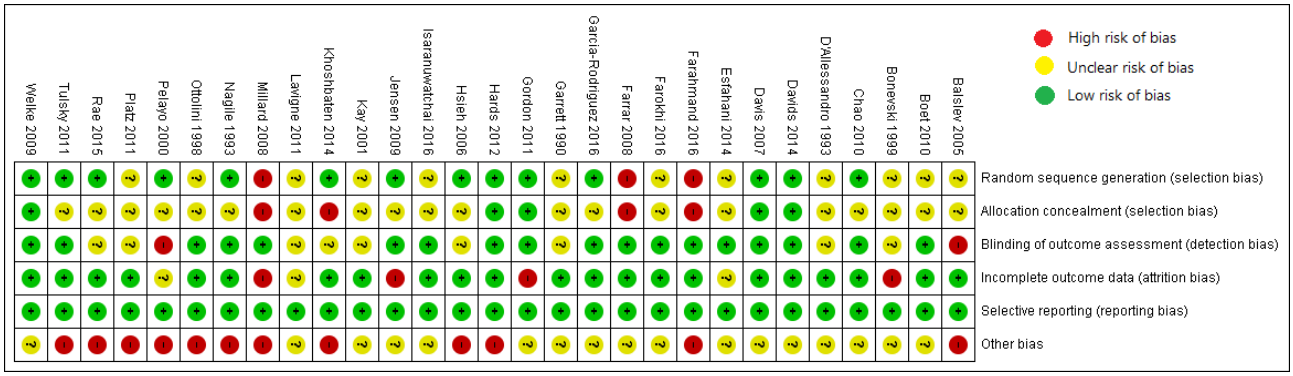
The assessment of risk of bias is described in detail in Multimedia Appendix 4 and shown on Figures 2 and 3. A total of 24 of the 29 included trials (83%) were judged to be at a high risk of bias [37-41,43-48,50-53,56-64]. In addition, 2 trials were judged to be of unclear risk of bias [42,49], whereas only 3 trials were judged to be at low risk of bias [54,55,65]. The assessment of the methodological quality of economic evaluation using the Evers checklist is presented in Multimedia Appendix 5.

Figure 2. Risk of bias graph.



XSL•FO
RenderX

Figure 3. Risk of bias summary.



Effects of the Interventions

The studies were divided into 4 comparisons, which evaluate the impact of OCDE compared with face-to-face learning, with no intervention, OCDE with text-based learning, and with another OCDE method.

Offline Computer-Based Digital Education Compared With Face-To-Face Learning

The characteristics of the studies are presented in (Multimedia Appendix 3); GRADE of evidence is presented in Summary of Findings (Multimedia Appendix 6).

Knowledge Gain

Overall, 4 studies [43,49,55,64] showed no significant difference in posttest knowledge scores between digital learning and face-to-face learning (ratio of the mean [RM] ranges from 1.0 to 1.13). In addition, 3 studies [41,60,63] showed the OCDE group to have significantly higher scores than the face-to-face group. However, the difference among the posttest mean scores of the participants was modest, as indicated by the small RM (RM=1.1, RM=1.17, and RM=1.13, respectively). One study [47] showed face-to-face participants to score higher than OCDE group (RM=0.95). The grade of evidence for this outcome is very low because of the high risk of bias in the included studies, the heterogeneity of participants and interventions, and the indirectness of evidence; therefore, it is uncertain whether there is difference in knowledge gain between OCDE and face-to-face learning.

Cognitive Skills Gain

Overall, 4 RCTs showed no significant difference between OCDE and face-to-face learning in posttest mean scores of skills gain [43,48,50,54] (RM range 0.94-1.0). In 2 studies [47,61], the mean posttest scores of face-to-face participants were significantly higher than those of the OCDE intervention (RM 0.91 and 0.95, respectively). In addition, 2 studies [62,53] showed the mean posttest score for the participants of the OCDE to be significantly higher than that of the face-to-face controls. The grade of evidence for this outcome is very low because of the high risk of bias in the included studies, the heterogeneity of participants and interventions, and the indirectness of evidence; therefore, it is uncertain whether there is a difference in cognitive skill gain between OCDE and face-to-face learning.

Patients' Outcomes

Patients' outcomes were examined in 2 trials [40,53]. Patients' outcomes significantly improved in the OCDE group. The grade of evidence for this outcome is low because of the high risk of bias and heterogeneity of participants and interventions; therefore, OCDE may improve patients' outcome compared with face-to-face learning.

Offline Computer-Based Digital Education Compared With No Intervention

The characteristics of the studies are presented in (Multimedia Appendix 3); GRADE of evidence is presented in Summary of Findings (Multimedia Appendix 6).

Knowledge Gain

Four RCTs [42,44,45,59] investigated the effect of OCDE compared with no intervention on knowledge gain. Of them, 3 trials [42,44,45] showed that OCDE was significantly more effective than no intervention with modest effect (RM range 1.11-1.36). The fourth trial [59] showed no significant difference in the posttest knowledge scores between intervention and control (RM=0.98). The grade of evidence for this outcome is low because of the high risk of bias and heterogeneity of participants and interventions; therefore, digital learning may improve knowledge gain compared with no intervention.

Overall, 3 trials [39,45,52] showed that OCDE had similar effect to no intervention in cognitive skills gain (RM=1.01), whereas Gordon et al [56] showed that participants in OCDE had significantly higher posttest scores compared with control (RM=1.25). The grade of evidence for this outcome is very low because of the high risk of bias, heterogeneity of participants and interventions, and indirectness of evidence; therefore, there is uncertainty about the effectiveness of OCDE compared with no intervention in cognitive skill gain.

Patients' Outcomes

Only 1 cRCT, at high risk of bias, investigated the effect of OCDE compared with no intervention on patients' outcome [38]. The trial was conducted among doctors of 24 pediatric practices where the number of doctors was not specified. It showed similar effectiveness in the treatment of patients with ADHD in intervention and control groups. The grade of evidence for this outcome was low because of the high risk of bias and the fact that evidence was drawn from a single study; therefore, OCDE may have an equal effect as no intervention in patients' outcomes.

XSL•FO
RenderX

Offline Computer-Based Digital Education Compared With Text-Based Learning: Cognitive Skills Gain

The characteristics of the studies are presented in Multimedia Appendix 3; GRADE of evidence is presented in Summary of Findings Table (Multimedia Appendix 6).

Overall, 3 RCTs with 4 interventions [37,46,58] investigated the effect of OCDE compared with text-based learning on cognitive skills gain. In the 3 comparisons [37,46,58], OCDE was significantly more effective than text (RM range 1.14-1.46). In the fourth comparison [37], there was no difference in the posttest scores between the intervention and control groups (RM=0.91). The grade of evidence is very low because of the high risk of bias, small number of participants, and indirectness of evidence; therefore, there is uncertainty about the effect of OCDE compared with text-based learning in cognitive skills gain.

Offline Computer-Based Digital Education Compared With Other Digital Learning

The characteristics of the studies are presented in Multimedia Appendix 3.

Knowledge Gain

Only 1 trial at low risk of bias investigated knowledge gain [65]. It showed no difference in the effects of 2 methods of digital learning (RM=0.98). The body of evidence is considered low grade as the evidence is driven from a single study with a small number of participants.

Cognitive Skills Gain

Only 1 RCT, at low risk of bias, in this comparison investigated cognitive skill gain [65]. It showed equal effects from 2 methods of OCDE (RM=0.98). The body of evidence is considered low grade as the evidence is driven from a single study with a small number of participants.

Patients' Outcomes

One RCT compared offline computer-based CME with feedback with the same CME without feedback in improving screening behavior as patients' outcome [57]. The RCT showed better patients' outcomes for CME with audit compared with the same program without audit. The body of evidence is considered very low grade as the evidence is driven from a single study at high risk of bias with a small number of participants.

Discussion

Principal Findings

This systematic review showed that the effectiveness of OCDE compared with other methods of education, on medical doctors' knowledge and cognitive skill gain, is uncertain. OCDE may improve doctors' knowledge compared with no intervention, but its effect on doctors' cognitive skills is uncertain. OCDE may have little or no effect in improving patients' outcomes.

The evidence for this review is driven from 29 RCTs, which covered a wide range of offline digital learning interventions in a variety of clinical and nonclinical medical disciplines. The studies investigated multiple outcomes of the intervention in 1690 doctors and dentists; therefore, they provide a considerable body of evidence. However, heterogeneity of participants, interventions, and methods of assessment of outcomes, in addition to the poor methodological quality of the trials, resulted in uncertainty about the effectiveness of OCDE compared with other instruction methods.

The quality of evidence for all outcomes was rated as low or very low (for different outcomes) on the GRADE scale because of the poor methodological quality of the included studies, as 24 out of the 29 included studies were judged to be at high risk of bias in addition to the marked clinical heterogeneity of the body of evidence.

It is worth noting that in all trials that compared OCDE with other types of learning, the measured outcome was the participants' improved knowledge or skills (contents) rather than the methods of learning. This surrogate outcome (indirectness of evidence) is valid for the evaluation of the methods of learning as long as the assumption that the participants in the intervention and control groups had equal baseline knowledge with respect to the contents of the interventions is valid. Nevertheless, bias can be introduced if the participants have different levels of knowledge about the contents (eg, same content was taught in medical school, participants at different level of training) and no pre-intervention test was performed or was completed with an invalidated tool, which is the case in most of the trials included in this review. Furthermore, the body of evidence in this review has been drawn from small individual studies, as 12 of the included studies had less than 50 participants.

The external validity of the interventions in this review has been compromised by the recruitment of volunteers, which might have resulted in the selection of participants who were more computer literate, and therefore overestimated the effect of offline OCDE by excluding participants who did not know how to use the technology or were unwilling to do so. We believe variation in computer literacy and cultural differences may influence the generalization of our results to LMICs. Furthermore, most of the included studies were experimental trials conducted in ideal university hospital settings rather than implemented in programs in the field; therefore, the true applicability of OCDE could not be examined by this review.

The results of this review are inconsistent with previous evidence about the effectiveness of digital learning for health care professionals in improving knowledge and skills gain. A systematic review of 15 RCTs on the effects of digital learning (both on the Web and offline) showed digital learning to outperform or have equal effects as face-to-face learning in knowledge gain and practice improvement [10]. Similar effects of Web-based continuing education compared with face-to-face learning for medical doctors were found by Wulto et al [66] in their systematic review, which included 16 RCTs. Another systematic review [67], which examined the effectiveness of computer-based programs on the dentists' performance, time spent, and attitude toward the programs, showed that in all the 12 included studies except 1, the computer-based programs were either better or similar to face-to-face learning in knowledge gain and that dentists had a positive attitude toward

XSL•FO
RenderX

the program. More recent systematic reviews that investigated the effectiveness of a specific type of digital learning on the knowledge and skills gain of health professionals have shown similar results [68,69]. However, our conclusion of the effect of OCDE has been based on grading the evidence base of the effectiveness of OCDE, which we believe gives a more accurate evaluation of the effectiveness of the intervention on the desired outcomes. To that end, our conclusion agrees with a recently published Cochrane systematic review [70], which considered grading of the body of evidence an integral part of its conclusion.

Strengths and Limitations

To complete this review, we followed the robust methodology outlined by the Cochrane collaboration for searching, assessing, and reporting of the body of evidence for the effectiveness of OCDE in improving medical doctors' knowledge and cognitive skills.

The review comprehensively evaluated the OCDE for medical doctors and dentists. The participants of the included studies are representative of the target population of medical doctors and dentists in training and non-training posts. Furthermore, the interventions accommodated participants from almost all clinical disciplines in addition to 4 fields of general knowledge and skills, including evidence-based medicine, research methodology, advanced life support, and cardiopulmonary resuscitation.

However, the heterogeneity of the participants, interventions, and outcomes precluded meta-analysis and subgroup analysis, which would have improved our certainty about the effectiveness of the intervention.

Implications for Practice and Research

The uncertainty associated with the effectiveness of OCDE for medical doctors' education calls for limited-scale implementation of OCDE in the context of experimental settings and research.

Research in digital education should be employed to investigate effectiveness in updating medical doctors' knowledge and skills, considering the patient as the center of care and the improvement of patients' health as the main outcome, especially in LMICs. Future trials should follow a robust methodology, focusing on avoiding major biases by employing valid methods for randomization and allocation concealment, in addition to the use of validated tests to assess the outcomes. As most of the participants are not blinded to the intervention in these types of studies, there is high risk of attrition bias for any outcome that relied on active participation and follow-up (eg, demonstrating skills or taking a knowledge test). Such bias can be reduced by securing the anonymity of the participants, for example, replacing their names with numbers or letters.

The indirectness of evidence will continue to downgrade the evidence base of the effectiveness of OCDE unless validated pre and posttests are used to evaluate the outcomes in addition to attentive selection of trial participants who have no previous knowledge about the subject of the education (eg, new imaging technique).

Furthermore, we believe that the development of a common taxonomy for digital learning will facilitate easier comparison among studies and therefore better the quality of evidence.

Evaluating the CE of the various methods of digital learning is an important field for future research, considering the need of such interventions in LMICs.

Conclusions

The effectiveness of OCDE when compared with other methods of education, on medical doctors' knowledge and cognitive skill gain, is uncertain. OCDE may improve doctors' knowledge when compared with no intervention, but its effect on doctors' cognitive skills is uncertain. OCDE may have little or no effect in improving patients' outcome.

Acknowledgments

This review was conducted in collaboration with the Health Workforce Department at the World Health Organization. The authors would also like to thank Mr Carl Gornitzki, Ms GunBrit Knutssön, and Mr Klas Moberg from University Library, Karolinska Institute for developing and undertaking our search strategy. We gratefully acknowledge King Saud University, Deanship of Scientific Research, Research Chair for Evidence-Based Health Care and Knowledge Translation for funding the reviewers from King Saud University and the funding from Nanyang Technological University and National Healthcare Group Polyclinics, Singapore for this project. We gratefully acknowledge funding from the Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, e-learning for health professionals education grant.

Authors' Contributions

HW is the main review author and has been involved with all aspects of the review. HW, MS, PP, and LTC wrote the protocol and commented on the draft of the review. HW, SE, YA, AF, AJ, NZ, MT, KB, and ARS screened eligible studies, conducted the quality assessment of eligible studies, and extracted data from the included studies. HW, SE, YA, AF, AJ, MT, KB, and ARS assessed the risk of bias of the included studies. HW and ARS conducted the data synthesis and performed the GRADE data synthesis. HW, SE, YA, MT, KB, AF, and PP wrote the draft of the review.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Search Strategy.

[PDF File (Adobe PDF File), 20KB - jmir_v21i2e12998_app1.pdf]

Multimedia Appendix 2

Excluded studies.

[PDF File (Adobe PDF File), 43KB - jmir_v21i2e12998_app2.pdf]

Multimedia Appendix 3

Characteristics of the included studies.

[PDF File (Adobe PDF File), 40KB - jmir_v21i2e12998_app3.pdf]

Multimedia Appendix 4

Risk of bias summary.

[PDF File (Adobe PDF File), 28KB - jmir_v21i2e12998_app4.pdf]

Multimedia Appendix 5

Evers checklist for economic evaluation.

[PDF File (Adobe PDF File), 21KB - jmir_v21i2e12998_app5.pdf]

Multimedia Appendix 6

Summary of findings tables.

[PDF File (Adobe PDF File), 30KB - jmir_v21i2e12998_app6.pdf]

References

1. Davis N, Davis D, Bloch R. Continuing medical education: AMEE Education Guide No 35. Med Teach 2008;30(7):652-666. [doi: 10.1080/01421590802108323] [Medline: 18777424]
2. Paixão MP, Miot HA, Wen CL. Tele-education on leprosy: evaluation of an educational strategy. Telemed J E Health 2009;15(6):552-559. [doi: 10.1089/tmj.2008.0137] [Medline: 19659411]
3. Lawton S, Wimpenny P. Continuing professional development: a review. Nurs Stand 2003;17(24):41-44. [Medline: 12655859]
4. Marinopoulos SS, Dorman T, Ratanawongsa N, Wilson LM, Ashar BH, Magaziner JL, et al. Effectiveness of continuing medical education. Evid Rep Technol Assess (Full Rep) 2007 Jan(149):1-69. [Medline: 17764217]
5. O'Neil KM, Addrizzo-Harris DJ, American College of Chest Physicians Health and Science Policy Committee. Continuing medical education effect on physician knowledge application and psychomotor skills: effectiveness of continuing medical education: American College of Chest Physicians Evidence–Based Educational Guidelines. Chest 2009 Mar;135(3 Suppl):37S-41S. [doi: 10.1378/chest.08-2516] [Medline: 19265074]
6. Thepwongsa I, Kirby CN, Schattner P, Piterman L. Online continuing medical education (CME) for GPs: does it work? A systematic review. Aust Fam Physician 2014 Oct;43(10):717-721 [FREE Full text] [Medline: 25286431]
7. Masic I. E-learning as new method of medical education. Acta Inform Med 2008;16(2):102-117 [FREE Full text] [doi: 10.5455/aim.2008.16.102-117] [Medline: 24109154]
8. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet–based learning for health professions education: a systematic review and meta-analysis. Acad Med 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]
9. Sinclair PM, Kable A, Levett-Jones T, Booth D. The effectiveness of Internet–based e-learning on clinician behaviour and patient outcomes: a systematic review. Int J Nurs Stud 2016 May;57:70-81. [doi: 10.1016/j.ijnurstu.2016.01.011] [Medline: 27045566]
10. Lam-Antoniades M, Ratnapalan S, Tait G. Electronic continuing education in the health professions: an update on evidence from RCTs. J Contin Educ Health Prof 2009;29(1):44-51. [doi: 10.1002/chp.20005] [Medline: 19288566]
11. Bouhnik D, Marcus T. Interaction in distance-learning courses. J Am Soc Inf Sci 2006 Feb 1;57(3):299-305 [FREE Full text] [doi: 10.1002/asi.20277]

XSL•FO
RenderX

12. Andreazzi DB, Rossi F, Wen CL. Interactive tele-education applied to a distant clinical microbiology specialization university course. Telemed J E Health 2011 Sep;17(7):524-529. [doi: 10.1089/tmj.2011.0010] [Medline: 21749260]
13. Sezer B. Faculty of medicine students' attitudes towards electronic learning and their opinion for an example of distance learning application. Comput Human Behav 2016 Feb:932-939 [FREE Full text] [doi: 10.1016/j.chb.2015.10.018]
14. Cantoni V, Cellario M, Porta M. Perspectives and challenges in e-learning: towards natural interaction paradigms. J Vis Lang Comput 2004 Oct:333-345 [FREE Full text] [doi: 10.1016/j.jvlc.2003.10.002]
15. Cook DA. Web-based learning: pros, cons and controversies. Clin Med (Lond) 2007;7(1):37-42 [FREE Full text] [doi: 10.7861/clinmedicine.7-1-37] [Medline: 17348573]
16. Alur P, Fatima K, Joseph R. Medical teaching websites: do they reflect the learning paradigm? Med Teach 2002 Jul;24(4):422-424. [doi: 10.1080/01421590220145815] [Medline: 12193328]
17. Carroll C, Booth A, Papaioannou D, Sutton A, Wong R. UK health-care professionals' experience of on-line learning techniques: a systematic review of qualitative data. J Contin Educ Health Prof 2009;29(4):235-241. [doi: 10.1002/chp.20041] [Medline: 19998474]
18. Greenhalgh T. Computer assisted learning in undergraduate medical education. Br Med J 2001 Jan 6;322(7277):40-44 [FREE Full text] [Medline: 11141156]
19. Triola MM, Huwendiek S, Levinson AJ, Cook DA. New directions in e-learning research in health professions education: report of two symposia. Med Teach 2012;34(1):e15-e20. [doi: 10.3109/0142159X.2012.638010] [Medline: 22250691]
20. Balslev T, de Grave W, Muijtjens AM, Eika B, Scherpbier AJ. The development of shared cognition in paediatric residents analysing a patient video versus a paper patient case. Adv Health Sci Educ Theory Pract 2009 Oct;14(4):557-565. [doi: 10.1007/s10459-008-9138-y] [Medline: 18846433]
21. Cook DA, McDonald FS. E-learning: is there anything special about the. Perspect Biol Med 2008;51(1):5-21. [doi: 10.1353/pbm.2008.0007] [Medline: 18192762]
22. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5. United States; 2011. URL: http://handbook-5-1.cochrane.org/
23. Car J, Carlstedt-Duke J, Tudor CL, Posadzki P, Whiting P, Zary N, et al. Digital education for health professions: methods for overarching evidence syntheses J Med Internet Res. Journal of medical internet research 2019:-(forthcoming)(forthcoming). [doi: 10.2196/preprints.12913]
24. Ward JP, Gordon J, Field MJ, Lehmann HP. Communication and information technology in medical education. Lancet 2001 Mar 10;357(9258):792-796. [doi: 10.1016/S0140-6736(00)04173-8] [Medline: 11253986]
25. Rasmussen K, Belisario JM, Wark PA, Molina JA, Loong SL, Cotic Z, et al. Offline eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010405 [FREE Full text] [doi: 10.7189/jogh.04.010405] [Medline: 24976964]
26. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. Br Med J 2003 Oct 4;327(7418):785-789 [FREE Full text] [doi: 10.1136/bmj.327.7418.785] [Medline: 14525877]
27. Davey P, Brown E, Charani E, Fenelon L, Gould IM, Holmes A, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev 2013;4:CD003543. [doi: 10.1002/14651858.CD003543.pub3] [Medline: 23633313]
28. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care 2005;21(2):240-245. [doi: 10.1017/S0266462305050324] [Medline: 15921065]
29. GRADE Working Group. 2014. URL: http://www.gradeworkinggroup.org/ [accessed 2019-02-09] [WebCite Cache ID 763K8bxGJ]
30. Schünemann H, Oxman A, Higgins J, Vist G, Glasziou P, Guyatt G. Cochrane Handbook for Systematic Reviews of Interventions. United States; 2011. Presenting results and Summary of findings' tables URL: https://handbook-5-1.cochrane.org/chapter_11/11_10_chapter_information.htm [accessed 2019-02-09] [WebCite Cache ID 763Ocethg]
31. Bharati A, Datta A, Gupta S, Chun-Lema P, Ryan J. Resident education in ultrasonography: assessment of identification of basic cardiac anatomy and pathology after a multimedia tutorial. Ann Emerg Med S 2012;60(4):S41. [doi: 10.1016/j.annemergmed.2012.06.089]
32. Juyal M, Ware A, Houk J, Mina R. Electronic learning module enhances rheumatology education. 2013 Presented at: 2013 ACR/ARHP Annual Meeting; October 25-30, 2013; San Diego, California, US URL: https://acrabstracts.org/abstract/electronic-learning-module-enhances-rheumatology-education/
33. Khan T, Gupta N, Wani SR. Comparison of classroom didactic teaching (DT) and computer based self-learning (Sl) in recognizing narrow band imaging (NBI) patterns for diminutive polyp (DP) histology characterization. Gastrointest Endosc 2015;81(5):AB325. [doi: 10.1016/j.gie.2015.03.1449]
34. Lichtman AS, Goff B, Schreuder HW, Lenihan JJ, Lefebvre G, Parker W. A validation study: does use of an interactive computer based laparoscopic hysterectomy trainer expand cognitive surgical knowledge of ob/gyn trainees beyond that of traditional educational tools? J Minim Invasive Gynecol 2014;21(6):S87. [doi: 10.1016/j.jmig.2014.08.304]

XSL•FO
RenderX

35. Muhammad A, Nutland S, Unsworth J. Electronic vs standard learning in hospital trust induction for junior doctors - a randomized trial. Br J Surg 2010;97(S2):49-50 [FREE Full text]
36. Springer. 2015 Apr. Abstracts from the 38th Annual Meeting of the Society of General Internal Medicine URL: https://link.springer.com/content/pdf/10.1007%2Fs11606-015-3271-0.pdf
37. Rae A, Khatib M, Sarker S, Bello F. The Effect of a Computer Based Open Surgery Simulation of an Inguinal Hernia Repair on the Results of Cognitive Task Analysis Performance of Surgical Trainees: An Educational Trial. 2015 Presented at: ASGBI International Surgical Congress 2015; April 22-24, 2015; Manchester, England.
38. Lavigne JV, Dulcan MK, LeBailly SA, Binns HJ, Cummins TK, Jha P. Computer-assisted management of attention-deficit/hyperactivity disorder. Pediatrics 2011 Jul;128(1):e46-e53 [FREE Full text] [doi: 10.1542/peds.2010-2684] [Medline: 21669891]
39. Kay EJ, Silkstone B, Worthington HV. Evaluation of computer aided learning in developing clinical decision-making skills. Br Dent J 2001 May 26;190(10):554-557 [FREE Full text] [doi: 10.1038/sj.bdj.4801033a] [Medline: 11411890]
40. Millard FB, Thistlehwaite J, Spagnolo C, Kennedy LR, Baune BT. Dementia diagnosis: a pilot randomised controlled trial of education and IT audit to assess change in GP dementia documentation. Aust J Prim Health 2008 Jan 10;14(3):141-149 [FREE Full text] [doi: 10.1071/PY08046]
41. Farrar M, Connolly AM, Lawson J, Burgess A, Lonergan A, Bye AM. Teaching doctors how to diagnose paroxysmal events: a comparison of two educational methods. Med Educ 2008 Sep;42(9):909-914. [doi: 10.1111/j.1365-2923.2008.03146.x] [Medline: 18715488]
42. Chao TT, Wendel GD, McIntire DD, Corton MM. Effectiveness of an instructional DVD on third- and fourth-degree laceration repair for obstetrics and gynecology postgraduate trainees. Int J Gynaecol Obstet 2010 Apr;109(1):16-19. [doi: 10.1016/j.ijgo.2009.10.016] [Medline: 20022005]
43. D'Alessandro MP, Galvin JR, Erkonen WE, Albanese MA, Michaelsen VE, Huntley JS, et al. The instructional effectiveness of a radiology multimedia textbook (HyperLung) versus a standard lecture. Invest Radiol 1993 Jul;28(7):643-648. [doi: 10.1097/00004424-199307000-00017] [Medline: 8344816]
44. Garrett TJ, Selnow G, Dobkin JF, Healton C. Computer-assisted instruction in aids infection control for physicians. Teach Learn Med 2009:2015-2218 [FREE Full text] [doi: 10.1080/10401339009539463]
45. Hsieh NK, Herzig K, Gansky SA, Danley D, Gerbert B. Changing dentists' knowledge, attitudes and behavior regarding domestic violence through an interactive multimedia tutorial. J Am Dent Assoc 2006 May;137(5):596-603. [doi: 10.14219/jada.archive.2006.0254] [Medline: 16739538]
46. Ottolini MC, Greenberg L. Development and evaluation of a CD-ROM computer program to teach residents telephone management. Pediatrics 1998 Mar;101(3):E2. [doi: 10.1542/peds.101.3.e2] [Medline: 9481021]
47. Platz E, Liteplo A, Hurwitz S, Hwang J. Are live instructors replaceable? Computer vs classroom lectures for EFAST training. J Emerg Med 2011 May;40(5):534-538. [doi: 10.1016/j.jemermed.2009.08.030] [Medline: 19892506]
48. Boet S. Self debriefing versus instructor debriefing: a prospective randomized trial. Can J Anesth/J Can Anesth 2010 Nov 13;57(S1):S26-S26 [FREE Full text] [doi: 10.1007/s12630-010-9415-0]
49. Garcia-Rodriguez JA, Donnon T. Using comprehensive video-module instruction as an alternative approach for teaching IUD insertion. Fam Med 2016 Jan;48(1):15-20 [FREE Full text] [Medline: 26950661]
50. Hards A, Davies S, Salman A, Erik-Soussi M, Balki M. Management of simulated maternal cardiac arrest by residents: didactic teaching versus electronic learning. Can J Anaesth 2012 Sep;59(9):852-860. [doi: 10.1007/s12630-012-9752-2] [Medline: 22777579]
51. Isaranuwatchai W, Alam F, Hoch J, Boet S. A cost-effectiveness analysis of self-debriefing versus instructor debriefing for simulated crises in perioperative medicine in Canada. J Educ Eval Health Prof 2017;13:44 [FREE Full text] [doi: 10.3352/jeehp.2016.13.44] [Medline: 28028288]
52. Naglie G, Silberfeld M, O'Rourke K, Fried B, Corber W, Bombardier C, et al. A randomized trial of a decisional aid for mental capacity assessments. J Clin Epidemiol 1993 Mar;46(3):221-230. [doi: 10.1016/0895-4356(93)90069-D] [Medline: 8455046]
53. Tulsky JA, Arnold RM, Alexander SC, Olsen MK, Jeffreys AS, Rodriguez KL, et al. Enhancing communication between oncologists and patients with a computer-based training program: a randomized trial. Ann Intern Med 2011 Nov 1;155(9):593-601 [FREE Full text] [doi: 10.7326/0003-4819-155-9-201111010-00007] [Medline: 22041948]
54. Welke TM, LeBlanc VR, Savoldelli GL, Joo HS, Chandra DB, Crabtree NA, et al. Personalized oral debriefing versus standardized multimedia instruction after patient crisis simulation. Anesth Analg 2009 Jul;109(1):183-189. [doi: 10.1213/ane.0b013e3181a324ab] [Medline: 19535709]
55. Davis J, Chryssafidou E, Zamora J, Davies D, Khan K, Coomarasamy A. Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: a randomised controlled trial. BMC Med Educ 2007 Jul 20;7:23 [FREE Full text] [doi: 10.1186/1472-6920-7-23] [Medline: 17659076]
56. Gordon M, Chandratilake M, Baker P. Improved junior paediatric prescribing skills after a short e-learning intervention: a randomised controlled trial. Arch Dis Child 2011 Dec;96(12):1191-1194. [doi: 10.1136/archdischild-2011-300577] [Medline: 22080459]

XSL•FO
RenderX

57. Bonevski B, Sanson-Fisher RW, Campbell E, Carruthers A, Reid AL, Ireland M. Randomized controlled trial of a computer strategy to increase general practitioner preventive care. Prev Med 1999 Dec;29(6 Pt 1):478-486. [doi: 10.1006/pmed.1999.0567] [Medline: 10600428]
58. Balslev T, de GW, Muijtjens AM, Scherpbier AJ. Comparison of text and video cases in a postgraduate problem-based learning format. Med Educ 2005 Nov;39(11):1086-1092. [doi: 10.1111/j.1365-2929.2005.02314.x] [Medline: 16262803]
59. Jensen ML, Mondrup F, Lippert F, Ringsted C. Using e-learning for maintenance of ALS competence. Resuscitation 2009 Aug;80(8):903-908. [doi: 10.1016/j.resuscitation.2009.06.005] [Medline: 19570601]
60. Pelayo-Alvarez M, Albert-Ros X, Gil-Latorre F, Gutierrez-Sigler D. Feasibility analysis of a personalized training plan for learning research methodology. Med Educ 2000 Feb;34(2):139-145. [doi: 10.1046/j.1365-2923.2000.00383.x] [Medline: 10652067]
61. Esfahani MN, Behzadipour M, Nadoushan AJ, Shariat SV. A pilot randomized controlled trial on the effectiveness of inclusion of a distant learning component into empathy training. Med J Islam Repub Iran 2014;28:65 [FREE Full text] [Medline: 25405130]
62. Farahmand S, Jalili E, Arbab M, Sedaghat M, Shirazi M, Keshmiri F, et al. Distance learning can be as effective as traditional learning for medical students in the initial assessment of trauma patients. Acta Med Iran 2016 Sep;54(9):600-604 [FREE Full text] [Medline: 27832693]
63. Farokhi RM, Zarifsanaiey N, Haghighi F, Mehrabi M. E-learning or in-person approaches in continuous medical education: a comparative study. IIOAB J 2016;7(2):472-477 [FREE Full text]
64. Khoshbaten M, Soleimanpour H, Ala A, Vahdati SS, Ebrahimian K, Safari S, et al. Which form of medical training is the best in improving interns' knowledge related to advanced cardiac life support drugs pharmacology? An educational analytical intervention study between electronic learning and lecture-based education. Anesth Pain Med 2014 Feb;4(1):e15546 [FREE Full text] [doi: 10.5812/aapm.15546] [Medline: 24719802]
65. Davids MR, Chikte UM, Halperin ML. Effect of improving the usability of an e-learning resource: a randomized trial. Adv Physiol Educ 2014 Jun;38(2):155-160 [FREE Full text] [doi: 10.1152/advan.00119.2013] [Medline: 24913451]
66. Wutoh R, Boren SA, Balas EA. eLearning: a review of internet-based continuing medical education. J Contin Educ Health Prof 2004;24(1):20-30. [doi: 10.1002/chp.1340240105] [Medline: 15069909]
67. Rosenberg H, Grad HA, Matear DW. The effectiveness of computer-aided, self-instructional programs in dental education: a systematic review of the literature. J Dent Educ 2003 May;67(5):524-532 [FREE Full text] [Medline: 12809187]
68. Ilgen JS, Sherbino J, Cook DA. Technology-enhanced simulation in emergency medicine: a systematic review and meta-analysis. Acad Emerg Med 2013 Feb;20(2):117-127 [FREE Full text] [doi: 10.1111/acem.12076] [Medline: 23406070]
69. Mundell WC, Kennedy CC, Szostek JH, Cook DA. Simulation technology for resuscitation training: a systematic review and meta-analysis. Resuscitation 2013 Sep;84(9):1174-1183. [doi: 10.1016/j.resuscitation.2013.04.016] [Medline: 23624247]
70. Vaona A, Banzi R, Kwag KH, Rigon G, Cereda D, Pecoraro V, et al. E-learning for health professionals. Cochrane Database Syst Rev 2018 Jan 21;1:CD011736. [doi: 10.1002/14651858.CD011736.pub2] [Medline: 29355907]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
CE: cost-effectiveness
CME: continuing medical education
CPD: continuing professional development
cRCT: cluster randomized controlled trial
GRADE: Grading of Recommendations Assessment, Development and Evaluation
LMICs: low- and middle-income countries
MCQs: multiple choice questions
MEDLINE: Medical Literature Analysis and Retrieval System Online
OCDE: offline computer-based digital education
OSCE: objective structured clinical examination
RCT: randomized controlled trial
RM: ratio of the means
ICC: intracluster correlation coefficients

XSL•FO
RenderX

Edited by A Marusic; submitted 30.11.18; peer-reviewed by D Adeloye, Y Subhi; comments to author 29.12.18; revised version received 06.01.19; accepted 31.01.19; published 01.03.19.

Please cite as:
Wahabi HA, Esmaeil SA, Bahkali KH, Titi MA, Amer YS, Fayed AA, Jamal A, Zakaria N, Siddiqui AR, Semwal M, Car LT, Posadzki P, Car J
Medical Doctors' Offline Computer-Assisted Digital Education: Systematic Review by the Digital Health Education Collaboration
J Med Internet Res 2019;21(3):e12998
URL: https://www.jmir.org/2019/3/e12998/
doi:10.2196/12998
PMID:30821689

©Hayfaa Abdelmageed Wahabi, Samia Ahmed Esmaeil, Khawater Hassan Bahkali, Maher Abdelraheim Titi, Yasser Sami Amer, Amel Ahmed Fayed, Amr Jamal, Nasriah Zakaria, Amna Rehana Siddiqui, Monika Semwal, Lorainne Tudor Car, Paul Posadzki, Josip Car. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Digital Education for Health Professions on Smoking Cessation Management: Systematic Review by the Digital Health Education Collaboration

Monika Semwal[1], PhD, BPharm, MPharm; Penny Whiting[2,3], PhD; Ram Bajpai[1], PhD; Shweta Bajpai[1], PhD; Bhone Myint Kyaw[4], MBBS, PhD; Lorainne Tudor Car[4,5], MD, MSc, PhD

[1]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Bristol Medical School, University of Bristol, Bristol, United Kingdom

[3]National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West, University Hospitals Bristol NHS Foundation Trust, Bristol, United Kingdom

[4]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[5]Department of Primary Care and Public health, School of Public health, Imperial College London, London, United Kingdom

Corresponding Author:
Lorainne Tudor Car, MD, MSc, PhD
Family Medicine and Primary Care
Lee Kong Chian School of Medicine
Nanyang Technological University
11 Mandalay Road
Singapore, 308232
Singapore
Phone: 65 6904 1258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Tobacco smoking, one of the leading causes of preventable death and disease, is associated with 7 million deaths every year. This is estimated to rise to more than 8 million deaths per year by 2030, with 80% occurring in low- and middle-income countries. Digital education, teaching, and learning using digital technologies have the potential to increase educational opportunities, supplement teaching activities, and decrease distance barriers in health professions education.

Objective: The primary objective of this systematic review was to evaluate the effectiveness of digital education compared with various controls in improving learners' knowledge, skills, attitudes, and satisfaction to deliver smoking cessation therapy. The secondary objectives were to assess patient-related outcomes, change in health professionals' practice or behavior, self-efficacy or self-rated competence of health professionals in delivering smoking cessation therapy, and cost-effectiveness of the interventions.

Methods: We searched 7 electronic databases and 2 trial registers for randomized controlled trials published between January 1990 and August 2017. We used gold standard Cochrane methods to select and extract data and appraise eligible studies.

Results: A total of 11 studies (number of participants, n=2684) were included in the review. All studies found that digital education was at least as effective as traditional or usual learning. There was some suggestion that blended education results in similar or greater improvements in knowledge (standardized mean difference, SMD=0.19, 95% CI −0.35 to 0.72), skill (SMD=0.58, 95% CI 0.08-1.08), and satisfaction (SMD=0.62, 95% CI 0.12-1.12) compared with digital education or usual learning alone. There was also some evidence for improved attitude (SMD=0.45, 95% CI 0.18-0.72) following digital education compared with usual learning. Only 1 study reported patient outcomes and the setup cost of blended education but did not compare outcomes among groups. There were insufficient data to investigate what components of the digital education interventions were associated with the greatest improvements in learning outcomes.

Conclusions: The evidence suggests that digital education is at least as effective as usual learning in improving health professionals' knowledge and skill for delivering smoking cessation therapy. However, limitations in the evidence base mean that these conclusions should be interpreted with some caution.

Trial Registration: PROSPERO CRD42016046815; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=46815

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e13000) doi:10.2196/13000

KEYWORDS

evidence-based practice; health personnel; learning; systematic review; smoking cessation

Introduction

Tobacco smoking, one of the leading causes of preventable death and disease, is associated with 7 million deaths every year. This is estimated to rise to more than 8 million deaths per year by 2030, with 80% occurring in low- and middle-income countries [1]. There is promising evidence to show that interventions delivered by health professionals are effective in preventing and stopping tobacco smoking [2-4]. Health professionals' advice has been shown to increase attempts to stop smoking and use medications aimed at stopping smoking. An intervention as brief as 3 min can result in a significant increase in smoking cessation rates [5]. However, the lack of relevant knowledge and skill is a significant barrier reported by health professionals, preventing them from providing smoking cessation advice to their patients who are tobacco smokers [6,7]. Timely and cost-effective training and education are essential to ensure that health professionals have appropriate knowledge and skill to deliver smoking cessation–related interventions.

Digital education, teaching, and learning using digital technologies have the potential to increase educational opportunities, supplement teaching activities, and decrease distance barriers in health professional education [8]. Encompassing a broad spectrum of interventions, digital education can combine self-directed learning with practical skill-based training to successfully deliver smoking cessation therapy to health professionals [9,10]. The different intervention modalities such as mobile phone apps, computer-assisted learning, simulation-based learning, and social networking can create active learning environments that provide real-time feedback and enable health professionals to participate in in-depth discussions of relevant topics. Digital education also offers the opportunity for professionals to become specialists in comprehensive smoking cessation by delivering more smoking cessation training than usually provided in the traditional classroom setting [8,11,12].

Previous systematic reviews that have evaluated digital education for smoking cessation and prevention have focused on interventions aimed at patients rather than training health professionals [13-16]. This systematic review is 1 of a series of reviews evaluating the scope for implementation and potential impact of a wide range of digital health education interventions for pre and postregistration health professionals. The primary objective of this systematic review is to evaluate the effectiveness of digital education compared with various controls in improving learners' knowledge, skills, attitudes, and satisfaction to deliver smoking cessation therapy. The secondary objectives are to assess patient-related outcomes, change in health professionals' practice or behavior, self-efficacy or self-rated competence of health professionals in delivering smoking cessation therapy, and cost-effectiveness of the interventions.

If digital education is at least as effective as standard face-to-face learning methods, then there is the potential for digital education to be used to deliver training in smoking cessation therapy with associated benefits such as being able to reach a much larger audience and allowing more flexibility in when and where the training is undertaken. The evidence could be used to make recommendations regarding the optimal digital education approach to train health professionals to deliver smoking cessation therapy.

Methods

The Cochrane recommendations for the conduct of systematic review were followed, and this review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidance [17,18]. A protocol detailing the review methods was also produced and followed throughout the review. This protocol was registered with PROSPERO (ID: CRD42016046815). For a detailed description of the methodology, please refer to the study by Car et al [19].

Data Sources

This review is a part of a global evidence synthesis initiative on digital health professions education for which a wider search strategy was developed (Multimedia Appendix 1). The following databases were searched from January 1990 to August 2017: Medical Literature Analysis and Retrieval System Online (Ovid), Excerpta Medica dataBASE (Ovid), Central Register of Controlled Trials (CENTRAL; Cochrane Library), PsycINFO (Ovid), Education Resources Information Center (Ovid), Cumulative Index to Nursing and Allied Health Literature (EBSCO), and Web of Science Core Collection (Thomson Reuters). The search was limited to studies reported after 1990, as before this, the use of computers was limited to very basic tasks.

No language or publication restrictions were applied. We searched reference lists of all included studies and relevant systematic reviews. The International Clinical Trials Registry Platform Search Portal and Current Controlled Trials metaRegister of Controlled Trials were also searched to identify unpublished or ongoing trials, as well as meeting abstracts and PhD theses. We contacted study authors of included studies to ask if they were aware of other relevant studies and to provide full reports where these were not identified by the searches. Search results from different sources were combined in a single library and duplicate records were removed.

Study Selection

We included individual or cluster randomized trials (cRCTs) that compared digital education to self and usual or traditional learning or other forms of digital education to train pre or postregistration health professionals to deliver smoking cessation therapy. Health professionals with qualifications listed in the Health Field of Education and Training (091) in the International

XSL•FO
RenderX

Standard Classification of Education [20] were included. We excluded studies of students and/or practitioners of traditional, alternative, and complementary medicine. Digital education interventions could be delivered as the main mode of the education intervention or as a part of a complex, multicomponent intervention (ie, blended education). We accepted any type of digital education such as offline and Web-based, computer-based digital education, Serious Gaming and Gamification, massive open online courses, virtual reality environments, virtual patient simulations, psychomotor skill trainers, and mobile learning or mLearning [21-25]. No restrictions on outcomes were applied.

Moreover, 2 reviewers (MS and SB) independently screened titles and abstracts identified by the searches. Full texts of potentially relevant articles were obtained and independently assessed for inclusion by 2 reviewers (MS and SB). Where data were missing or incomplete, authors were contacted for additional information [26,27]. Any disagreements were resolved through discussions between the 2 reviewers with a third reviewer acting as an arbiter (RB).

Data Extraction and Quality Assessment

A total of 2 reviewers (MS and SB) independently extracted data using a standardized data extraction form, which was piloted and amended on the basis of feedback. Data were extracted on study design, participants' demographics, type of digital education, intervention content, and outcomes. We contacted study authors of the included studies in case of any unclear or missing information. Disagreements between reviewers were resolved by discussion. A third reviewer (RB) acted as an arbiter in cases where disagreements persisted.

Data on the following primary outcomes were extracted:

1. Learners' knowledge postintervention. Knowledge is defined as the learners' factual or conceptual understanding measured using difference in pre and posttest scores.
2. Learners' skill postintervention. Skill is defined as the learners' ability to demonstrate a procedure or technique in an educational setting.
3. Learners' attitudes postintervention toward digital education or toward new clinical knowledge and skill or patients (eg, awareness of moral and ethical responsibilities involved in patient contact). Attitude is defined as the tendency to respond positively or negatively toward the intervention.
4. Learners' satisfaction postintervention with the learning intervention (eg, retention rates, dropout rates, and survey satisfaction scores). This is defined as the level of approval when comparing the perceived performance of digital education with one's expectations.

We also extracted data on the following secondary outcomes:

1. Patient-related outcomes (eg, heaviness of smoking index, number of patients who are stopping smoking).
2. Change in health professionals' practice or behavior.
3. Self-efficacy referred to as self-rated competence of health professionals in delivering smoking cessation therapy.
4. Cost and cost effectiveness of the intervention.

For continuous outcomes, we extracted data to calculate standardized mean difference (SMD) and associated 95% CIs in change from baseline or at follow-up between intervention and control groups. For studies that reported median and range for the various outcomes, we converted this to mean and standard deviation [28]. For dichotomous outcomes, we extracted data to calculate relative risks (RRs) and 95% CIs. Where studies reported more than 1 measure for each outcome, the primary measure as defined by the study authors was extracted.

Risk of Bias Assessment

The methodological quality of included randomized controlled trials was independently assessed by 2 reviewers (MS and SB) using the Cochrane risk of bias tool, which includes the following domains: random sequence generation, allocation concealment, blinding of participants to the intervention, blinding of outcome assessment, attrition, and selective reporting. We also assessed the additional domain of baseline imbalances [17]. The following additional criteria were included for the assessment of cRCTs: recruitment bias that can occur when individuals are recruited to the trial after the clusters have been randomized, loss of clusters, incorrect analysis, and comparability with individually randomized trials [17].

Data Synthesis and Analysis

We grouped studies according to outcomes assessed—skill, knowledge, attitude, satisfaction, practice and behavior change, self-efficacy, patient outcomes, and cost. Within these outcomes, we further grouped studies on the basis of intervention (digital education or blended education) and comparison (usual learning or traditional education, blended education, or other forms of digital education)

Heterogeneity was assessed visually using forest plots and by considering differences in participants, interventions, and outcomes across studies. Due to substantial differences among studies, we used a narrative approach to data synthesis. We were unable to identify a clinically meaningful interpretation of effect size in the literature for digital education interventions. Therefore, in line with other research in the field, we present outcomes using postintervention SMD and interpret the effect size using Cohen *rule of thumb* (ie, with 0.2 representing a small effect, 0.5 a moderate effect, and 0.8 a large effect) [17]. For dichotomous outcomes, we summarized RRs and associated 95% CIs across studies. Subgroup analyses were not feasible because of the small number of studies.

Results

Our search strategy for a series of systematic reviews focusing on different digital health professional education modalities yielded 30,532 unique references. Upon screening of titles and abstracts, we excluded 30,051 ineligible references and retrieved full texts for 22 potentially eligible studies. We excluded 10 studies that did not meet the inclusion criteria: 3 were not randomized trials, 4 did not evaluate digital education intervention, and 3 did not target health professionals. A total of 11 studies (12 reports; 2684 health professionals) were included in the review—8 individually randomized trials and

XSL•FO
RenderX

3 cluster randomized trials. Furthermore, 1 study was reported in 2 separate journal articles [29]. The flow of studies through the systematic review process is shown in Figure 1. Characteristics of the 11 included studies are summarized in Table 1.

All included studies were published in English. In addition, 7 studies focused on postregistration health professionals—3 were restricted to doctors [26,29,34] and 4 included mixed populations of doctors and other health professionals [27,31,35,36]. The remaining 4 studies included preregistration health professionals—2 studies included medical students [33,37] and 2 included pharmacy students [30,32]. Furthermore, 10 studies were conducted in high-income countries, 2 in Australia [26,34], 6 in the United States [29,31-33,35,36], and single studies in the United Kingdom [27] and Switzerland [37]. In addition, 1 study was conducted in Thailand, a middle-income country [30]. A total of 4 studies compared digital education with usual learning, 2 studies compared different digital education interventions [34,35], and 2 studies compared blended education with digital education [36,37]. The 3 cRCTs compared blended education with usual learning [27,29,33].

Several modalities were used to deliver the digital education intervention. Web-based systems were used in 6 studies where participants could access learning materials through a Web gateway [27-31,37]. Of these, 2 studies used computer-based programs, 1 with computerized feedback [34], and the other was an interactive multimedia program [30]. Technology-enabled student response systems or clickers were used to provide instructions in 1 study [32]. Moreover, 5 studies blended digital education components utilizing CD-ROM, computer or Web-based interface with usual learning modalities such as face-to-face interactions, lectures, and seminars [27,29,33,36,37]. Although most interventions focused on improving knowledge about smoking cessation and skill in delivering smoking cessation therapy [26,29,30,33-37], the content of the smoking cessation education varied widely. Table 1 provides a detailed overview of the interventions compared in each study.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. RCT: randomized controlled trial.



XSL•FO
RenderX

Table 1. Characteristics of included studies.

Study, country, design	Participants (N)	Participants details	Intervention	Control	Learning outcomes
Digital education versus usual learning					
Chaikoolvatana 2009, Thailand (RCT[a]) [30]	85	Pharmacy students	Interactive computer program for smoking cessation counseling	Classroom lectures	Knowledge and attitude toward intervention
Gordon 2013, United States of America (RCT) [31]	215	Respiratory therapists, nurses, and nurse practitioners	Web-based smoking cessation education program	Usual traditional learning	Behavior, attitude, and self-rated efficacy in providing smoking cessation therapy
Young 2002, Australia (RCT) [26]	53	Family physicians	Web-based distance learning module for delivering smoking cessation advice	Preventive care guidelines sent via postal mail	Knowledge, skill, readiness to change, and self-rated competence
Galal 2015, United States of America (RCT) [32]	214	Pharmacy students	Use of student response systems (SRS) or "clickers" for instruction in a smoking cessation module	Instruction without student response systems	Learner's attitude toward intervention
Blended education versus usual learning					
Butler 2013, United Kingdom (cRCT[b]) [27]	53	General practitioners, and nurses	Web-based learning program with face-to-face trainings to deliver behavior change counseling in smoking cessation	Usual traditional learning	Patient reported changes in smoking behavior after health professionals training
Hymowitz 2007, United States of America (cRCT) [29]	16	Pediatric residents	Hybrid CD-ROM/website training program and seminars series to deliver smoking cessation therapy	Usual learning with reading material on smoking cessation	Skill, practice, and behavior change
Ockene 2015, United States of America (cRCT) [33]	1503	Medical students	Web-based multi-modal education and face-to-face trainings for smoking cessation counseling	Usual learning with traditional tobacco education in the medical curricula	Smoking cessation counseling skill and self-rated competence
Digital education versus digital education					
Bonevski 1999, Australia (RCT) [34]	19	General practitioners	Computer-based program with performance specific feedback system for screening smoking behavior	Computer program without feedback system for smoking cessation behavior	Screening smoking behavior (smoking status classification)
Stoner 2014, United States of America (RCT) [35]	92	Physicians, nurse practitioner, and physician assistants	Web-based multimedia training program for screening, brief intervention, and referral to treatment	Website with hyperlinks to downloadable reading materials	Knowledge, satisfaction, self-efficacy, and change in clinical practice
Blended education versus digital education					
Brunette 2015, United States of America (RCT) [36]	46	Psychiatrists, advanced nurse practitioners	Videoconference educational outreach and use of printed slides, handouts, and questionnaires for cessation pharmacotherapy	Interactive in-person lecture with slides and handouts	Knowledge and attitude toward intervention
Stolz 2012, Switzerland (RCT) [37]	129	Medical students	Self-directed Web-based module	Lectures with video demonstration on smoking cessation	Knowledge, skill, satisfaction, and self-rated smoking cessation counseling skill

[a]RCT: randomized controlled trial.

[b]cRCT: cluster randomized controlled trial.

XSL•FO
RenderX

Comparison interventions also varied across the studies. A total of 7 studies included a traditional learning control group: face-to-face workshops or lectures in 3 studies [33], reading materials in 1 study [29], and preventive care guidelines sent via postal mail in 1 study [26]. Furthermore, 2 studies did not specify details of the traditional learning intervention [27,31]. In addition, 4 studies evaluated others forms of digital education in the control group: computer-based offline learning in 3 studies [25,26,33] and Web-based learning in 1 study [35].

Methodological Quality of Included Studies

A total of 3 studies were judged as high risk of bias for at least 1 domain; all other studies were rated as unclear risk of bias. Concealment of treatment allocation and blinding of outcome assessors were particularly, poorly reported with only 2 studies judged as low risk of bias for each of these domains. Information on incomplete outcome data and selective outcome reporting was better reported with 2 studies judged as unclear risk of bias for incomplete outcome data and 1 for selective outcome reporting. Furthermore, 2 studies were judged as high risk of bias for incomplete outcome reporting and 1 for selective outcome reporting. No other domains were judged high risk of bias for any of the included trials (Figure 2). For cRCTs, none of the studies were judged as high risk of bias for any domain; all studies were judged as either low or unclear risk of bias for recruitment, loss of clusters, and incorrect analysis.

Primary Outcomes

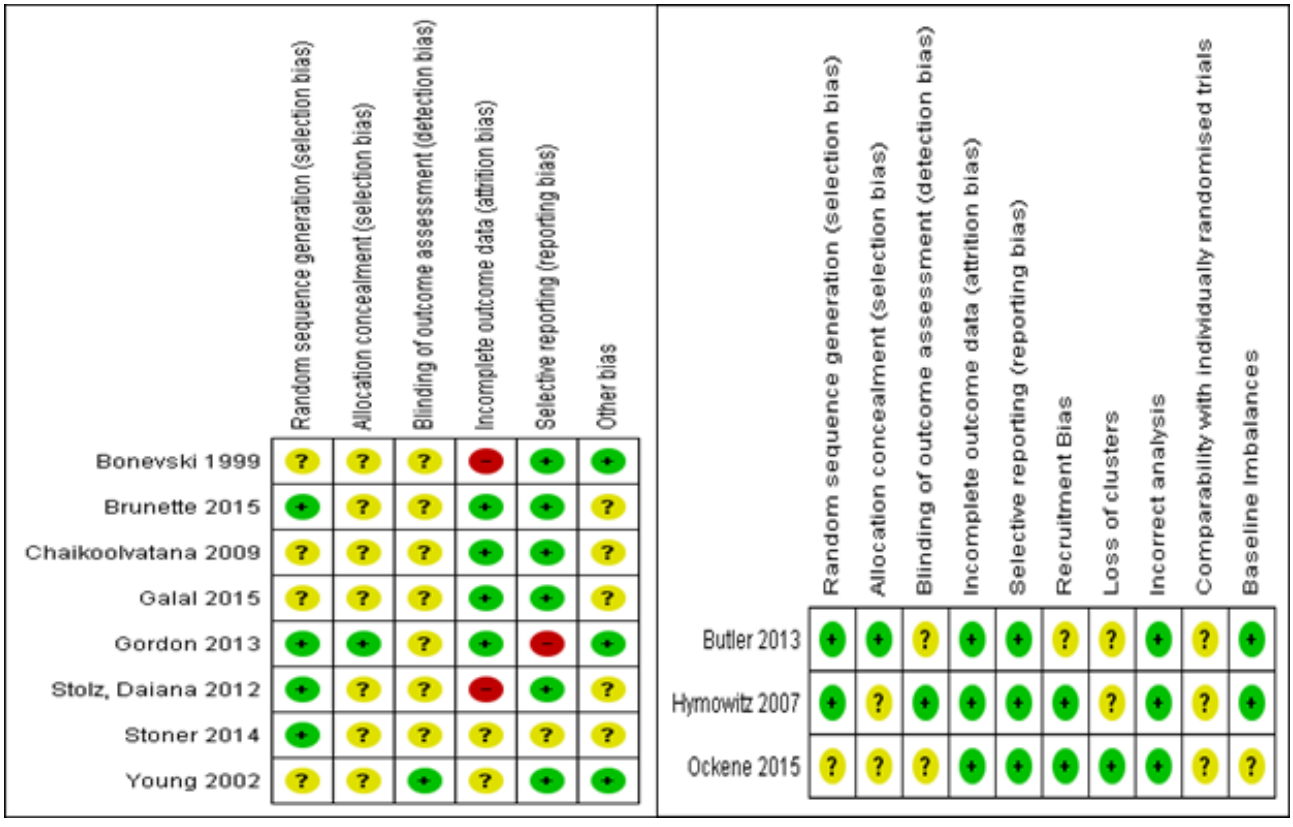
Knowledge

A total of 5 studies (313 participants) assessed postintervention knowledge gain using multiple-choice questions (MCQs) [26,30,35,37] or questionnaires [36] (Multimedia Appendix 2). None of the studies used validated instruments to measure knowledge. Moreover, 3 studies provided quantitative data and found that knowledge gain was similar with digital education compared with usual learning and for blended education compared with digital education (Figure 3). Furthermore, 2 studies did not provide numerical data. One study compared 2 different types of digital education (Web-based multimedia training compared with a website with hyperlinks to downloadable reading material) and found no difference between the interventions [35]. The other study compared blended education with digital education and found no difference in postintervention knowledge between groups [36].

Skill

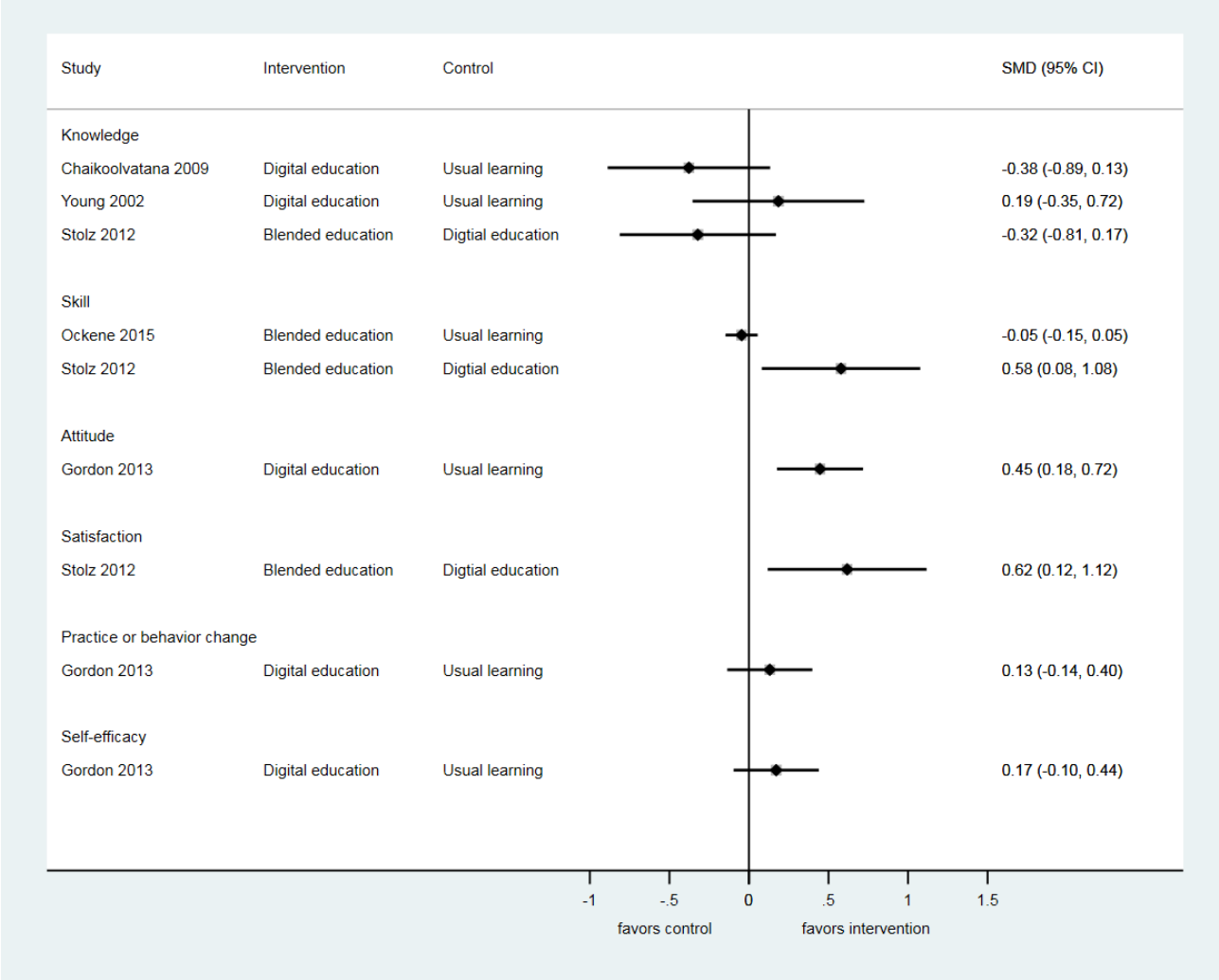
A total of 5 studies (3293 participants) assessed postintervention skill to deliver smoking cessation therapy using objective structured clinical examination scores, [29,33,37] a Likert scale [26], or practitioner checklist [34] (Multimedia Appendix 2). Only 1 study used a validated assessment tool [33]. Furthermore, 1 study found that blended education was associated with greater improvement in skill for delivering smoking cessation advice compared with digital education alone (SMD=0.58, 95% CI 0.08-1.08) [37]. A second study found greater improvement in skill with blended education compared with usual learning (RR=2.04, 95% CI 1.51-2.76) [29]. However, a further study that compared blended education with usual learning found no difference between groups (SMD=−0.05, 95% CI −0.15 to 0.05) [33]. A study that compared 2 types of digital education involving a computer-based program with or without a performance specific feedback system found no difference in skill between groups (RR=1.01, 95% CI 0.96-1.08) [34].

Figure 2. Risk of bias summary: reviewers' judgements about each risk of bias item for each included study (RCTs and cRCTs).



XSL•FO
RenderX

Figure 3. Forest-plot of outcomes showing standardized mean deviations (SMDs) and 95% CI for studies comparing digital education or blended education with usual learning.



Another study reported small improvements in postintervention skill of health professionals compared with usual learning; however, no numerical data were reported for quantitative analysis [26].

Attitude

A total of 4 studies (532 participants) assessed postintervention attitude toward educational interventions [30,32] and new knowledge [31,36] using Likert scales [30-32] and a questionnaire [36] (Multimedia Appendix 2). None of the studies used a validated instrument.

Moreover, 1 study reported improved attitudes following digital education compared with usual learning (SMD=0.45; 95% CI 0.18-0.72) [31]. Furthermore, 3 studies did not provide quantitative data. In addition, 2 of these compared digital education to usual learning and reported positive attitudes toward digital education [30,32]. Another study reported no difference in postintervention attitude between blended education and digital education [36].

Satisfaction

A total of 3 studies (415 participants) assessed postintervention satisfaction with the educational interventions using nonvalidated MCQs [37] and Likert scales [31,35] (Multimedia Appendix 2). Moreover, 1 study reported greater satisfaction after blended education compared with digital education (SMD=0.62, 95% CI 0.12-1.12) [37]. In addition, 2 studies did not provide any quantitative outcome data. Out of these, 1 reported higher postintervention satisfaction in digital education compared with usual learning [31]. Another study reported that postintervention satisfaction was higher with Web-based digital education compared with digital education through websites with hyperlinks [35].

Secondary Outcomes

Practice and Behavior Change

A total of 4 studies (650 participants) assessed postintervention practice and behavior change using nonvalidated surveys [29] and Likert scales [26,31,35,36] (Multimedia Appendix 2). Furthermore, 1 study reported higher postintervention readiness to change practice and behavior to help patients quit smoking with blended education compared with the usual learning (RR=1.58, 95% CI 1.25-2.00) [29]. Another study reported no difference in tobacco cessation–related behaviors between the digital education and usual learning groups (SMD=0.13, 95% CI −0.14 to 0.40) [31]. Furthermore, 1 study reported no difference in postintervention practice and behavior change between 2 modalities of digital education [35]. Similarly, no

XSL•FO
RenderX

difference was observed between digital education compared with usual learning in another study [26,36]. No numerical data were reported for these 2 studies to be included in the quantitative analysis.

Self-Efficacy

A total of 6 studies (1988 participants) assessed postintervention self-efficacy using nonvalidated questionnaires [30,37], Likert scales [26,31,35], and a 6 item checklist [33] (Multimedia Appendix 2). Moreover, 1 study reported no difference in postintervention self-rated smoking cessation counseling skill between blended education and digital education (RR=0.38, 95% CI 0.12-1.13) [37]. Another study reported no difference in self-efficacy scores toward providing tobacco cessation interventions in the digital education and usual learning group (SMD=0.17, 95% CI web 0.10 to 0.44) [31]. In 1 study, higher number of participants receiving blended education reported self-efficacy for performing smoking cessation-related counseling compared with usual learning (P<.05) [33]. In another study, change in self-efficacy between baseline and postintervention was significantly greater with digital education compared with usual learning (P=.03) [26]. Moreover, 1 study reported no difference in postintervention self-efficacy between 2 modalities of digital education [35]. In another study, where self-efficacy was measured only in the intervention group, computer-based program was reported to be effective in improving smoking cessation counseling skill by 73.34% of intervention group participants [30]. No numerical data were reported for these 4 studies to be included in the quantitative analysis [29,31,35,37].

Patient Outcomes and Cost

There was 1 study which assessed postintervention patient outcomes such as smoking index, general health score, quality of life score, and the cost associated with the blended education intervention (Multimedia Appendix 2). The total cost including intervention and health care cost per practice was estimated to be US $2384. However, no quantitative data were reported [27].

Discussion

Overview

Our review included 11 studies that investigated the effectiveness of digital education for training health professionals to deliver smoking cessation therapy. No difference was found between digital education and traditional or usual learning. There was some suggestion that blended education results in greater improvements in satisfaction, skill, and knowledge compared with digital education alone. There was also some evidence for improved attitude following digital education compared with usual learning. Only 1 study reported patient outcomes and the setup cost of digital education. There were insufficient data to investigate what components of the digital education interventions were associated with the greatest improvements in learning outcomes. Studies were poorly reported, heterogeneous, assessed a broad range of different outcomes, and compared different types of interventions on the range of pre and postregistration health professionals. The findings of this review should therefore be interpreted with some caution.

Strengths and Weaknesses

As far as we are aware, this is the first review to address the topic of digital education to train health professionals to deliver smoking cessation therapy. We followed best practice methods for systematic reviews, which attempted to minimize risk of bias and errors in the review process. We conducted a comprehensive sensitive search across a broad range of databases and included additional steps to identify unpublished studies such as searching trials registers, meeting abstracts and PhD theses, screened references of included studies, and contacted authors of abstracts for further information. It was not possible to formally assess the risk of publication bias because of the small number of heterogeneous studies included in our review, but given our extensive search, we consider it unlikely that relevant studies have been missed. Moreover, 2 independent reviewers were involved in all stages of the review process, standardized data extraction forms were used, and we used an accepted tool to assess the risk of bias in the included studies. This identified potential limitations in the included studies, particularly in reporting, which meant that many of the risk of bias domains were judged as unclear for the majority of studies. The included studies evaluated a broad range of interventions and outcomes; therefore, it was not appropriate to calculate summary effect estimates. The included digital education interventions mostly comprised asynchronous Web-based programs aimed at postregistration health professionals. For busy health professionals, digital education is a convenient avenue for fulfilling continuing medical education requirements and promoting knowledge and skill in particular areas in which they may not have previously had training in. Furthermore, 2 studies included an interactive or feedback component in the digital education intervention [34,36]. Evidence suggests that interventions with feedback and interactivity can enhance engagement and consequently the effectiveness of learning [38,39]. However, there were insufficient data available in our review to perform subgroup analyses or a more advanced statistical analysis to evaluate what components of the included interventions contributed to the greatest improvement in outcomes.

There were a number of limitations with the included studies. Reporting of the digital education interventions, especially the description of the intervention, aims and outcomes, pedagogical approach, and use of validated outcome assessment instruments was inconsistent across the studies. There was also a lack of baseline assessment in some studies, meaning that only postintervention data could be used in the analysis, potentially biasing results. Several studies did not report numerical data; therefore, they could not be formally included in our synthesis. We have included results available from these studies, but had additional data been available in the included studies, these would have allowed us to conduct a more informative analysis.

Implications for Practice

When considering the implications of the findings of our review for practice, it is important to consider the implications of the differences in the effectiveness of different types of education.

XSL•FO
RenderX

For forms of digital education that have additional benefits compared with standard education (eg, ability to target larger numbers of people, self-paced learning), showing that these are as effective as standard education is likely to be sufficient to recommend the use of these types of education. In contrast, for types of digital education where technology is supplementing standard face-to-face learning (eg, blended education or use of "clickers"), outcomes would need to be better with digital education than with standard education. For example, the study evaluating the use of clickers to give instructions during classroom lectures reported a positive impact of technology on learners' attitudes, which consequently improved learning outcomes [32]. Our findings suggest that computer-based education, both Web-based and offline, is at least as effective as usual or traditional learning for smoking cessation therapy training in health professionals. This suggests that computer-based learning is an appropriate method to deliver training for health professionals. We found that blended education appears to offer additional benefits compared with digital education or traditional education alone. However, given the additional costs of this type of education, further studies on the cost effectiveness of blended education are needed before this can be recommended for use in practice. There were insufficient data from the studies included in our review to make specific recommendations regarding what types of digital education or components of digital education are likely to be most effective.

Implications for Research

There is a need for further robust studies in a range of settings to determine the true potential of digital education to train health professionals to deliver smoking cessation therapy. More studies evaluating patient outcomes such as postintervention smoking cessation rates, smoking status, and abstinence are needed to assess the effectiveness of digital education for health professionals. Before recommending implementation of digital education programs, information on cost effectiveness, sustainability, as well as the direct and indirect costs such as time to develop as well as implement a smoking cessation module is needed. This will help policy makers to make practical recommendations and allocate resources appropriately. More research is needed to understand the feasibility of integrating digital smoking cessation training methods into the curriculum and continued medical education; research is also needed to understand the short-term and long-term effects in different geographical, socioeconomic, and cultural settings. In addition, sustainability, cost savings, and accreditation of digital smoking cessation therapies in health professional training need to be further researched.

Many digital education interventions in the included studies were based on smoking cessation guidelines. However, validation of the course content and use of underpinning learning theories to develop the pedagogy were lacking in most of the studies. Largely, the studies focused on integrating the new technology into the existing curriculum as opposed to using learning theories to design the digital education intervention for successful delivery of education. An increasing body of evidence reveals that theory-based interventions have greater impact than those that are not based on theory. Appropriate use of learning theory and pedagogy framework along with sound methodology can enable more robust studies to be conducted on digital education with research questions adequately addressed through theoretically informed research design, data collection, and analysis [40-42]. Future studies should therefore focus on developing, delivering, and evaluating digital education with a strong pedagogical foundation.

Digital education can contribute significantly to the World Health Organization (WHO) mission to transform and scale up health professionals' education by filling the medical education divide between low- and high-income countries [43]. Given the greater prevalence of smoking in low-income countries and lesser awareness of harms of smoking compared with high-income countries, being able to train health professionals to deliver smoking cessation advice is of particular importance for low-income countries. However, none of the studies included in our review assessed the use of digital education in such resource-constraint settings. This is an important area for future research.

We did not identify any studies on advanced educational technologies for digital education such as mobile learning, virtual patients, virtual reality environments, or serious gaming, which have the potential to transform education for health professionals. The lack of smoking cessation studies evaluating these educational technologies makes it difficult to make recommendations for integrating such pivotal digital technologies into health professional education. In the future, it is important to have a more detailed reporting of different components of digital education interventions to allow for a more thorough analysis of the most active and effective components.

Conclusions

Digital education appears to be at least as effective as usual or traditional learning in improving health professionals' knowledge and skill for delivering smoking cessation therapy. This suggests that digital education is an appropriate method to deliver training for health professionals on how to deliver smoking cessation therapy. However, limitations in the evidence base mean that these conclusions should be interpreted with some caution. There was insufficient evidence to determine what components of digital education are associated with the greatest improvements in outcomes, although there was some evidence that blended education may be more effective than either digital education or usual learning alone.

Acknowledgments

This review was conducted in collaboration with the Health Workforce Department at WHO. We would also like to thank Mr Carl Gornitzki, Ms GunBrit Knutssön, and Mr Klas Moberg from the university library, Karolinska Institutet, Sweden, for developing the search strategy, and the peer reviewers for their comments. We gratefully acknowledge funding from the Lee

XSL•FO
RenderX

Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, eLearning for health professional's education grant.

Authors' Contributions

LTC conceived the idea for the review. MS and PW wrote the review. RB, SB, and BMK provided comments on the review.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE (Ovid) search strategy.

[PDF File (Adobe PDF File), 22KB - jmir_v21i3e13000_app1.pdf]

Multimedia Appendix 2

Results on included studies.

[PDF File (Adobe PDF File), 33KB - jmir_v21i3e13000_app2.pdf]

References

1. World Health Organization. Tobacco Facts. 2017. Tobacco Free Initiative URL: https://www.who.int/mediacentre/factsheets/ [accessed 2019-02-17] [WebCite Cache ID 76GlaSSoF]
2. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2017 Dec 31;3:CD001292. [doi: 10.1002/14651858.CD001292.pub3] [Medline: 28361496]
3. Stead LF, Carroll AJ, Lancaster T. Group behaviour therapy programmes for smoking cessation. Cochrane Database Syst Rev 2017 Dec 31;3:CD001007. [doi: 10.1002/14651858.CD001007.pub3] [Medline: 28361497]
4. Stead LF, Hartmann-Boyce J, Perera R, Lancaster T. Telephone counselling for smoking cessation. Cochrane Database Syst Rev 2013 Aug 12(8):CD002850. [doi: 10.1002/14651858.CD002850.pub3] [Medline: 23934971]
5. 2008 PHS Guideline Update Panel, Liaisons, and Staff. Treating tobacco use and dependence: 2008 update US Public Health Service Clinical Practice Guideline executive summary. Respir Care 2008 Sep;53(9):1217-1222 [FREE Full text] [Medline: 18807274]
6. Colomar M, Tong VT, Morello P, Farr SL, Lawsin C, Dietz PM, et al. Barriers and promoters of an evidenced-based smoking cessation counseling during prenatal care in Argentina and Uruguay. Matern Child Health J 2015 Jul;19(7):1481-1489 [FREE Full text] [doi: 10.1007/s10995-014-1652-3] [Medline: 25500989]
7. Wells QS, Freiberg MS, Greevy RA, Tyndale RF, Kundu S, Duncan MS, et al. Nicotine metabolism-informed care for smoking cessation: a pilot precision RCT. Nicotine Tob Res 2018 Nov 15;20(12):1489-1496. [doi: 10.1093/ntr/ntx235] [Medline: 29059367]
8. Frehywot S, Vovides Y, Talib Z, Mikhail N, Ross H, Wohltjen H, et al. E-learning in medical education in resource constrained low- and middle-income countries. Hum Resour Health 2013;11:4 [FREE Full text] [doi: 10.1186/1478-4491-11-4] [Medline: 23379467]
9. Makhdoom N, Khoshhal K, Algaidi S, Heissam K, Zolaly M. 'Blended learning' as an effective teaching and learning strategy in clinical medicine: a comparative cross-sectional university-based study. J Taibah Univ Sci 2013 Apr;8(1):12-17 [FREE Full text] [doi: 10.1016/j.jtumed.2013.01.002]
10. Duque G, Demontiero O, Whereat S, Gunawardene P, Leung O, Webster P, et al. Evaluation of a blended learning model in geriatric medicine: a successful learning experience for medical students. Australas J Ageing 2013 Jun;32(2):103-109. [doi: 10.1111/j.1741-6612.2012.00620.x] [Medline: 23773249]
11. Rowe M, Frantz J, Bozalek V. The role of blended learning in the clinical education of healthcare students: a systematic review. Med Teach 2012;34(4):e216-e221. [doi: 10.3109/0142159X.2012.642831] [Medline: 22455712]
12. Schmelz AN, Nixon B, McDaniel A, Hudmon KS, Zillich AJ. Evaluation of an online tobacco cessation course for health professions students. Am J Pharm Educ 2010 Mar 10;74(2):36 [FREE Full text] [Medline: 20414451]
13. Fanshawe TR, Halliwell W, Lindson N, Aveyard P, Livingstone-Banks J, Hartmann-Boyce J. Tobacco cessation interventions for young people. Cochrane Database Syst Rev 2017 Dec 17;11:CD003289. [doi: 10.1002/14651858.CD003289.pub6] [Medline: 29148565]
14. Taylor GM, Dalili MN, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017 Sep 04;9:CD007078. [doi: 10.1002/14651858.CD007078.pub5] [Medline: 28869775]
15. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]

XSL•FO
RenderX

16. Ghorai K, Akter S, Khatun F, Ray P. mHealth for smoking cessation programs: a systematic review. J Pers Med 2014;4(3):412-423 [FREE Full text] [doi: 10.3390/jpm4030412] [Medline: 25563359]
17. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [Medline: 22008217]
18. 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]
19. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, et al. Digital education for health professions: methods for overarching evidence syntheses. J Med Internet Res 2018;21(2):1. [doi: 10.2196/12913]
20. UNESCO Institute for Statistics. 2013. International Standard Classification of Education: Fields of Education Training URL: http://uis.unesco.org/en/topic/international-standard-classification-education-isced [accessed 2019-02-21] [WebCite Cache ID 76MI8cXeK]
21. Gentry S, L'Estrade EB, Gauthier A, Alvarez J, Wortley D, van Rijswijk J. Serious gaming and gamification interventions for health professional education. Cochrane Database Syst Rev 2016(6):10.1002/14651858.CD012209. [doi: 10.1002/14651858.CD012209]
22. Paul P, Toon E, Hadadgar A, Jirwe M, Saxena N, Lim KT. Online- and local area network (LAN)-based eLearning interventions for medical doctors' education. Cochrane Database Syst Rev 2016(3):-. [doi: 10.1002/14651858.CD012108]
23. Semwal M, Wahabi H, Posadzki P, Divakar U, Lim K, Audouard-Marzin Y. Offline and computer-based eLearning interventions for medical doctors' education. Cochrane Database Syst Rev 2016(12):-. [doi: 10.1002/14651858]
24. Kononowicz A, Woodham L, Georg C, Edelbring S, Stathakarou N, Davies D. Virtual patient simulations for health professional education. Cochrane Database Syst Rev (5) 2016:-. [doi: 10.1002/14651858.CD012194]
25. Tudor Car L, Riboli-Sasco E, Marcano BJ, Nikolaou C, Majeed A, Zary N. Mobile learning for delivering health professional education. Cochrane Database Syst Rev (9) 2015. [doi: 10.1002/14651858.CD011861]
26. Young J, Ward J. Can distance learning improve smoking cessation advice in family practice? A randomized trial. J Contin Educ Health Prof 2002;22(2):84-93. [doi: 10.1002/chp.1340220204] [Medline: 12099124]
27. Butler CC, Simpson SA, Hood K, Cohen D, Pickles T, Spanou C, et al. Training practitioners to deliver opportunistic multiple behaviour change counselling in primary care: a cluster randomised trial. Br Med J 2013;346:f1191 [FREE Full text] [Medline: 23512758]
28. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014 Dec 19;14:135 [FREE Full text] [doi: 10.1186/1471-2288-14-135] [Medline: 25524443]
29. Hymowitz N, Schwab JV, Haddock CK, Pyle SA, Schwab LM. The pediatric residency training on tobacco project: four-year resident outcome findings. Prev Med 2007 Dec;45(6):481-490. [doi: 10.1016/j.ypmed.2007.07.024]
30. Chaikoolvatana A, Kitiwongsoonthorn U. Evaluation of a computer interactive multimedia program in smoking cessation counseling (CIMPSCC) for pharmacy students. J Med Assoc Thai 2009;92(11):1516-1523. [Medline: 19938745]
31. Gordon JS, Mahabee-Gittens EM, Andrews JA, Christiansen SM, Byron DJ. A randomized clinical trial of a web-based tobacco cessation education program. Pediatrics 2013 Feb;131(2):e455-e462 [FREE Full text] [doi: 10.1542/peds.2012-0611] [Medline: 23319529]
32. Galal SM, Mayberry JK, Chan E, Hargis J, Halilovic J. Technology vs pedagogy: instructional effectiveness and student perceptions of a student response system. Curr Pharm Teach Learn 2015 Sep;7(5):590-598. [doi: 10.1016/j.cptl.2015.06.004]
33. Ockene J, Hayes R, Churchill L, Crawford S, Jolicoeur D, Murray D. Teaching medical students to help patients quit smoking: outcomes of a 10-school randomized controlled trial. J Gen Intern Med 2016;31(2):172-181. [Medline: 26391030]
34. Bonevski B, Sanson-Fisher RW, Campbell E, Carruthers A, Reid AL, Ireland M. Randomized controlled trial of a computer strategy to increase general practitioner preventive care. Prev Med 1999 Dec;29(6 Pt 1):478-486. [doi: 10.1006/pmed.1999.0567] [Medline: 10600428]
35. Stoner S, Mikko A, Carpenter K. Web-based training for primary care providers on screening, brief intervention, and referral to treatment (SBIRT) for alcohol, tobacco, and other drugs. J Subst Abuse Treat 2014;47(5):362-370. [doi: 10.1016/j.jsat.2014.06.009] [Medline: 25115136]
36. Brunette M, Dzebisashvili N, Xie H, Akerman S, Ferron J, Bartels S. Expanding cessation pharmacotherapy via videoconference educational outreach to prescribers. Nicotine Tob Res 2015;17(8):960-967. [doi: 10.1093/ntr/ntv006] [Medline: 26180220]
37. Stolz D, Langewitz W, Meyer A, Pierer K, Tschudi P, S'ng C. Enhanced didactic methods of smoking cessation training for medical students-a randomized study. Nicotine Tob Res 2012;14(2):224-228. [doi: 10.1093/ntr/ntr186] [Medline: 22090454]
38. Krause U, Stark R, Mandl H. The effects of cooperative learning and feedback on e-learning in statistics. Learn Instr 2009 Apr;19(2):158-170 [FREE Full text] [doi: 10.1016/j.learninstruc.2008.03.003]
39. Hernandez-Torrano D, Ali S, Chan C. First year medical students' learning style preferences and their correlation with performance in different subjects within the medical course. BMC Med Educ 2017;17(1):131. [doi: 10.1186/s12909-017-0965-5] [Medline: 28789690]

XSL•FO
RenderX

40. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
41. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296 [FREE Full text] [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]
42. Cook DA, Ellaway RH. Evaluating technology-enhanced learning: a comprehensive framework. Med Teach 2015;37(10):961-970. [doi: 10.3109/0142159X.2015.1009024] [Medline: 25782599]
43. World Health Organization. 2016. High-Level Commission on Health Employment and Economic Growth URL: https://www.who.int/hrh/com-heeg/en/ [accessed 2019-02-21] [WebCite Cache ID 76MIC0qpN]

Abbreviations

cRCT: cluster randomized trial
MCQ: multiple-choice question
RR: relative risk
SMD: standardized mean difference
WHO: World Health Organization

Edited by A Marusic; submitted 05.12.18; peer-reviewed by I Buljan, S Peters, S Gallagher, S Amantini; comments to author 29.12.18; revised version received 01.02.19; accepted 01.02.19; published 04.03.19.

Please cite as:
Semwal M, Whiting P, Bajpai R, Bajpai S, Kyaw BM, Tudor Car L
Digital Education for Health Professions on Smoking Cessation Management: Systematic Review by the Digital Health Education Collaboration
J Med Internet Res 2019;21(3):e13000
URL: http://www.jmir.org/2019/3/e13000/
doi:10.2196/13000
PMID:30829576

©Monika Semwal, Penny Whiting, Ram Bajpai, Shweta Bajpai, Bhone Myint Kyaw, Lorainne Tudor Car. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Health Professions' Digital Education: Review of Learning Theories in Randomized Controlled Trials by the Digital Health Education Collaboration

Shweta Bajpai[1,2], PhD; Monika Semwal[1,2], PhD; Ram Bajpai[1], PhD; Josip Car[1,3], MD, PhD; Andy Hau Yan Ho[1,4,5], PhD, EdD

[1]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Ageing Research Institute for Society and Education, Nanyang Technological University, Singapore, Singapore

[3]Global eHealth Unit, Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

[4]Psychology Programme, School of Social Sciences, Nanyang Technological University, Singapore, Singapore

[5]Palliative Care Centre for Excellence in Research and Education, Singapore, Singapore

Corresponding Author:
Andy Hau Yan Ho, PhD, EdD
Psychology Programme
School of Social Sciences
Nanyang Technological University
48 Nanyang Avenue
HSS-04-03
Singapore, 639818
Singapore
Phone: 65 63168943
Fax: 65 67955797
Email: andyhyho@ntu.edu.sg

Abstract

Background: Learning theory is an essential component for designing an effective educational curriculum. Reviews of existing literature consistently lack sufficient evidence to support the effectiveness of digital interventions for health professions' education, which may reflect disconnections among learning theories, curriculum design, use of technology, and outcome evaluation.

Objective: The aim of this review was to identify, map, and evaluate the use of learning theories in designing and implementing intervention trials of health professions' digital education, as well as highlight areas for future research on technology-enhanced education via the establishment of a development framework for practice and research.

Methods: We performed a systematic search of Medical Literature Analysis and Retrieval System Online, Excerpta Medica database, Cochrane Central Register of Controlled Trials (Cochrane Library), PsycINFO, Cumulative Index to Nursing and Allied Health Literature, Education Resources Information Center, and Web of Science for randomized controlled trials (RCTs) published between 2007 and 2016.

Results: A total of 874 RCTs on digital health education were identified and categorized into online-offline, mobile digital education, and simulation-based modalities for pre and postregistration health professions' education. Of these, 242 studies were randomly selected for methodological review and thematic analysis. Data were extracted by one author using a standardized form, with a (48/242, 20%) random sample extracted by a second author, in duplicate. One-third (81/242, 33.4%) of the studies reported single or multiple learning theories in design, assessment, conceptualization, or interpretation of outcomes of the digital education interventions. Commonly reported learning theories were problem-based learning (16/81, 20%), social learning theory (11/81, 14%), and cognitive theory of multimedia learning (10/81, 12%). Most of these studies assessed knowledge (118/242, 48.8%), skills (62/242, 25.6%), and performance (59/242, 24.3%) as primary outcomes with nonvalidated assessment tools (151/242, 62.4%). Studies with reported learning theories (χ^2_1=8.2; *P*=.002) and validated instruments (χ^2_1=12.6; *P*=.006) have shown effective acquisition of learning outcomes.

Conclusions: We proposed a Theory-Technology Alignment Framework to safeguard the robustness and integrity of the design and implementation of future digital education programs for the training of health professionals.

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e12912) doi:10.2196/12912

KEYWORDS

learning theory; health professions; digital education; digital health education; digital education interventions

Introduction

Background

Digital education is one of the most popular and rapidly evolving approaches to teaching and learning in health professions' education. It offers a wide range of intervention modalities using information and communication technologies, such as computer-assisted learning, mobile learning, and digital simulation–based learning, which enable individuals to acquire knowledge and skills in a timely and cost-effective manner with greater personal control [1,2]. The Mayo Clinic, a sprawling national health care system in the United States, reported that a sizeable portion of the national expenditure (approximately US $1 billion over 3 to 5 years) goes to digital medical education [3], and other advanced nations are following a similar developmental trend. Despite such huge investments, there is a lack of sufficient evidence to support the effectiveness of digital interventions for health professions' education [4]. Well-conducted randomized controlled trials (RCTs) are the optimal design for evaluating the effectiveness of an intervention; this also applies to digital education interventions in health professions' education. However, in the absence of clear theoretical foundations to measure change in learning outcomes, the methodological integrity of RCTs may be compromised, weakening the process of research, and ultimately weakening the validity of results [5-9].

Implementation of learning theories in the design of digital education could reduce such uncertainties as they provide structured theoretical and practical foundations that help educators develop the curricula, pedagogies, and assessments that are most relevant and conducive to student learning [10,11]. Learning theories also help learners understand their own learning processes, recognize ways to ensure short- and long-term maintenance of learning, and engage in effective practices to achieve intended learning outcomes [12-14]. Furthermore, learning theories inform and inspire pedagogy strategies that serve to ensure good teaching practices for both traditional and digital education, enabling educators to identify and understand the complexity and specificity of knowledge acquisition, in addition to providing insights for effective curriculum design and appropriate measurement of learning outcomes [15]. Although there is a wealth of learning theories available to help guide and evaluate traditional education designs, these theories have not been consistently applied or realized in the development of digital education [16]. Without a robust learning theory foundation or pedagogy framework to guide or evaluate digital health education, its effectiveness for achieving optimal learning outcomes is highly questionable [17,18]. In fact, an increasing body of evidence reveals that theory-based learning intervention has greater impact over nontheory-based interventions [12,19-24].

Objectives

There is a paucity of research on the extent to which digital intervention for health professions' education design integrates educational theory. The lack of understanding of how learners acquire knowledge via different digital modalities makes it difficult to determine the appropriate outcomes to measure when evaluating the effectiveness of such interventions. Therefore, this methodological review aimed to address this important but often neglected area of digital health professions' education. We carried out a critical analysis of digital interventions within health professions' education to determine the extent to which learning theories were explicitly used in the intervention design and evaluation and examine how these theories have been implemented. Finally, on the basis of the resulting findings, we conceptualized a development framework for augmenting learning theories in the design of digital interventions for health professions.

Methods

Defining Intervention Modalities

Digital education includes a variety of technologies such as offline and online computer-based digital education, digital game–based learning (DGBL), massive open online courses, virtual reality (VR), virtual patient simulation (VPS), psychomotor skills trainers, and mobile digital education [25]. For this study, we classified digital education into 3 broad sections on the basis of the nature of the educational content and delivery. We grouped studies that used online modes with those that used offline modes, such as CD-ROM or universal serial bus sticks, as *online-offline* interventions. Studies that used mobile phones, tablets, personal digital assistants, and other handheld devices for delivering educational content were grouped as *mobile digital education* interventions. Finally, studies that utilized simulation in the learning intervention, such as VR, VPS, and DGBL, were categorized as *digital simulation–based education.*

Study Design and Data Sources

Our study is a methodological review that adopts both a quantitative and a qualitative evaluative approach. This review is part of a global evidence synthesis initiative on digital health professions' education [26]. A systematic literature search for digital health education RCTs and quasi-experimental studies was carried out using the following databases from January 1990 to August 2016: Medical Literature Analysis and Retrieval System Online (Ovid), Excerpta Medica database (Elsevier), Cochrane Central Register of Controlled Trials (Cochrane Library, Wiley), PsycINFO (Ovid), Educational Research Information Centre (Ovid), Cumulative Index to Nursing and Allied Health Literature (EBSCO), and Web of Science Core Collection (Thomson Reuters; Multimedia Appendix 1).

XSL•FO
RenderX

Study Selection

A total of 874 intervention studies, published between January 2007 and August 2016, on different areas of digital education, were identified through our database search. We included only pre- and post-registration health professions as listed in the Health Field Education and Training (091) of the International Standard Classification of Education (United Nations Educational, Scientific and Cultural Organization Institute for Statistics, 2013) [27]. Learners of traditional, alternative, and complementary medicine were excluded. A convenience sample of 25.2% (220/874) was drawn randomly from these studies to understand the reporting pattern of learning theories in digital education interventions. We selected only one-fourth of these studies in the methodological review, as our objective was to understand and identify the general trend of reporting, and such a sampling approach was deemed adequate and feasible in previous methodological review studies on education interventions [28]. We also extracted and included the data of 22 randomly selected pilot studies in the review to understand and highlight the reporting style of learning theories. Therefore, a total of 242 (220+22 pilot studies) unique studies were used in this analysis. Microsoft Excel was used to generate random numbers.

Data Extraction

Data were extracted for each included study. We developed the data extraction form through pilot testing and revised it further according to the feedback from coauthors. We extracted information including random number, mode of digital learning, first author, year of publication, title of the study, name of the journal, sample size, study population, setting, country where research was originally conducted, primary outcomes, measurement instrument, validation of measurement instrument, and the theory mentioned in the included studies. Data were extracted by the first author (SB) and verified by the second author (MS), and discrepancies were resolved through discussion, with adjudication by a senior author (AH) when necessary.

Identification and Analysis of Theories

This study investigated the use of learning theory reported in the interventions of 3 modalities of digital health professions' education. To be judged as having a theory used, 3 criteria had to be met. First, any study that explicitly named a learning theory used in the design of intervention or learning evaluation was considered as theory used. Second, any study that described the use of pedagogy and the theoretical framework relevant to learning theory in intervention design and learning evaluation was considered as theory used. Third, any study that did not explicitly mention a learning theory but had clearly employed a learning theory in intervention design and learning evaluations was considered as theory used. In instances where a study merely mentioned that its intervention or evaluation design was based on *pedagogical or learning principles* but did not name the theory or describe its relevant features, was not considered as having used a theory and was excluded from this review. Decisions on theory used were based on consensus of the working group in case of uncertainty (SB, MS).

Statistical Analysis

Data were directly entered into Microsoft Excel and subsequently cleaned for invalid entries. The analysis was mostly descriptive; we summarized data as frequency and percentage for categorical items and median and interquartile range (IQR) for continuous items. We analyzed predefined study characteristics of all included digital learning interventions as previously described in data extraction section. We quantified associations using the Chi-square test and 2-sided $P<.05$ was considered as statistical significance. All analyses were performed using Stata software (version 14.0, StataCorp).

Results

General Characteristics of Included Digital Education Intervention Studies

We evaluated 242 studies, published between 2007 and 2016, from 3 modalities of digital health professions' education interventions: online-offline–based digital education (154/242, 63.6%), mobile digital education (21/242, 8.7%), and digital simulation–based education (67/242, 27.7%; see Table 1). Most of the studies were published between 2012 and 2016 (155/242, 64.0%) and conducted in high-income countries including the United States (102/242, 42.1%), United Kingdom (25/242, 10.3%), Germany (14/242, 5.7%), Canada (13/242, 5.3%), and Australia (11/242, 4.5%). The study population in the majority of the studies was preregistration health professionals (148/242, 61.1%) and the median (IQR) of the study size was 72 (43-120). Only one-third of the studies (81/242, 33.4%) mentioned any type of learning theory applied in the intervention design. More than half of the studies (151/242, 62.4%) used nonvalidated measurement instruments to assess primary outcomes. Most studies assessed knowledge (118/242, 48.7%), skills (62/242, 25.6%), and performance (59/242, 24.4%) as primary outcomes (Figure 1). The preferred choice of measurement structures was self-reported multiple-choice questions (MCQs; 69/242, 28.5%), questionnaire (66/242, 27.3%), scales (48/242, 19.8%), and a combination of the tools (18/242, 7.4%). However, in some studies (17/242, 7.0%) the measurement assessment tools were not clearly specified (Multimedia Appendix 2).

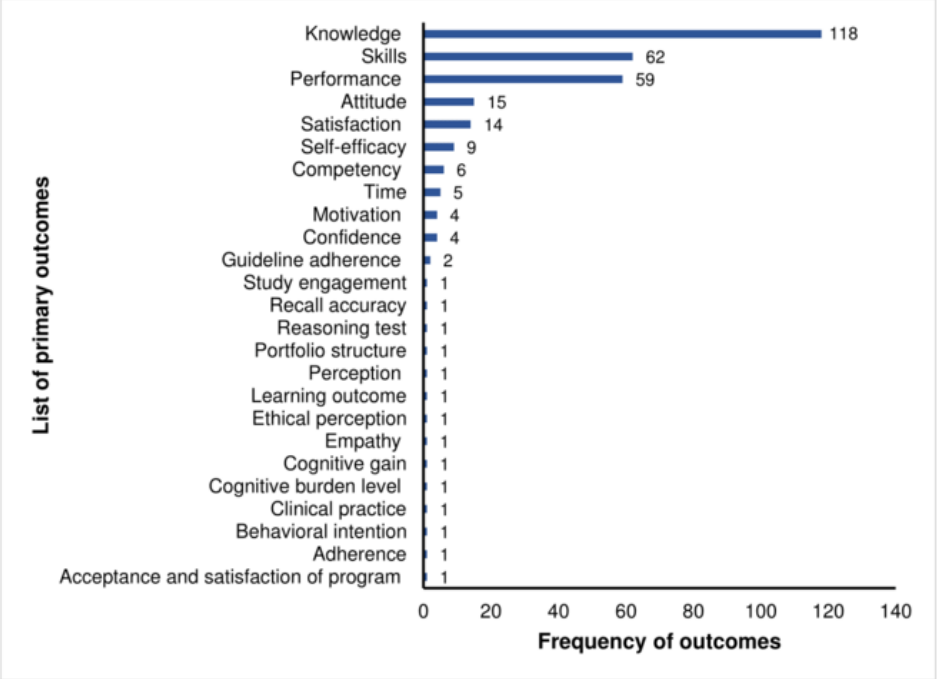
XSL•FO
RenderX

Table 1. Characteristics of included digital health professions' education intervention studies.

Study characteristics	Type of digital domain			Total (N=242)
	Online-offline–based education (N=154)	Mobile digital education (N=21)	Digital simulation–based education (N=67)	
Year of publication, n (%)				
2007	4 (2.6)	1 (4.7)	3 (4.4)	8 (3.3)
2008	10 (6.5)	1 (4.7)	5 (7.4)	16 (6.6)
2009	10 (6.5)	1 (4.7)	4 (5.9)	15 (6.2)
2010	16 (10.4)	2 (9.5)	6 (8.9)	24 (9.9)
2011	17 (11)	1 (4.7)	6 (8.9)	24 (9.9)
2012	25 (16.2)	0 (0)	12 (17.9)	37 (15.2)
2013	16 (10.3)	4 (19)	7 (10.4)	27 (11.1)
2014	21 (13.6)	3 (14.2)	9 (13.4)	33 (13.6)
2015	22 (14.2)	4 (19)	10 (14.9)	36 (14.8)
2016	13 (8.4)	4 (19)	5 (7.4)	22 (9)
Type of population, n (%)				
Undergraduate	86 (55.8)	14 (66.6)	5 (7.4)	148 (61.1)
Postgraduate	45 (29.2)	5 (23.8)	14 (20.9)	64 (26.4)
Mixed population	23 (14.9)	2 (9.5)	48 (71.6)	30 (12.4)
Setting, n (%)				
Hospital	60 (38.9)	10 (47.6)	16 (23.8)	86 (35.5)
University	94 (61)	11 (52.3)	51 (76.1)	156 (64.4)
Study size, median (interquartile range)	84 (47-138)	63 (42-72)	52 (30-93)	72 (43-120)
Top 5 countries of publication, n (%)				
United States	69 (44.8)	8 (38)	25 (37.3)	102 (42.2)
United Kingdom	20 (12.9)	0 (0)	5 (7.4)	25 (10.3)
Germany	10 (6.5)	0 (0)	4 (5.9)	14 (5.7)
Canada	10 (6.5)	1 (4.7)	2 (2.9)	13 (5.3)
Australia	6 (3.9)	1 (4.7)	4 (5.9)	11 (4.5)
Statistical significance of primary outcomes, n (%)				
No	74 (48)	8 (38)	25 (37.3)	107 (44.2)
Yes	71 (46.1)	11 (52.3)	37 (55.2)	119 (49.1)
Mixed	9 (5.8)	2 (9.5)	5 (7.4)	16 (6.6)
Reported validity of the instrument used, n (%)				
No	86 (55.8)	16 (76.1)	49 (73.1)	151 (62.4)
Yes	68 (44.1)	5 (23.8)	18 (26.8)	91 (37.6)
Reported learning theory for the design of intervention, n (%)				
No	99 (64.2)	12 (57.1)	50 (74.6)	161 (66.5)
Yes	55 (35.7)	9 (42.8)	17 (25.3)	81 (33.4)

XSL•FO
RenderX

Figure 1. Frequency distribution of reported primary outcomes in digital health professions' education intervention studies.



Categorization of Reported Learning Theories in Digital Education Interventions

In 33.5% (81/242) studies that reported learning theories, a total of 42 theories were applied in the design of digital health professions' education interventions (Table 2). The most commonly applied learning theories were problem-based learning (16/81, 20%), social learning theory (11/81, 14%), Mayer's cognitive theory of multimedia learning (10/81, 12%), and adult learning theory (8/81, 10%). Furthermore, 7 theories including cognitive load theory (7/81, 9%), Kirkpatrick's framework (6/81, 7%), cognitive theory of learning (6/81, 7%), constructive theory of learning (5/81, 6%), Bloom's taxonomy (5/81, 6%), collaborative learning (5/81, 6%), and social cognitive learning (4/81, 5%) were employed 5 to 7 times across different intervention studies. The remaining 32 theories were sparsely reported. In terms of theory application and integration, 48 out of 81 studies (59%) reported only 1 theory in the design of the digital education programs, 22 out of 81 studies (27%) reported integrating 2 theories, and 8 out of 81 studies (10%) used 3 to 6 theories to develop the interventions.

Given the vast number of theories reported, we further organized each theory into a set of thematic categories on the basis of its nature and characteristics. Among the 42 identified theories, only 13 referred to a specific learning theory. Of these, 7 theories were categorized under cognitivism, which focuses on the inner mental activity and information processing of learners. Furthermore, 3 theories were organized into the constructivism category. The remaining 3 theories were standalone idiosyncratic models of learning. A total of 9 theories were categorized under design-based learning, which comprises a combination of multiple learning theories for explaining learning processes and pedagogy practices. A total of 9 theories were categorized under behavior-change theories that explain the processes of health-related behavioral and attitudinal transformation. Furthermore, 3 theories were organized into the social sciences category, 3 into the decision-making and therapeutic framework category, 3 into the learning style category, and 2 theories into the motivational theory category. (Multimedia Appendix 3).

XSL•FO
RenderX

Table 2. List and frequency of reported learning theories (n=42) by modality in digital health professions' education intervention studies (total reported studies=81).

Name of theory	Online-offline–based education (n=55), n (%)	Mobile digital education (n=9), n (%)	Digital simulation–based education (n=17), n (%)	Total (N=81), n (%)
Problem-based learning	11 (20)	1 (11.1)	5 (29.4)	17 (21)
Social learning theory	9 (16.3)	0 (0)	2 (11.7)	11 (13.5)
Mayer's cognitive theory of multimedia learning	4 (7.2)	3 (33.3)	3 (17.6)	10 (12.3)
Adult learning theory	7 (12.7)	1 (11.1)	0 (0)	8 (9.8)
Cognitive load	2 (3.6)	1 (11.1)	4 (23.5)	7 (8.6)
Kirkpatrick's framework	6 (10.9)	0 (0)	0 (0)	6 (7.4)
Cognitive theory	3 (5.4)	2 (22.2)	1 (5.8)	6 (7.4)
Constructive theory	4 (7.2)	1 (11.1)	0 (0)	5 (6.1)
Bloom's taxonomy	3 (5.4)	1 (11.1)	1 (5.8)	5 (6.1)
Collaborative learning	5 (9)	0 (0)	0 (0)	5 (6.1)
Social cognitive theory	4 (7.2)	0 (0)	0 (0)	4 (4.9)
Theory of self-efficacy	4 (7.2)	0 (0)	0 (0)	4 (4.9)
Information processing	1 (1.8)	1 (11.1)	1 (5.8)	3 (3.7)
Health belief model	3 (5.4)	0 (0)	0 (0)	3 (3.7)
Situated learning	2 (3.6)	0 (0)	1 (5.8)	3 (3.7)
Dual coding theory	1 (1.8)	2 (22.2)	0 (0)	3 (3.7)
Kolb's experiential learning	0 (0)	0 (0)	3 (17.6)	3 (3.7)
Innovation diffusion theory	2 (3.6)	0 (0)	0 (0)	2 (2.4)
Cooperative learning	1 (1.8)	0 (0)	1 (5.8)	2 (2.4)
Social constructivism	0 (0)	1 (11.1)	1 (5.8)	2 (2.4)
Theory of reasoned action	2 (3.6)	0 (0)	0 (0)	2 (2.4)
Cognitive dissonance theory	2 (3.6)	0 (0)	0 (0)	2 (2.4)
Cognitive apprenticeship model	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Theory of behavior change	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Cognitive flexibility theory	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Cognitive behavioral therapy theory	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Enquiry-based learning	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Practice-based learning	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Theory of reflective practice	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Bowen's teaching strategy	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Banning's theoretical framework	0 (0)	1 (11.1)	0 (0)	1 (1.2)
Positive psychological theoretical framework	0 (0)	0 (0)	1 (5.8)	1 (1.2)
System approach model	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Persuasive communication model	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Social support theory	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Social marketing theory	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Theory of self-determination	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Wittrock's generative learning theory	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Elaboration theory	1 (1.8)	0 (0)	0 (0)	1 (1.2)

XSL•FO
RenderX

Name of theory	Online-offline–based education (n=55), n (%)	Mobile digital education (n=9), n (%)	Digital simulation–based education (n=17), n (%)	Total (N=81), n (%)
Taxonomy of significant learning	0 (0)	0 (0)	1 (5.8)	1 (1.2)
ARCS model of motivational design[a]	1 (1.8)	0 (0)	0 (0)	1 (1.2)
Connectivism	1 (1.8)	0 (0)	0 (0)	1 (1.2)

[a]This model is based on four steps for promoting and sustaining motivation in the learning process: attention, relevance, confidence, and satisfaction.

Comparison of Learning Theory Used With Study Characteristics and Thematic Analysis

There were no significant differences in the publication years, types of digital domain, research settings, sample population, and sample size among studies that did or did not report learning theory used. However, studies that reported the validated measurement instrument (χ^2_1=12.6; *P*=.006) and significant primary outcomes (χ^2_1=8.2; *P*=.002) were statistically associated with the reporting of learning theory (Multimedia Appendix 4).

Furthermore, thematic analysis revealed that out of the 81 studies that reported learning theories, 70% (57/81) of the studies were judged to have a relatively clear use of theory in interventions (Multimedia Appendix 5). Among these 57 studies, 27 studies clearly provided the description of use of theory to develop an instructional design, 10 studies used theory-based assessment tools to measure learning outcomes, 4 studies used a theoretical framework to evaluate learning system, and 4 studies used theory to justify their findings. Another 4 studies used theory only for developing study hypothesis and objectives, but the effects were not observed. Furthermore, 7 other studies simply stated that a theory was used in the conceptualization of the intervention without a clear description, whereas 1 study reported partial use of theory to develop the intervention. On the other hand, 7 out of 81 studies (9%) reported learning theory but did not provide any explicit explanation about theory usability in methods and only mentioned a theory name in the introduction and discussion sections. Moreover, 17 out of 81 studies (21%) only named a learning theory vaguely or implicitly; therefore, inference was drawn on the basis of the available description on the theory used. Finally, from conceptualization to conclusion of results, no study was completely based on a learning theory framework.

A closer examination of different digital modalities revealed that among offline-online intervention studies, (n=55), 37 out of 55 studies (67%) clearly elaborated on the purpose of the theory used, 29 out of 55 studies (53%) used validated measurement instruments to assess learning outcomes, and 17 out of 55 studies (31%) reported effectiveness of digital interventions compared with the conventional method of learning. Among the mobile digital intervention studies, all included studies (n=9) discussed the purpose of the theory used, 3 studies used a validated instrument to assess the outcomes, and 2 studies reported significant results. Among the digital simulation–based intervention studies (n=17), 11 studies out of 17 (65%) were judged to have employed theory in their interventions, 9 studies assessed outcomes implying valid tools, and 5 studies reported statistically significant results (Multimedia Appendix 5).

Discussion

Principal Findings

This is the first ever study that comprehensively reviewed the application of learning theory within the design and evaluation of digital health professions' education interventions, inclusive of online-offline–based education, mobile digital education, and digital simulation–based education modalities. Our analyses highlight 4 serious concerns that may hamper the developmental integrity of this fast growing industry, including (1) design and implementation of digital health professions' education interventions without integration of appropriate pedagogy frameworks (161/242, 66.5%), (2) poor selection of learning theories for developing or supporting education interventions, (3) not selecting appropriate theory(s) for different modalities of digital health education such as computer-assisted learning, Web-based learning, mobile learning, and others, and (4) inappropriate usage of learning theory(s) and a nonvalidated assessment instrument that results in the mismatch of learning outcomes. The issues identified with the lack of learning theories and inappropriate application of learning theories to the development of digital education for health professions impede high-quality research into the efficacy of digital learning in this area.

Specifically, our findings reveal that the interventions for health professions' education did not utilize a learning theory or pedagogical framework in curriculum design, program implementation, or learning evaluation. In fact, only one-third of the intervention studies included in our review were informed by learning theories; nevertheless, the purpose of theory utilization was unclear in many instances. This phenomenon may be explained by the intentionality of interventions. All published articles focused primarily on the use of technology and its effectiveness on learning rather than carefully applying learning theory and pedagogy in the design of digital curriculums [18]. Consequently, most studies are comparative in nature, with an emphasis to contrast technology-assisted education with traditional education methods in producing learning outcomes, ignoring the appropriate use of learning theories for education design [29,30] and failing to adequately investigate the mechanisms that make digital education effective.

We also observed a lack of clarity and explanations for the selection of learning theories and pedagogy that informed digital education design, implementation, and evaluation [24]. Poor selection of theory could be a result of nonavailability of reporting guidelines [16]. Given the vast number of learning theories available in existing literature with overlapping characteristics, it may prove difficult to choose a specific theory for developing an effective curriculum while keeping a specific

XSL•FO
RenderX

set of learners in mind [31,32]. It is also important to note that no learning theory has been established precisely for illuminating the inner working of digital education, despite rapid developments in the field, and this could result in overt simplification of the use of new technology to shape digital education without truly realizing its impact on learning and pedagogy [18,20,33-35]. Unmistakably, a comprehensive understanding of technological prowess combined with relevant and well-versed learning principles is urgently needed to improve the quality of education delivered to digital learners [36,37].

Another important finding derived from our review is the improper selection and reporting of measurement tools and the use of a nonvalidated instrument. Although certain psychological or attitudinal constructs of learning outcomes, such as self-efficacy, often demand the use of specific self-assessment instruments, validation may not always be feasible or practical. In addition, nonvalidated self-assessment tools have also been widely criticized for their lack of accuracy [38]. It is thus recommended that if validation has not been carried out for an established scale at the time of education intervention, especially that related to psychological and attitudinal constructs, then confirmatory factor analysis needs to be conducted upon data collection to ensure validity and reliability of instrument used [39].

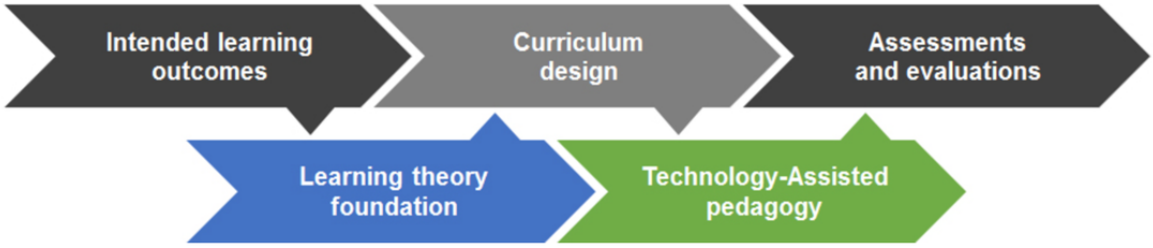
Moreover, most studies in our review assessed knowledge followed by skills and performance as primary outcomes. We observed a high degree of incongruities (71/81, 88%) among the choice of learning outcomes, the underlying learning theory, and the use of measurement instruments for assessing such learning outcomes (measurement instruments; Multimedia Appendix 5). In most of the studies, the measurement tools do not adequately fit the learning outcomes of the reported learning theory. For instance, it is difficult, if not impossible, to assume that MCQs in nonstandardized tests (151/242, 62.4%), which are strongly associated with assessing lower cognitive processing such as fact recall [40], can adequately assess problem-solving skills, knowledge application, motivation, teamwork, and creativity, all of which are primary learning outcomes of the most frequently reported learning theories in our review. Specifically, problem-based learning, the top reported learning theory in our review, aims to foster teamwork and creative real-world problem-solving abilities [41]; on the other hand, social learning theory, the second most reported theory, aims to produce behavioral, motivational, and attitudinal changes [42]; finally, Mayer's cognitive theory of multimedia learning, the third most reported theory, aims to enhance information processing and creative thinking [43]. As learning outcomes are used and interpreted in various ways, it is important to choose an appropriate, valid, and theory-based instrument that essentially assesses what it intended to assess [44]. Poor evaluation processes in digital health professions' education can compromise curriculum design, mislead learning outcomes [24], and lead to poor clinical practices, which ultimately put patient care at stake.

Notably, our review showed significant association among the application of learning theory, validity of the instrument used, and statistical significance of primary outcomes. Findings revealed that studies that did not apply learning theory (110/161, 68.3%) were significantly less likely to use validated measurement instrument. Similarly, primary outcomes were statistically significant in a majority of the studies (52/81, 64%) that applied learning theory and validated assessment tools. In addition, more than half of the studies reporting a nonvalidated assessment tool may indicate a poor understanding of validity theory (ie, what constitutes a valid outcome measure) and poor contextualization and application of validity and reliability in medical education [45]. Therefore, the design of the intervention from hypothesis to measuring outcomes could be improved when the research begins with an appropriate theory or pedagogical framework.

Limitations of the Study

Despite the important findings reported, this review comprises some limitations. First, this is not a systematic review of theory-used in digital learning, which might increase the likelihood of missing some important study(s) pertaining to the topic. Second, we reviewed only one-fourth of published studies from 3 modalities; however, digital learning includes various modalities, but feasibility restriction does not permit us to include all published studies from different modalities of digital education. Nonetheless, future research may aim to focus on these issues, including proper reporting of theory-used and measurement tools for assessment. Given the multimodal nature of digital education, it is recommended that unless we have the solid theoretical guidelines for development, implementation, and evaluation, we are unlikely to achieve the desired learning goals [16].

Figure 2. Theory-Technology Alignment Framework for health professions' digital education.



XSL•FO
RenderX

Theory-Technology Alignment Framework

The appropriate application of learning theories in digital education is essential to ensure curriculum integrity and critical to successful learning outcomes [46,47]. Therefore, we outlined a Theory-Technology Alignment Framework (TTAF) to inform the development of digital education for health professionals (Figure 2). In conventional practices of designing teaching, learning, and assessment activities [48], educators begin with a set of *intended learning outcomes* that they would like students to acquire upon the completion of education and training. On the basis of these outcomes, they proceed to *curriculum design* by incorporating various teaching and learning activities that they believe would lead to the intended learning outcomes, and thereafter the *assessments and evaluations* procedures are established to evaluate whether students have attained the intended outcomes, as well as assess standards of performance. We argue that this widely adopted practice known as constructive alignment [49] is too simplistic in the context of digital education design and neglects the intricate connections and interplays among learning processes, technology, and pedagogy practices. We propose that to achieve a holistic learning experience via digital education, intended learning outcomes and curriculum design must be informed and aligned by an appropriate *learning theory foundation*, one that includes a collection of the most relevant learning theories to ensure the effective choice and application of teaching and learning activities. This would empower educators to better conceptualize and create a conducive pedagogical framework and effective learning environment to help students achieve success. Once a fitting teaching framework and learning environment are clearly delineated and established, they can then be augmented by using appropriate *technology-assisted pedagogy*, which effectively aligns with and supports learning theory foundation and curriculum design for developing theory-driven assessments and evaluations for adequately measuring students' performance in accordance to the original intended learning outcomes.

Putting the TTAF in practice, a digital education intervention for medical trainees may aspire toward the *intended learning outcomes* of (1) evidence-based clinical diagnosis with (2) team-based problem solving. To achieve these outcomes, one can apply a *learning theory foundation* that incorporates team-based learning theory with the Mayer's cognitive theory of multimedia learning to inform *curriculum design*. This may comprise pedagogical instructions that utilize visual and auditory information processing for making proper patient diagnosis in a team-based environment. Furthermore, a set of theory-informed teaching and learning activities, matching *technology-assisted pedagogy*, such as the use of digital patient records including medical charts, x-rays, and computed tomography scans, coupled with voice recordings of patient intake assessments and patient-physician communications, can be used to facilitate the first intended learning outcome. Moreover, an online discussion forum with a built-in monitoring and feedback mechanism for course instructors can be created for trainee groups to deliberate on the digital patient medical information they have and to share their analysis of problem, to come to a joint diagnosis; this will serve to facilitate the second intended learning outcome. With this alignment and integration of curriculum design and technology-assisted pedagogy, an appropriate *assessments and evaluations* protocol can be developed to ensure that all intended learning outcomes are properly assessed. To assess the first outcome, evaluation on individual trainee's understanding and application of digital patient medical information in clinical diagnosis may be conducted through short-answer questions and higher order MCQs that test the cognitive processes of information analysis and knowledge application. To assess the second outcome, continuous monitoring of team-based discussion, as well as evaluation of written reflections on group processes, and group-based problem-solving capacity via stimulated online clinical meetings may be conducted.

In short, we recommend that instead of placing technology at the center of digital education, one must begin at the fundamental roots of learning theory and pedagogy. Therefore, the TTAF could be useful for designing an effective digital education intervention.

Conclusions

Our study has opened the doors for future research by highlighting many problems in digital health professions' education research and its different modalities. This multifaceted problem can be tackled through effective utilization of appropriate learning theories to foster stronger integration among intended outcomes, curriculum design, pedagogy activities, and evaluations with pertinent computer-assisted technologies. Imperatively, high-quality digital research using a clear theoretical framework, well-defined outcomes, and standardized assessment tools with adequate validity evidences is urgently warranted with proper reporting in methodology. Our review serves as an important guideline for researchers, educators, policy makers, and program designers to develop an effective intervention for the training of health professionals, and if applied proficiently, it would assist in advancing the field of digital research in medical education by addressing the various methodological shortcomings that exist in current interventional studies.

Acknowledgments

This study was funded by the Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, eLearning for health professions' education grant, and the Ageing Research Institute for Society and Education, Nanyang Technological University, Singapore.

The authors wish to acknowledge Dr Krister Sven Ake Jarbrink for his input in reviewing the early draft of the manuscript. The authors would also like to thank Mr Carl Gornitzki, Ms GunBrit Knutssön, and Mr Klas Moberg from the university library,

XSL•FO
RenderX

Karolinska Institute, Sweden, for developing the search strategy. The authors would also like to thank anonymous reviewers for their invaluable comments and feedback to improve the manuscript.

Authors' Contributions

SB conducted the study selection, screened, extracted and synthesized data, and wrote and revised the drafts. MS extracted the data and revised the review. RB analyzed the data and provided statistical and methodological guidance and critical input into the study. AHYH critically revised the manuscript, interpreted the results for the work, and was the content expert as a psychologist. JC conceptualized the study, obtained the funding, provided critical input at all stages of the study, and approved the final version of the paper. All authors read and approved the final manuscript before submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE (Ovid) search strategy.

[DOCX File, 17KB - jmir_v21i3e12912_app1.docx]

Multimedia Appendix 2

Frequency distribution of reported measurement instruments in digital medical education intervention studies.

[DOCX File, 31KB - jmir_v21i3e12912_app2.docx]

Multimedia Appendix 3

Categorization and theoretical description of reported learning theories and related theories in digital medical education intervention studies (n=42).

[DOCX File, 22KB - jmir_v21i3e12912_app3.docx]

Multimedia Appendix 4

Comparison of study characteristics with the reporting of learning theory in digital medical education intervention studies.

[DOCX File, 16KB - jmir_v21i3e12912_app4.docx]

Multimedia Appendix 5

Thematic analysis of theory used in digital health professions' education intervention studies (n=81).

[DOCX File, 33KB - jmir_v21i3e12912_app5.docx]

References

1. Guze PA. Using technology to meet the challenges of medical education. Trans Am Clin Climatol Assoc 2015;126:260-270 [FREE Full text] [Medline: 26330687]
2. Sangra A, Vlachopoulos D, Cabrera N. Building an inclusive definition of e-learning: an approach to the conceptual framework. Int Rev Res Open Dist Learn 2012;13(2):145-159 [FREE Full text]
3. Lucas P. Digital Commerce 360. 2017. Mayo ponies up $1 billion to expand digital healthcare URL: https://www.digitalcommerce360.com/2017/02/27/mayo-ponies-1-billion-expand-digital-healthcare/ [accessed 2018-11-24] [WebCite Cache ID 74AEV5zWa]
4. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. J Am Med Assoc 2008 Sep 10;300(10):1181-1196. [doi: 10.1001/jama.300.10.1181] [Medline: 18780847]
5. Kabisch M, Ruckes C, Seibert-Grafe M, Blettner M. Randomized controlled trials: part 17 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2011 Sep;108(39):663-668 [FREE Full text] [doi: 10.3238/arztebl.2011.0663] [Medline: 22013494]
6. Park JH, Son JY, Yoo NJ, Kim S. A review on usage and effectiveness of e-Learning in medical education. Korean J Med Educ 2010 Jun;22(2):91-100 [FREE Full text] [doi: 10.3946/kjme.2010.22.2.91] [Medline: 25813808]
7. Chumley-Jones HS, Dobbie A, Alford CL. Web-based learning: sound educational method or hype? A review of the evaluation literature. Acad Med 2002 Oct;77(10 Suppl):S86-S93. [doi: 10.1097/00001888-200210001-00028] [Medline: 12377715]

XSL•FO
RenderX

8. Wutoh R, Boren SA, Balas EA. eLearning: a review of internet-based continuing medical education. J Contin Educ Health Prof 2004;24(1):20-30. [doi: 10.1002/chp.1340240105] [Medline: 15069909]
9. Liu JL, Wyatt JC. The case for randomized controlled trials to assess the impact of clinical information systems. J Am Med Inform Assoc 2011;18(2):173-180 [FREE Full text] [doi: 10.1136/jamia.2010.010306] [Medline: 21270132]
10. Orchard J, Winch C. What training do teachers need? Why theory is necessary to good teaching. Impact 2015;2015(22):16-18 [FREE Full text] [doi: 10.1111/2048-416X.2015.12002.x]
11. Sandars J, Patel RS, Goh PS, Kokatailo PK, Lafferty N. The importance of educational theories for facilitating learning when using technology in medical education. Med Teach 2015;37(11):1039-1042. [doi: 10.3109/0142159X.2015.1019438] [Medline: 25776228]
12. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
13. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296 [FREE Full text] [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]
14. Michie S. Designing and implementing behaviour change interventions to improve population health. J Health Serv Res Policy 2008 Oct;13(Suppl 3):64-69. [doi: 10.1258/jhsrp.2008.008014] [Medline: 18806194]
15. Cook DA. The research we still are not doing: an agenda for the study of computer-based learning. Acad Med 2005 Jun;80(6):541-548. [doi: 10.1097/00001888-200506000-00005] [Medline: 15917356]
16. Alonso F, Lopez G, Manrique D, Vines JM. An instructional model for web-based e-learning education with a blended learning process approach. Br J Educ Technol 2005 Mar;36(2):217-235. [doi: 10.1111/j.1467-8535.2005.00454.x]
17. Aparicio M, Bacao F, Oliveira T. An e-Learning theoretical framework. J Educ Technol Society 2016;19(1):292-307 [FREE Full text]
18. Cook DA, Bordage G, Schmidt HG. Description, justification and clarification: a framework for classifying the purposes of research in medical education. Med Educ 2008 Feb;42(2):128-133. [doi: 10.1111/j.1365-2923.2007.02974.x] [Medline: 18194162]
19. Davies P, Walker AE, Grimshaw JM. A systematic review of the use of theory in the design of guideline dissemination and implementation strategies and interpretation of the results of rigorous evaluations. Implement Sci 2010 Feb 9;5:14 [FREE Full text] [doi: 10.1186/1748-5908-5-14] [Medline: 20181130]
20. Mayes T, de Freitas S. Coventry Univeristy Open Collections. London; 2004. Review of e-learning theories, frameworks and models URL: https://curve.coventry.ac.uk/open/file/8ff033fc-e97d-4cb8-aed3-29be7915e6b0/1/Review%20of%20e-learning%20theories.pdf [accessed 2019-02-28] [WebCite Cache ID 76WfNgujZ]
21. Colquhoun HL, Brehaut JC, Sales A, Ivers N, Grimshaw J, Michie S, et al. A systematic review of the use of theory in randomized controlled trials of audit and feedback. Implement Sci 2013 Jun 10;8:66 [FREE Full text] [doi: 10.1186/1748-5908-8-66] [Medline: 23759034]
22. McGowan IS. Towards a theory-based design framework for an effective Elearning computer programming course. In: Proceedings of the 13th International Conference on Cognition and Exploratory Learning in Digital Age. 2016 Presented at: CELDA 2016; October 28-30, 2016; Mannheim, Germany p. 10-16 URL: https://eric.ed.gov/?id=ED571406
23. de Leeuw RA, Westerman M, Nelson E, Ket JC, Scheele F. Quality specifications in postgraduate medical e-learning: an integrative literature review leading to a postgraduate medical e-learning model. BMC Med Educ 2016 Dec 8;16:168 [FREE Full text] [doi: 10.1186/s12909-016-0700-7] [Medline: 27390843]
24. Cook DA, Ellaway RH. Evaluating technology-enhanced learning: a comprehensive framework. Med Teach 2015;37(10):961-970. [doi: 10.3109/0142159X.2015.1009024] [Medline: 25782599]
25. Tudor CL, Riboli-Sasco E, Marcano BJ, Nikolaou C, Majeed A, Zary N, et al. Mobile learning for delivering health professional education. Cochrane Database Syst Rev 2015;9:1-13. [doi: 10.1002/14651858.CD011861]
26. Car J, Carlstedt-Duke J, Tudor CL, Posadzki P, Whiting P, Zary N, et al. Digital education in health professions: the need for overarching evidence synthesis. J Med Internet Res 2019;21(2):1-11 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
27. UN Statistics Division. 2013. International Standard Classification of Education: Fields of Education Trainin (ISCED-F 2013) g URL: https://unstats.un.org/unsd/classifications/Family/Detail/1046 [accessed 2019-03-04] [WebCite Cache ID 76cpcnXws]
28. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review. Med Educ 2007 Aug;41(8):737-745 [FREE Full text] [doi: 10.1111/j.1365-2923.2007.02777.x] [Medline: 17661881]
29. Childs S, Blenkinsopp E, Hall A, Walton G. Effective e-learning for health professionals and students--barriers and their solutions. A systematic review of the literature--findings from the HeXL project. Health Info Libr J 2005 Dec;22(Suppl 2):20-32. [doi: 10.1111/j.1470-3327.2005.00614.x] [Medline: 16279973]
30. Jwayyed S, Stiffler KA, Wilber ST, Southern A, Weigand J, Bare R, et al. Technology-assisted education in graduate medical education: a review of the literature. Int J Emerg Med 2011;4:51 [FREE Full text] [doi: 10.1186/1865-1380-4-51] [Medline: 21824405]

XSL•FO
RenderX

31. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A, Psychology Theory Group. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
32. Anders A. Theories and applications of Massive Online Open Courses (MOOCs): the case for hybrid design. Int Rev Res Open Dist Learn 2015 Dec 3;16(6):39-61. [doi: 10.19173/irrodl.v16i6.2185]
33. Jackson LD. Revisiting adult learning theory through the lens of an adult learner. Adult Learning 2009 Jun;20(3-4):20-22. [doi: 10.1177/104515950902000307]
34. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. Acad Med 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]
35. Goldie JG. Connectivism: a knowledge learning theory for the digital age? Med Teach 2016 Oct;38(10):1064-1069. [doi: 10.3109/0142159X.2016.1173661] [Medline: 27128290]
36. Zehry K, Halder N, Theodosiou L. E-learning in medical education in the United Kingdom. Procedia Soc Behav Sci 2011;15:3163-3167 [FREE Full text] [doi: 10.1016/j.sbspro.2011.04.265]
37. Nuninger W. Common scenario for an efficient use of online learning: some guidelines for pedagogical digital device development. In: Handbook of Research on Innovative Pedagogies and Technologies for Online Learning in Higher Education. USA: Pennsylvania, USA; 2016:331-366.
38. Axboe MK, Christensen KS, Kofoed P, Ammentorp J. Development and validation of a self-efficacy questionnaire (SE-12) measuring the clinical communication skills of health care professionals. BMC Med Educ 2016 Oct 18;16(1):272 [FREE Full text] [doi: 10.1186/s12909-016-0798-7] [Medline: 27756291]
39. Said H, Badru BB, Shahid M. Confirmatory factor analysis (Cfa) for testing validity and reliability instrument in the study of education. Aust J Basic Appl Sci 2011;5(12):1098-1103 [FREE Full text]
40. Martinez ME. Cognition and the question of test item format. Educ Psychol 1999 Sep;34(4):207-218. [doi: 10.1207/s15326985ep3404_2]
41. Hodges HF. Preparing new nurses with complexity science and problem-based learning. J Nurs Educ 2011 Jan;50(1):7-13. [doi: 10.3928/01484834-20101029-01] [Medline: 21053861]
42. Bandura A. Social Learning Theory. Englewood Cliffs, NJ: Prentice-Hall; 1977.
43. Mayer RE. Multimedia Learning, Second Edition. USA: Cambridge University Press; 2009.
44. Kane MT. Validating the interpretations and uses of test scores. J Educ Meas 2013 Mar 14;50(1):1-73. [doi: 10.1111/jedm.12000]
45. Beckman TJ, Ghosh AK, Cook DA, Erwin PJ, Mandrekar JN. How reliable are assessments of clinical teaching? A review of the published instruments. J Gen Intern Med 2004 Sep;19(9):971-977 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.40066.x] [Medline: 15333063]
46. Alshalabi IA, Hamada S, Elleithy K. Research learning theories that entail M-learning education elated to computer science and engineering courses. Int J Eng Sci 2013;2(3):88-95 [FREE Full text]
47. Koohang A, Harman K. Open source: a metaphor for e-learning. Inform Sci 2005;8:75-86 [FREE Full text]
48. Biggs J. Enhancing teaching through constructive alignment. High Educ 1996 Oct;32(3):347-364. [doi: 10.1007/BF00138871]
49. Biggs J. Constructive alignment in university teaching. HERDSA Rev High Educ 2014;1:5-22 [FREE Full text]

Abbreviations

DGBL: digital game–based learning
IQR: interquartile range
MCQs: multiple-choice questions
RCT: randomized controlled trials
TTAF: Theory-Technology Alignment Framework
VPS: virtual patient simulation
VR: virtual reality

XSL•FO
RenderX

Edited by A Marusic; submitted 11.12.18; peer-reviewed by S Gallagher, V Haldane, I Buljan; comments to author 04.01.19; revised version received 26.01.19; accepted 09.02.19; published 12.03.19.

Please cite as:
Bajpai S, Semwal M, Bajpai R, Car J, Ho AHY
Health Professions' Digital Education: Review of Learning Theories in Randomized Controlled Trials by the Digital Health Education Collaboration
J Med Internet Res 2019;21(3):e12912
URL: http://www.jmir.org/2019/3/e12912/
doi:10.2196/12912
PMID:30860483

©Shweta Bajpai, Monika Semwal, Ram Bajpai, Josip Car, Andy Hau Yan Ho. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Offline Digital Education for Medical Students: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration

Bhone Myint Kyaw[1], MBBS, MSc, PhD; Pawel Posadzki[2], PhD; Gerard Dunleavy[2], MSc; Monika Semwal[2], PhD; Ushashree Divakar[2], MSc; Vasilis Hervatis[3], PhD; Lorainne Tudor Car[1,4], MD, MSc, PhD

[1]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Centre for Population Health Sciences (CePHaS), Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[3]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[4]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Lorainne Tudor Car, MD, MSc, PhD
Family Medicine and Primary Care
Lee Kong Chian School of Medicine
Nanyang Technological University
11 Mandalay Road, Level 18, Clinical Science Building
Singapore,
Singapore
Phone: 65 69041258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Medical schools in low- and middle-income countries are facing a shortage of staff, limited infrastructure, and restricted access to fast and reliable internet. Offline digital education may be an alternative solution for these issues, allowing medical students to learn at their own time and pace, without the need for a network connection.

Objective: The primary objective of this systematic review was to assess the effectiveness of offline digital education compared with traditional learning or a different form of offline digital education such as CD-ROM or PowerPoint presentations in improving knowledge, skills, attitudes, and satisfaction of medical students. The secondary objective was to assess the cost-effectiveness of offline digital education, changes in its accessibility or availability, and its unintended/adverse effects on students.

Methods: We carried out a systematic review of the literature by following the Cochrane methodology. We searched seven major electronic databases from January 1990 to August 2017 for randomized controlled trials (RCTs) or cluster RCTs. Two authors independently screened studies, extracted data, and assessed the risk of bias. We assessed the quality of evidence using the Grading of Recommendations, Assessment, Development, and Evaluations criteria.

Results: We included 36 studies with 3325 medical students, of which 33 were RCTs and three were cluster RCTs. The interventions consisted of software programs, CD-ROMs, PowerPoint presentations, computer-based videos, and other computer-based interventions. The pooled estimate of 19 studies (1717 participants) showed no significant difference between offline digital education and traditional learning groups in terms of students' postintervention knowledge scores (standardized mean difference=0.11, 95% CI –0.11 to 0.32; small effect size; low-quality evidence). Meta-analysis of four studies found that, compared with traditional learning, offline digital education improved medical students' postintervention skills (standardized mean difference=1.05, 95% CI 0.15-1.95; large effect size; low-quality evidence). We are uncertain about the effects of offline digital education on students' attitudes and satisfaction due to missing or incomplete outcome data. Only four studies estimated the costs of offline digital education, and none reported changes in accessibility or availability of such education or in the adverse effects. The risk of bias was predominantly high in more than half of the included studies. The overall quality of the evidence was low (for knowledge, skills, attitudes, and satisfaction) due to the study limitations and inconsistency across the studies.

Conclusions: Our findings suggest that offline digital education is as effective as traditional learning in terms of medical students' knowledge and may be more effective than traditional learning in terms of medical students' skills. However, there is a need to further investigate students' attitudes and satisfaction with offline digital education as well as its cost-effectiveness, changes in its accessibility or availability, and any resulting unintended/adverse effects.

XSL•FO
RenderX

(J Med Internet Res 2019;21(3):e13165) doi:10.2196/13165

KEYWORDS

medical education; systematic review; meta-analysis; randomized controlled trials; students, medical

Introduction

There is a global shortage of 2.6 million medical doctors according to the World Health Organization [1]. In low- and middle-income countries, this shortage is further exacerbated by migrations, inadequacy of training programs, and poor infrastructure including limited access to fast and reliable internet connection [2-6]. Additionally, the content, structure, and delivery mode of medical curricula in these countries are often inadequate to equip medical students with the required knowledge, skills, and experience needed to meet their populations' evolving health care needs [7]. To tackle these multifaceted and intertwined problems, complex measures need to be taken to increase not only the number of medical doctors, but also the quality and relevance of their training [8]. Offline digital education offers a potential solution to overcome these problems.

Offline digital education, herein also referred to as computer-based learning or computer-assisted instruction, was one of the first forms of digital education, used before the internet became available on a global scale [9,10]. Unlike online digital education, offline digital education is independent of the internet or a local area network connection. Offline digital education can be delivered through CD-ROM, digital versatile disc (DVD)-ROM, external hard discs, universal serial bus (USB) memory sticks, or different software packages [11]. Offline digital education offers potential benefits over traditional modes of learning, including self-paced directed learning, stimulation of various senses (eg, with visual and spatial components) [12,13], and the ability to represent content in a variety of media (eg, text, sound, and motion) [14]. The educational content of the interventions is highly adaptable to the learners' needs, with the potential to be reviewed, repeated, and resumed at will. The interventions offer improved accessibility and flexibility and transcend the geographical, temporal, and financial boundaries that medical students may face. By reducing the costs of transportation or renting out classrooms and by freeing up the time of medical curriculum providers [14-16], the interventions may potentially offer substantial monetary savings.

A number of randomized control trials (RCTs) have evaluated the effectiveness of offline digital education in improving learning outcomes of medical students. Some of these trials were further evaluated in systematic reviews [13,17], but the findings were inconclusive. The primary objective of this systematic review was to evaluate the effectiveness of offline digital education compared with traditional learning or different forms of offline digital education in improving medical students' knowledge, skills, attitudes, and satisfaction. The secondary objective was to assess the economic impact of offline digital education, changes in its accessibility or availability, and its unintended/adverse effects.

Methods

Protocol

For this systematic review, we adhered to the published protocol [18]. The methodology has been described in detail by the Digital Health Education Collaboration [19]. The Digital Health Education collaboration is a global initiative focused on evaluating the effectiveness of digital health professions education through a series of methodologically robust systematic reviews.

Search Strategy and Data Sources

Electronic Searches

We developed a comprehensive search strategy for MEDLINE (Ovid; see Multimedia Appendix 1 for MEDLINE [Ovid] search strategy), Embase (Elsevier), Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley), PsycINFO (Ovid), Educational Research Information Centre (Ovid), Cumulative Index to Nursing and Allied Health Literature (Ebsco), and Web of Science Core Collection (Thomson Reuters). Databases were searched from January 1990 to August 2017. We selected 1990 as the starting year for our search because prior to this year, the use of computers was limited to very basic tasks. There were no language restrictions. We searched the reference lists of all the studies that we deemed eligible for inclusion in our review and the relevant systematic reviews. We also searched the International Clinical Trials Registry Platform Search Portal and metaRegister of Controlled Trials to identify unpublished trials.

We developed a common, comprehensive search strategy for a series of systematic reviews focusing on different types of digital education (ie, offline digital education, online digital education, and mobile learning) for preregistration as well as postregistration health care professionals. We retrieved 30,532 records from different bibliographic databases initially. In this review, we only included studies focusing on the effectiveness of offline digital education in medical students' education, and the findings on other types of digital education (such as virtual reality and mobile learning) within health professions education were reported separately [20-26].

For the purpose of this review, offline digital education can be defined as offline and stand-alone computer-based or computer-assisted learning where internet or intranet connection is not required for the learning activities. Traditional learning can be defined as learning via traditional forms of education such as paper- or text book–based learning and didactic or face-to-face-lecture. Blended digital education can be defined as any intervention that involves the combined use of offline digital education and traditional learning.

XSL•FO
RenderX

Inclusion Criteria

We included RCTs and cluster RCTs (cRCTs). Crossover trials were excluded due to a high likelihood of carry-over effect. We included studies with medical students enrolled in a preregistration, university degree program. Participants were not excluded based on age, gender, or any other sociodemographic characteristic.

We included studies in which offline digital education was used to deliver the learning content of the course. This included studies focused solely on offline digital education, or where offline digital education was part of a complex, multicomponent intervention. The main tasks of the learning activities were performed on a personal computer or laptop (with a hard keyboard). The delivery channel of the computer-based intervention was typically accessed via software programs, CD-ROM, DVD, hard disc, or USB memory stick. The focus was mainly on the learning activities that do not have to rely on any internet or online connection. Interventions where the internet connection was essential to provide learning content were excluded from this review.

We included the control groups that comprised traditional learning or traditional face-to-face learning such as lectures or discussions or text- or textbook-based learning as well as other offline digital education. We included studies that compared offline digital education or blended learning to traditional learning or a different form of offline digital education such as CD-ROM or PowerPoint presentations.

Learning outcomes were chosen based on the literature and relevance for medical students' education [27]. Eligible studies had to report at least one of the specified primary or secondary outcomes. Primary outcomes (measured using any validated or nonvalidated instruments) were medical students' knowledge scores (postintervention), medical students' cognitive skills (postintervention), medical students' postintervention attitudes toward the interventions or new clinical knowledge, and medical students' postintervention satisfaction with the interventions. Secondary outcomes included the economic impact of offline digital education (eg, cost-effectiveness, implementation cost, and return on investment), changes in its accessibility or availability, and any resulting adverse effects.

Data Collection and Analysis

Selection of Studies

The search results from different electronic databases were combined in a single Endnote (X.8.2) library, and duplicate records were removed [28]. Four review authors (BK, GD, MS, and UD) independently screened titles and abstracts of all the records to identify potentially eligible studies. We retrieved full-text copies of the articles deemed potentially relevant. Finally, two reviewers (BK and GD) independently assessed the full-text versions of the retrieved articles against the eligibility criteria. Any disagreements were resolved through discussion between the two reviewers, with a third review author (PP) acting as an arbiter, when necessary.

Data Extraction and Management

Five reviewers (BK, GD, MS, UD, and VH) independently extracted relevant characteristics related to participants, intervention, comparators, outcome measures, and results from all the included studies using a standard data-collection form. Any disagreements between the reviewers were resolved by discussion. We contacted the study authors for any missing information, particularly information required to judge the risk of bias.

Assessment of Risk of Bias in Included Studies

Four reviewers (BK, GD, MS, and UD) independently assessed the methodological risk of bias of included studies using the Cochrane methodology [29]. The following individual risk-of-bias domains were assessed in the included RCTs: random sequence generation, allocation concealment, blinding (outcome assessment), completeness of outcome data (attrition bias), selective outcome reporting (relevant outcomes reported), and other sources of bias (baseline imbalances).

For cRCTs, we assessed the risk of the following additional domains: recruitment bias, baseline imbalance, loss of clusters, incorrect analysis, and comparability with individually randomized trials recommended by Puffer et al [30]. Judgements concerning the risk of bias for each study were scored as high, low, or unclear. We incorporated the results of the risk-of-bias assessment into the review using a graph and a narrative summary. We also assessed publication bias using a funnel plot for comparisons with at least 10 studies.

Measures of Treatment Effect

For continuous outcomes, we reported mean postintervention scores and SD in each intervention group along with the number of participants and *P* values. We reported mean postintervention outcome data to ensure consistency across the included studies, as this was the most commonly reported form of findings. We presented outcomes using postintervention standardized mean difference (SMD) and interpreted the effect size using the Cohen rule of thumb (ie, with 0.2 representing a small effect, 0.5 representing a moderate effect, and 0.8 representing a large effect) [29,31]. For dichotomous outcomes, we calculated the risk ratio and 95% CIs. If studies had multiple arms, we compared the most active intervention arm to the least active control arm and assessed the difference in postintervention outcomes. We used the standard method recommended by Higgins et al to convert the results [29].

Data Synthesis

For meta-analysis, we used a random-effects model. For studies with the same continuous outcome measures, SMDs (for different scales) between groups, along with the 95% CIs, were estimated using Review Manager 5.3 [32]. In the analysis of continuous outcomes and cRCTs, we used the inverse variance method. We displayed the results of the meta-analyses in forest plots that provided effect estimates and 95% CIs for each individual study as well as a pooled effect estimate and 95% CI. For every step in the data analysis, we adhered to the statistical guidelines described by Higgins et al in 2011 [29].

XSL•FO
RenderX

Table 1. Summary of findings table: Effects of offline digital education on knowledge, skills, attitudes, and satisfaction. Patient or population: medical students, Settings: university or hospital, Intervention: offline digital education, Comparison: offline digital education versus traditional learning.

Outcomes	Illustrative comparative risks (95% CI)	Number of participants (number of studies)	Quality of the evidence (GRADE[a])	Comments
Knowledge: Assessed with multiple-choice questions, questionnaires, essays, quizzes, and practical section (from postintervention to 11-22 months of follow-up)	The mean knowledge score in offline digital education groups was 0.11 SD higher (–0.11 lower to 0.32 higher)	1717 (19)	Low[b,c,d]	The results from seven studies (689 participants) were not added to the meta-analysis due to incomplete or incomparable outcome data. These studies reported mixed findings: four studies (331 participants) favored offline digital education group, two studies reported no difference (289 participants), and one study favored the traditional learning group (69 participants).
Skills: Assessed with checklists, Likert-type scales, and questionnaires, (from postintervention to 1-10 months of follow-up)	The mean skills score in the offline digital education groups was 0.5 SD higher (0.25 higher to 0.75 higher)	415 (4)	Low[b,c,d]	The results of two studies (190 participants) were not added to the meta-analysis due to incomplete outcome data. One study (121 participants) favored offline digital education group. The other study (69 participants) reported no difference between the groups immediately postintervention and favored the offline digital education group at 1-month of follow-up.
Attitude: Assessed with Likert scale, questionnaires, and surveys (from postintervention to 5 weeks of follow-up)	Not estimable	493 (5)	Low[b,c,d]	One study (54 participants) reported higher postintervention attitude scores in offline digital education compared to traditional learning. We were uncertain about the effect of four studies (439 participants) due to incomplete outcome data.
Satisfaction: Assessed with Likert scales, questionnaires, and surveys (postintervention)	Not estimable	1442 (15)	Low[b,c,d]	Two studies (144 participants) favored traditional learning and two studies (103 participants) reported little or no difference between the groups. We were uncertain about the effect of 11 studies (1195 participants) due to incomplete outcome data.

[a]GRADE: Grading of Recommendations, Assessment, Development, and Evaluations.

[b]Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

[c]Rated down by one level for study limitations. The risk of bias was unclear for sequence generation and allocation concealment in majority of the studies.

[d]Rated down by one level for inconsistency. The heterogeneity is high with large variations in effect and lack of overlap among CIs.

We synthesized the findings from the included studies by the type of comparison: offline digital education (including PowerPoint and CD-ROM) versus traditional learning, offline digital education versus a different form of offline digital education, and blended learning versus traditional learning. Two authors (BK and GD) used the Grading of Recommendations, Assessment, Development, and Evaluations criteria to assess the quality of the evidence [33]. We considered the following criteria to evaluate the quality of the evidence, downgrading the quality where appropriate: limitations of studies (risk of bias), inconsistency of results (heterogeneity), indirectness of the evidence, imprecision (sample size and effect estimate), and publication bias. We prepared a summary of findings table [33] to present the results (Table 1). Where a meta-analysis was unfeasible, we presented the results in a narrative format, such as that used by Chan et al [34].

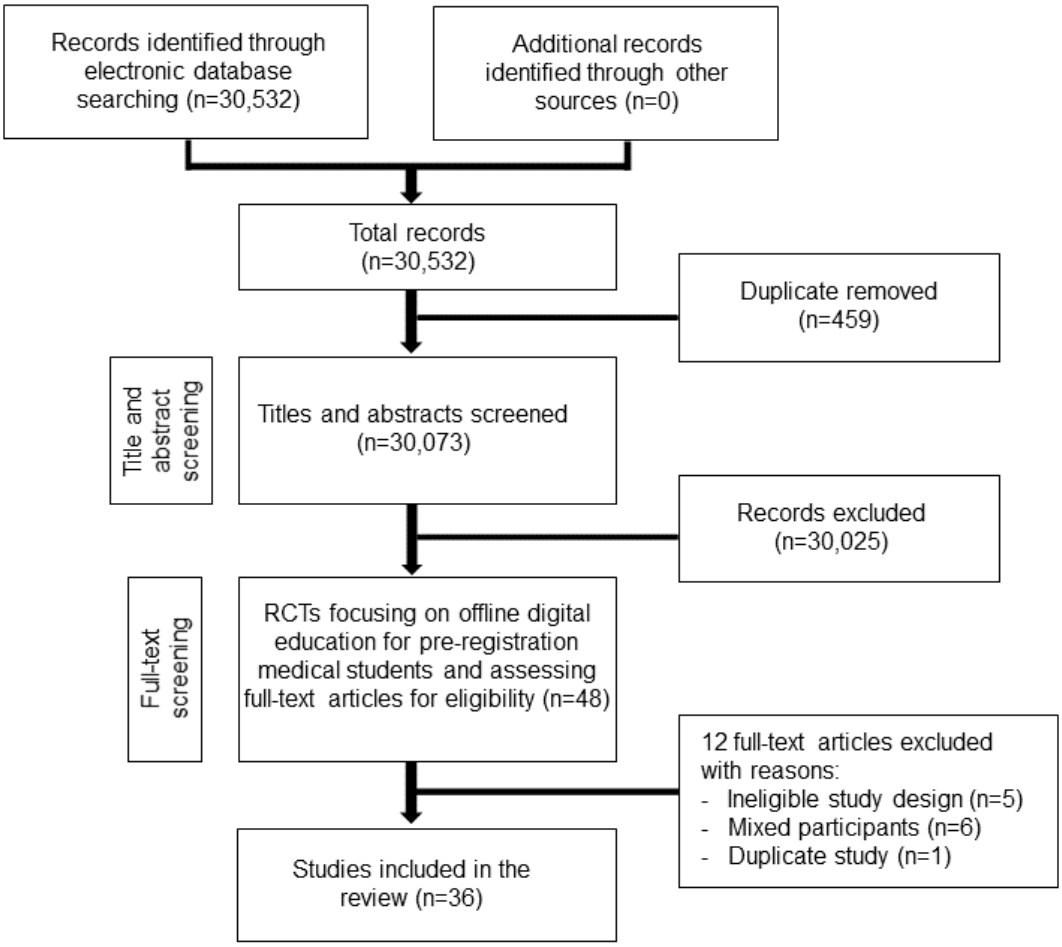
Results

Results of the Search

Our search strategy retrieved 30,532 unique references (Figure 1).

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. RCT: randomized controlled trials.



Included Studies

We included 36 studies from 35 reports involving 3325 participants [10,35-68] (Table 2), of which 33 studies were RCTs and the remaining three studies (from two reports) were cRCTs [56,68]. One study [68] reported the results of two cRCTs and we reported these results separately. Thirty-two studies (89%) were published before 2010, and only four studies (11%) were published after 2010 (Figure 2) [39,48,55,66].

The number of participants across the studies varied from 20 [35] to 241 [68], while individual studies focused on different areas of medical education. For the intervention groups, 20 studies used software programs [10,42,43,45,48-51,53,54,56,58-62,64-67], nine used CD-ROMs [35,36,40,44,46,55,57,68], four used PowerPoint presentations [37,38,41,63], two did not specify the type of intervention [47,52], and one used a computer-based video [39]. The duration of the interventions ranged from 10 minutes [39] to 3 weeks [41]. Four studies did not report the duration of the intervention [47,48,56,63]. The frequency of the intervention ranged from one [35,38-40,48,49,51,52,55,57-61,64,65,67-68] to six [50], and the intensity ranged from 10 minutes [39] to 11.1 hours [53]. Nine studies provided instructions on how to use the software [39,42,44,46,50,53,65,68]. Eight studies reported security arrangements [36,37,43,45,49,50,53,58]. For the control groups, 30 studies used traditional methods of learning such as face-to-face lectures, paper- or text book–based learning resources, laboratory courses, practical workshops, or small group tutorials [10,36-42,44,45,47-57,59-65,67,68]. Five studies used different forms of offline digital education as the controls [35,43,46,58,66]. One study compared blended learning (computer-assisted learning in addition to traditional learning) and traditional learning alone [44]. More information on the types of interventions is provided in Multimedia Appendix 2.

XSL•FO
RenderX

Table 2. Characteristics of the included studies.

Study, design, and country	Population (n), (medical student year)	Field of study	Outcomes
Ackermann et al 2010 [35], RCT[b], Germany	20 (not specified)	Surgery (orthopedic surgery)	Skill
Amesse 2008 [36], RCT, United States	36 (third year)	Radiology	Knowledge
Armstrong et al 2009 [37], RCT, United Kingdom	21 (fourth year)	Arterial blood gas interpretation	Knowledge and satisfaction
Carrero et al 2009 [38], RCT, Spain	68 (third year)	Basic life support algorithms	Knowledge
Cheng et al 2017 [39], RCT, United States	41 (second, third and fourth year)	Orthopedics	Skill
Davis et al 2008 [40], RCT, United Kingdom	229 (first year)	Evidence-based medicine	Knowledge and attitude
de Jong et al 2010 [41], RCT, The Netherlands	107 (third year)	Musculoskeletal problems	Knowledge and satisfaction
Desch et al 1991 [42], RCT, United States	78 (third year)	Pediatrics (neonatal management)	Knowledge, satisfaction, and cost
Devitt and Palmer 1999 [43], RCT, Australia	90 (second year)	Anatomy and physiology	Knowledge
Elves et al 1997 [44], RCT, United Kingdom	26 (third year)	Urology	Knowledge and satisfaction
Fasce et al 1995 [45], RCT, Chile	100 (fourth year)	Medicine (hypertension)	Knowledge, attitude, and satisfaction
Finley et al 1998 [46], RCT, Canada	40 (second year)	Medicine (auscultation of heart)	Knowledge and satisfaction
Gelb 2001 [47], RCT, United States	107 (not specified)	Anatomy	Knowledge and satisfaction
Green and Levi 2011 [48], RCT, United States	121 (second year)	Advanced care planning	Knowledge, skill, and satisfaction
Hilger et al 1996 [49], RCT, United States	77 (third year)	Medicine (pharyngitis)	Knowledge and attitude
Hudson 2004 [10], RCT, Australia	100 (third year)	Neuroanatomy and neurophysiology	Knowledge
Holt et al 2001 [50], RCT, United Kingdom	185 (first year)	Endocrinology	Knowledge, satisfaction, and cost
Lee et al 1997 [51], RCT, United States	82 (second year)	Biochemistry/acid-base problem solving	Knowledge and satisfaction
MacFadyen et al 1993 [52], RCT, Canada	54 (fourth year)	Clinical pharmacology	Knowledge and attitude
Mangione et al 1991 [53], RCT, United States	35 (third year)	Auscultation of the heart	Knowledge and attitude
McDonough and Marks 2002 [54], RCT, United Kingdom	37 (third year)	Psychiatry	Knowledge and satisfaction
Mojtahedzadeh et al 2014 [55], RCT, Iran	61 (third year)	Physiology of hematology and oncology	Knowledge and satisfaction
Nola et al 2005 [56], cRCT[c], Croatia	225 (not specified)	Pathology	Knowledge
Perfeito et al 2008 [57], RCT, Brazil	35 (fourth year)	Surgery	Knowledge and satisfaction
Pusic et al 2007 [58], RCT, Canada and United States	152 (final year)	Radiology	Knowledge and satisfaction
Ram 1997 [59], RCT, Malaysia	64 (final year)	Cardiology	Knowledge
Santer et al 1995 [60], RCT, United States	179 (third and fourth year)	Pediatrics	Knowledge and satisfaction
Seabra et al 2004 [61], RCT, Brazil	60 (second and third year)	Urology	Knowledge and satisfaction
Shomaker et al 2002 [62], RCT, United States	94 (second year)	Parasitology	Knowledge and satisfaction
Solomon et al 2004 [63], RCT, United States	29 (third year)	Learning concepts (digital and live lecture formats)	Knowledge
Stanford et al 1994 [64], RCT, United States	175 (first year)	Anatomy (cardiac anatomy)	Knowledge and satisfaction
Summers et al 1999 [65], RCT, United States	69 (first year)	Surgery	Knowledge and skill
Taveira-Gomes et al 2015 [66], RCT, Portugal	96 (fourth and fifth year)	Cellular biology	Knowledge

XSL•FO
RenderX

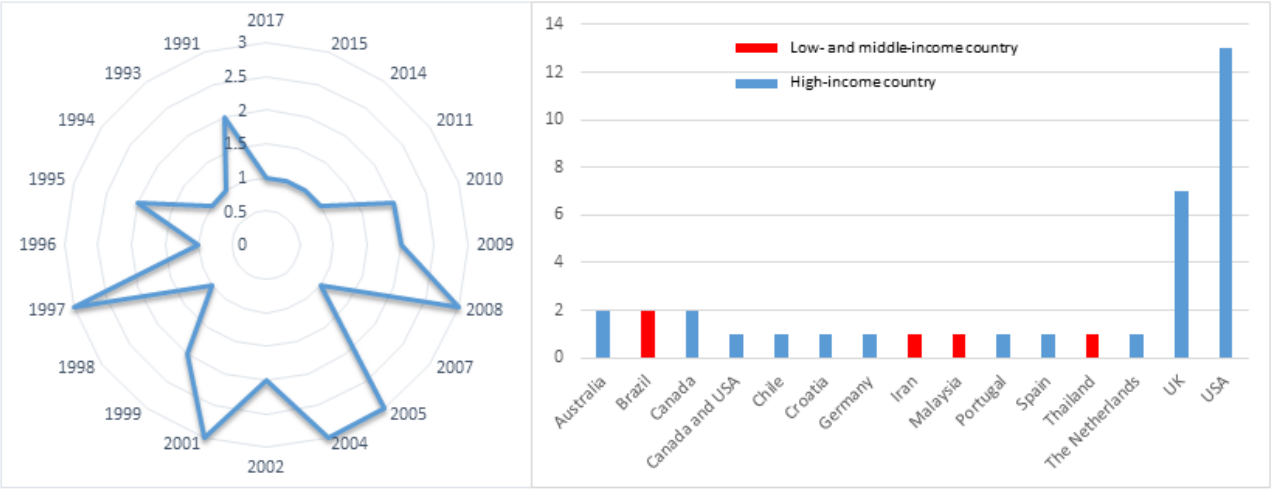
Study, design, and country	Population (n), (medical student year)	Field of study	Outcomes
Vichitvejpaisal et al 2001 [67], RCT, Thailand	80 (third year)	Arterial blood gas interpretation	Knowledge
Vivekananda-Schmidt et al 2005 [68], cRCT, United Kingdom (Newcastle)[c]	241 (third year)	Orthopedics (musculoskeletal examination skills)	Skill and cost
Vivekananda-Schmidt et al 2005 [68], cRCT, United Kingdom (London)[c]	113 (third year)	Orthopedics (musculoskeletal examination skills)	Skill and cost

[a]RCT: randomized controlled trial.

[b]cRCT: cluster randomized controlled trial.

[c]This study reported its results from two separate cRCTs. We analyzed data from the two cRCTs separately.

Figure 2. Number of publication(s) on offline digital education in relation to their year of publication and country of origin.



Knowledge

Overview

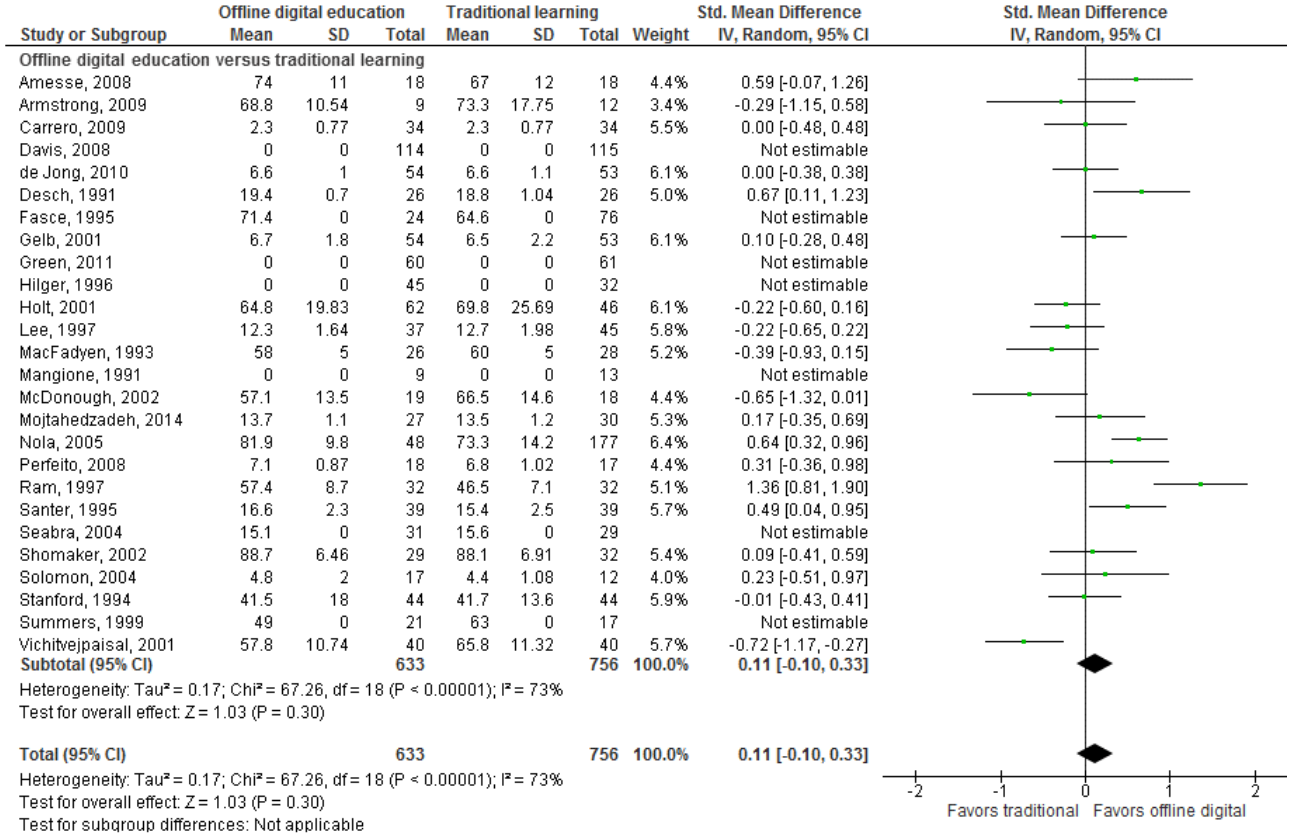
Knowledge was assessed as the primary outcome in 32 studies [10,36-38,40-67], and majority of the studies (69%) used multiple-choice questions or questionnaires to measure the outcome. Twenty-nine studies assessed knowledge using nonvalidated instruments [36-38,40-49,51-57,59-67], while three studies [10,50,58] used validated instruments such as multiple-choice questions, questionnaires, and tests. Twenty-six studies assessed postintervention knowledge scores, while six studies assessed both short-term and long-term knowledge retention, ranging from 1 week to 22 months of follow-up [52,60,62,65-67] (Multimedia Appendix 3).

Offline Digital Education Versus Traditional Learning

Meta-analysis of 19 studies showed that there was no significant difference between offline digital education and traditional learning in postintervention knowledge scores (SMD 0.11, 95% CI –0.11 to 0.32; 1717 participants; small effect size, low-quality evidence; Figure 3). There was a substantial amount of heterogeneity in the pooled analyses (I^2=73%). The remaining seven studies were not pooled due to incomplete or incomparable outcome data [40,45,48,49,53,61,65] and reported mixed findings. Four studies reported a significant difference in postintervention knowledge scores in favor of offline digital education (331 participants) [45,48,49,53]. Two studies reported no significant difference between the interventions (289 participants) [40,61], and one study (69 participants) reported a significant difference in postintervention knowledge scores in favor of traditional learning [65]. Taken together, these findings suggest that offline digital education had similar effects as traditional learning on medical students' postintervention knowledge scores.

XSL•FO
RenderX

Figure 3. Forest plot of studies comparing offline digital education with traditional, postintervention knowledge outcome. IV=interval variable; random=random effect model.



Blended Learning Versus Traditional Learning

One study (26 participants) assessed postintervention knowledge scores in blended learning (offline digital education plus traditional learning) versus traditional learning alone and reported a significant difference in favor of blended learning (SMD 0.81, 95% CI 0.01-1.62; large effect size) [44].

Offline Digital Education Versus Offline Digital Education

Five studies (478 participants) compared one form of offline digital education (eg, CD-ROM and software programs) to another form of offline digital education (eg, CD-ROM, software programs, or computer-assisted learning programs) [10,43,46,58,66].

Devitt et al reported higher knowledge scores in the computer-based group than a free-text entry program (SMD 1.62, 95% CI 0.93-2.31; large effect size) at 2 weeks postintervention [43]. Taveira-Gomes et al also reported higher postintervention knowledge scores in the computer-based software group (ie, the use of *flashcards-based learning materials on cellular structure*) compared to a computer-based method alone (ie, without the use of flashcards; SMD 2.17, 95% CI 1.67-2.67, large effect size) [66]. Pusic et al reported that the effectiveness of a simple linear computer program was equivalent to that of a more interactive, branched version of the program in terms of postintervention knowledge scores (SMD 0, 95% CI –0.33 to 0.33; small effect size) [58]. The effect of two studies was uncertain due to incomplete outcome data [10,46]. Overall, the findings were mixed and inconclusive.

Skills

Overview

Six studies from five reports (605 participants) assessed skills as a primary outcome [35,39,48,65,68]. Three studies used validated instruments to measure the outcome such as an Objective Structured Clinical Examination [68] and a checklist rating form [65]. The remaining three studies used nonvalidated instruments such as questionnaires [35], a six-point checklist scale [39], and self-assessment [48]. Four studies [35,39,48,68] assessed postintervention skills scores, while two studies [65,68] assessed both short-term and long-term skill retention, ranging from 1 month to 10 months of follow-up (Multimedia Appendix 3).

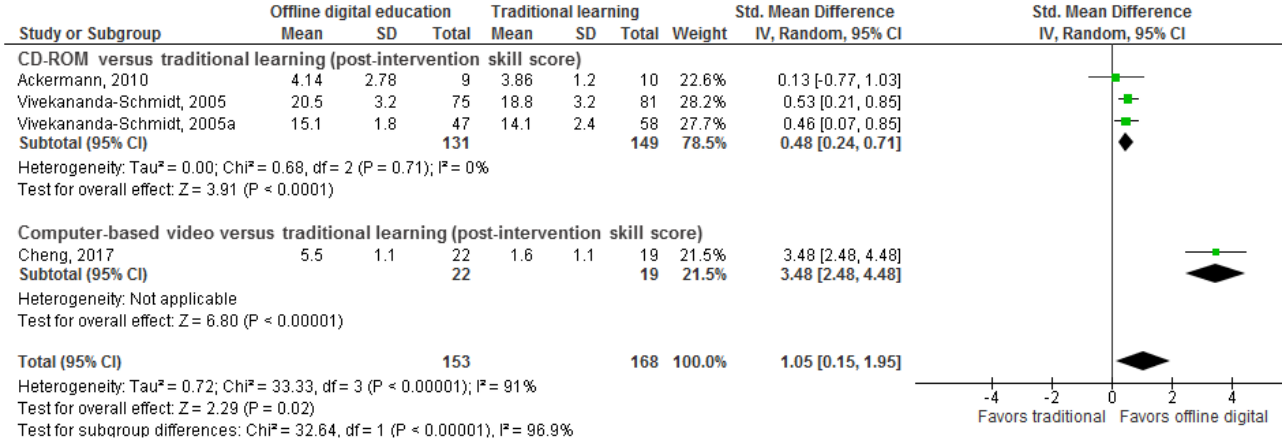
Offline Digital Education Versus Traditional Learning

Meta-analysis of four studies showed that, compared with traditional learning, offline digital education improved medical students' postintervention skill scores (SMD 1.05, 95% CI 0.15-1.95; I^2=91%; large effect size; low-quality evidence; Figure 4). There was, however, a considerable amount of heterogeneity in the pooled analyses (I^2=91%).

The results of two studies were not pooled due to incomplete outcome data [48,65]. These studies also reported higher postintervention skill scores with offline digital education than with traditional learning [48,65]. Taken together, these results suggest that offline digital education may improve postintervention skill scores compared to traditional learning.

XSL•FO
RenderX

Figure 4. Forest plot of studies comparing offline digital education with traditional, postintervention skill outcome. IV=interval variable; random=random effect model. Vivekananda-Schmidt, 2005 was conducted in Newcastle and Vivekananda-Schmidt, 2005a was conducted in London.



Attitude

Five studies (493 participants) reported students' postintervention attitude toward the intervention or new clinical knowledge [40,45,49,52,53]. Attitudes were measured using Likert-based questionnaires [40,49,52], surveys [45], and the Computer Anxiety Index [53]. None of the studies used validated tools to measure the outcome. MacFadyen et al reported higher postintervention attitude scores in the offline digital education group than in the traditional learning group (SMD 2.71, 95% CI 1.96-3.47; large effect size) [52]. One study reported higher attitude scores in the intervention group than in the traditional learning group (89% vs 47%) [45]. Two studies did not report numerical data for either of the study groups [40,49], while one study assessed participants' postintervention attitudes in the intervention group only [53]; hence, we were unable to judge the effect of these three interventions due to missing outcome data [40,45,49,53]. Taken together, the overall effect of the interventions seems uncertain due to the lack of outcome data in most of the included studies.

Satisfaction

Eighteen studies (1660 participants) assessed postintervention satisfaction [37,41,42,44-48,50,51,54,55,57,58,60-62,64]. Nine studies used Likert-type rating scales [42,46,48,51,54,55,58,61,62], eight studies used questionnaires [37,41,44,47,50,57,60,64], and one study [45] used a survey to assess participants' postintervention satisfaction. None of the studies used validated tools to measure the outcome.

Fifteen studies comparing offline digital education with traditional learning assessed satisfaction. Two studies [41,54] reported higher postintervention satisfaction scores in the traditional learning group than in offline digital education (risk ratio=0.46, 95% CI 0.30-0.69; small effect size; SMD –1.33, 95% CI –2.05 to –0.61, large effect size). Two other studies reported no significant difference in students' postintervention satisfaction between the groups [37,51]. The remaining 11 studies reported incomplete or incomparable outcome data [42,45,47,48,50,55,57,60-62,64]. Overall, we were uncertain about the effects of offline digital education on students' satisfaction scores, when compared with traditional learning.

Three studies comparing different forms of offline digital education and blended learning to traditional learning also assessed satisfaction. However, we were unable to judge the overall effect of the intervention in the three studies due to missing or incomparable outcome data [44,46,58].

Secondary Outcomes

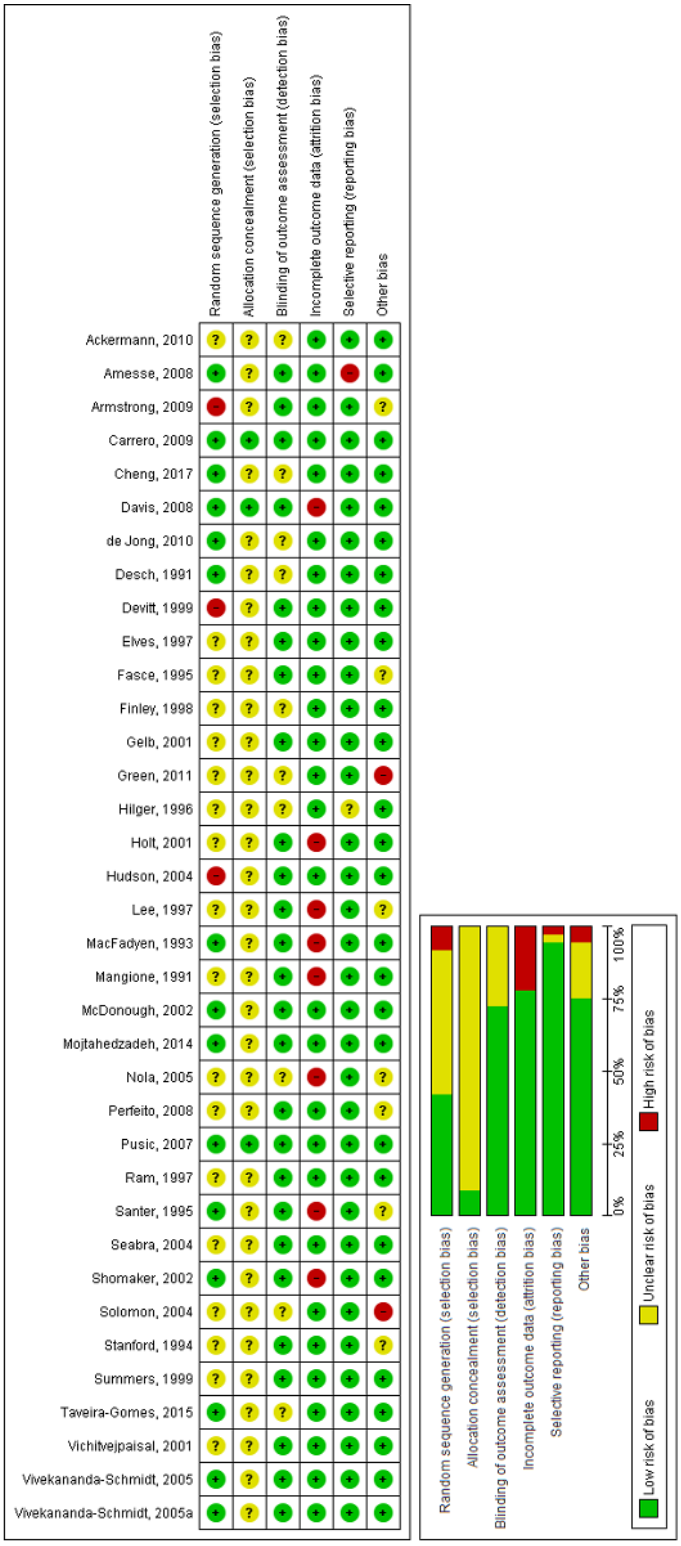
Four studies (617 participants) reported the cost of the offline digital education [42,50,68]. However, none of the included studies compared costs between the intervention and control groups.

Desch et al reported that the authoring system (computer-assisted instructional software program) costs US $600. Additionally, the study used US $1500 to hire a student to develop the program [42]. The microcomputers used by the students in the study by Desch et al were within a large microcomputer area in the medical library and were used for multiple purposes. Holt et al reported that the total cost of the equipment specially needed to set up the computer-assisted learning course (including slide and document scanners, sound recording, a laptop, and software) was approximately £3000 (~US $4530) [50]. Two studies reported that the cost of designing a virtual rheumatology CD was £11,740 (US $22,045) [68].

No studies reported adverse or unintended effects of the interventions or changes in the accessibility or availability of digital offline education.

XSL•FO
RenderX

Figure 5. Risk of bias summary: review authors' judgements about each risk of bias item across all included studies.

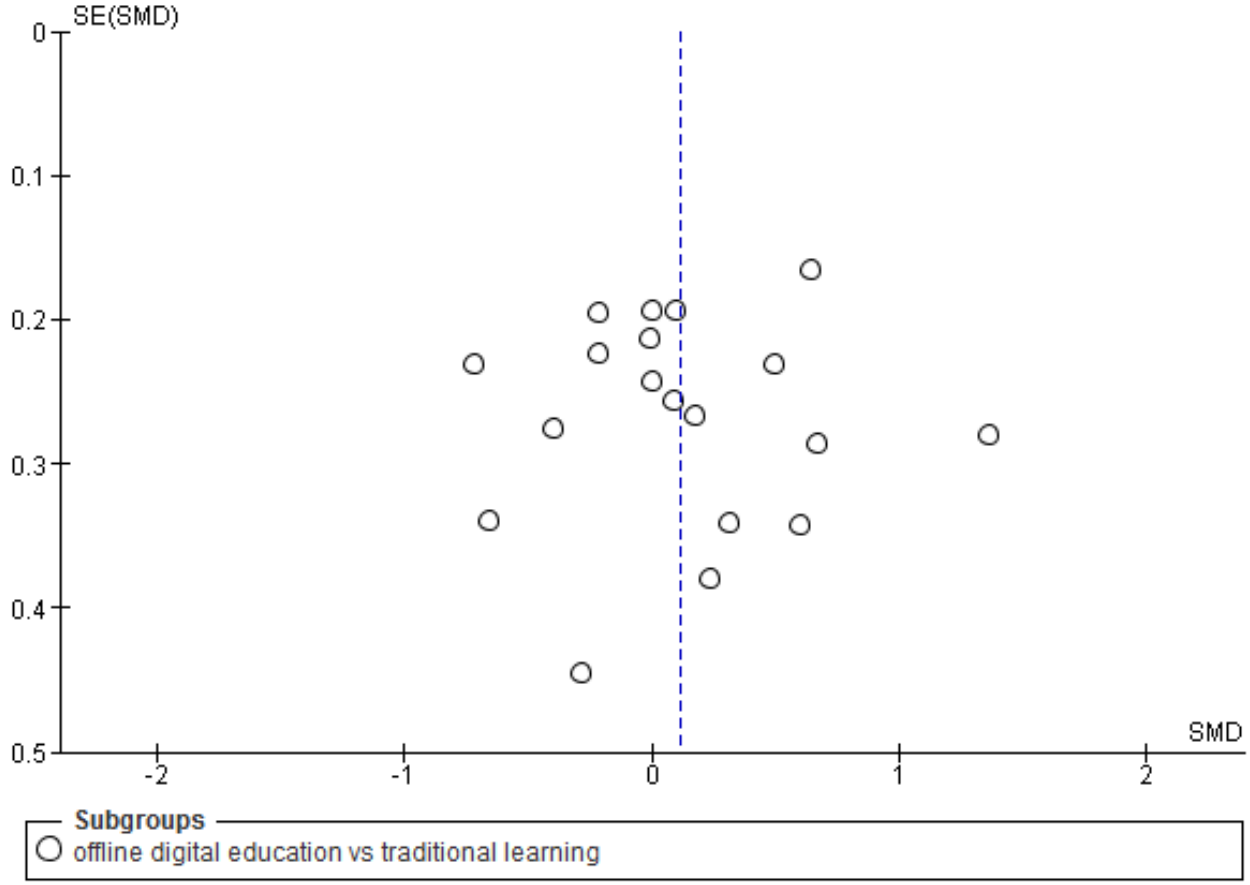


Risk of Bias in Included Studies

As presented in Figure 5, the risk of bias was generally unclear or high in most of the studies because of a lack of relevant information in the included studies. For 14 (39%) studies, we found that the risk of bias was low in at least four of six domains [38,39,41,42,54,55,58,59,61,65-68]. For 22 studies (61%), we found that the risk of bias was high, as the studies had an unclear risk of bias in at least three of six domains or a high risk in at least one domain [10,35-37,40,43,44,45,46,47,48-53,56,57,60,62-64]. A symmetrical funnel plot of studies comparing offline digital education and traditional learning suggests low risk of publication bias for the outcome knowledge (Figure 6). The overall risk of bias for cRCTs was unclear due to limited information from included studies (Multimedia Appendix 4).

XSL•FO
RenderX

Figure 6. Funnel plot of studies comparing offline digital education with traditional, postintervention knowledge outcome. SMD: standardized mean difference.



Discussion

Principal Findings

Our findings show that offline digital education is as effective as traditional learning in improving medical students' postintervention knowledge and may be more effective in improving skills, with effect sizes ranging from small (for knowledge) to large (for skills). We are uncertain about the effects for attitudes and satisfaction due to missing data or incomplete reporting. None of the studies reported on changes in accessibility or availability for education or adverse effects of the interventions. Only four studies reported the cost of offline digital education interventions; however, no estimates for comparator groups were provided.

Several limitations in the included literature need to be highlighted. For instance, we found that the evidence was of low quality due to the predominantly high risk of bias (studies' limitations) or inconsistency (high heterogeneity of the pooled analyses). Furthermore, the included studies were highly heterogeneous in terms of student populations (years 1-5), comparator groups (traditional learning and different forms of offline digital education), outcomes and measurement tools (multiple-choice questionnaires, surveys, Likert-type scales, questionnaires, essays, quizzes, practical sections), study designs, settings (university or hospital), and interventions (learning contents, types of delivery mode, duration, frequency, intensity, and security arrangements). In addition, the duration frequency of the interventions were highly variable. Although the included studies encompass a reasonable range of interventions and content, the data are mostly limited to high-income countries, thereby limiting the generalizability of our findings to other settings including low- and middle-income countries.

We also found that reporting in the included studies was often poor. For example, four studies (11%) reported on the cost of setup of the interventions only (without any comparison data). Moreover, we found that none of the studies used learning theories underpinning the development or application of the offline digital education. Most of the studies (92%) used nonvalidated measurement instruments to quantify the outcomes, thereby jeopardizing the reliability and credibility of digital education research. Furthermore, 20 studies (56%) used software/computer programs as the main mode of delivery of the learning content. However, the technical aspects of these programs such as design or functions were often omitted from the studies.

Offline digital education has the potential to play an important role in medical students' education, especially in low- and middle-income countries. Implementing offline digital education in medical education may require much less investment and infrastructure than alternative forms of digital education (eg, virtual reality or online computer-based education). Because of its scalability, offline digital education has the potential to reduce the shortage of medical doctors. It could be a major (for

XSL•FO
RenderX

low- and middle-income countries) or an alternative (for high-income countries) mode of delivering education for medical schools across the world, as more than 4 billion people still did not have access to the internet as of 2016 [6].

To the best of our knowledge, there are only two reviews available in the literature that examined the effectiveness of offline digital education among similar populations [13,17]. One of these reviews, published in 2001, suggested that offline digital education (computer-assisted learning) could reduce the costs of education and increase the number of medical students [13]. However, no formal assessments on cost-related outcomes were made, and further research was recommended. A review by Rasmussen et al stated that offline digital education was equivalent or possibly superior to traditional learning in improving knowledge, skills, attitudes, and satisfactions of preregistration health professionals [17], which is largely in line with our findings. However, Rasmussen et al applied a much narrower search timeframe and focused on all preregistration health care professionals (ie, including students from medical, dental, nursing, and allied health care fields) and could not provide specific recommendations for medical students' education. Our review provides up-to-date evidence with a comprehensive search strategy and a focus on medical students' education and includes meta-analyses of studies for knowledge and skills.

Strengths

Strengths of this systematic review include comprehensive searches with no language limitations, robust screening, independent data extractions, and risk-of-bias assessments. The review includes studies from the year 1990 in order to report the most comprehensive evidence and provides up-to-date evidence on the effectiveness of different types of offline digital education for medical students' education.

Limitations

Some limitations must be acknowledged while interpreting the results. First, we were unable to obtain missing information from the study authors despite multiple attempts. Second, we presented postintervention data rather than mean change scores, as the majority of the included studies (81%) reported postintervention data and only seven studies (19%) reported mean change scores. Third, we were unable to determine whether the study administrators received any incentives from the software or program developers, which might constitute bias. Lastly, we were unable to carry out prespecified subgroup analysis because of an insufficient number of studies under respective outcomes and because of the considerable heterogeneity of populations, interventions, comparators, and outcome measures used.

Implications for Research and Practice

We believe that offline digital education interventions can be practically introduced in medical students' education for improving their knowledge and skills in places where internet connectivity is limited, which may be of most concern in low- and middle-income countries. However, when interpreting the findings of this systematic review, stakeholders need to consider other factors such as students' geographical location or features of the intervention such as interactivity, duration, frequency, intensity, and delivery mode.

Future studies should evaluate the cost-effectiveness, sustainability, and indirect (and direct) costs of the interventions (eg, time to develop or implement the educational module). Future research should also report on potential (or actual) adverse effects of the interventions. In addition, most of the studies assessed short-term effectiveness of the interventions; hence, there is a need to evaluate knowledge and skill retention during longer follow-ups (eg, 6-12 months). Additionally, other aspects of the interventions such as different levels of interactivity or feedback in low- and middle-income countries still need to be explored. Addressing these gaps in evidence will help policy makers and curriculum planners allocate resources appropriately.

Conclusions

The findings from this review suggest that offline digital education is as effective as traditional learning in terms of medical students' knowledge and may be more effective in improving their skills. However, the evidence on other outcomes is inconclusive or limited. Future research should evaluate the effectiveness of offline digital education interventions in low- and middle-income countries and report on outcomes such as attitudes, satisfaction, adverse effects, and economic impact.

Acknowledgments

This review was conducted in collaboration with the Health Workforce Department at the World Health Organization. We would also like to thank Mr Carl Gornitzki, Ms GunBrit Knutssön, and Mr Klas Moberg from the University Library, Karolinska Institutet, Sweden, for developing the search strategy and the peer reviewers for their comments. We gratefully acknowledge funding from the Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore (eLearning for health professionals' education grant). Additionally, we would like to thank Ms Sabrina Andrea Spieck, Ms Maja Magdalena Olsson, and Mr Geronimo Jimenez for translating German and Spanish papers. We also would like to thank Dr Olena Zhabenko and Dr Ram Chandra Bajpai for their suggestions in data analysis.

Authors' Contributions

LC conceived the idea for the review. BK, PP, and GD wrote the review. LC, PP, and MS peer reviewed the review. LC provided methodological guidance on the review. PP, MS, UD, VH, and LTC provided comments on the review.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE (Ovid) Search Strategy.

[PDF File (Adobe PDF File), 72KB - jmir_v21i3e13165_app1.pdf]

Multimedia Appendix 2

Characteristics of the included studies.

[PDF File (Adobe PDF File), 54KB - jmir_v21i3e13165_app2.pdf]

Multimedia Appendix 3

Results of the included studies.

[PDF File (Adobe PDF File), 374KB - jmir_v21i3e13165_app3.pdf]

Multimedia Appendix 4

Risk of bias for cluster randomized controlled trials.

[PDF File (Adobe PDF File), 57KB - jmir_v21i3e13165_app4.pdf]

References

1. World Health Organization. 2016. High-Level Commission on Health Employment and Economic Growth Health workforce URL: http://www.who.int/hrh/com-heeg/en/ [accessed 2018-11-14] [WebCite Cache ID 73uaCutCo]
2. Anand S, Bärnighausen T. Human resources and health outcomes: cross-country econometric study. Lancet 2004;364(9445):1603-1609. [doi: 10.1016/S0140-6736(04)17313-3] [Medline: 15519630]
3. Campbell J, Dussault G, Buchan J, Pozo-Martin F, Guerra AM, Leone C, et al. World Health Organization. 2013. A universal truth: no health without a workforce URL: https://www.who.int/workforcealliance/knowledge/resources/GHWA_AUniversalTruthReport.pdf [accessed 2019-03-01] [WebCite Cache ID 76Xhm6eKa]
4. Chen LC. Striking the right balance: health workforce retention in remote and rural areas. Bull World Health Organ 2010 May;88(5):323, A [FREE Full text] [doi: 10.2471/BLT.10.078477] [Medline: 20461215]
5. Hongoro C, McPake B. How to bridge the gap in human resources for health. Lancet 2004;364(9443):1451-1456. [doi: 10.1016/S0140-6736(04)17229-2] [Medline: 15488222]
6. Luxton E. World Economic Forum. 4 billion people still don't have internet access. Here's how to connect them URL: https://www.weforum.org/agenda/2016/05/4-billion-people-still-don-t-have-internet-access-here-s-how-to-connect-them/ [accessed 2019-02-27] [WebCite Cache ID 74aFf1JDH]
7. Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet 2010 Dec 04;376(9756):1923-1958. [doi: 10.1016/S0140-6736(10)61854-5] [Medline: 21112623]
8. World Health Organization. 2015. eLearning for undergraduate health professional education: a systematic review informing a radical transformation of health workforce development URL: http://www.who.int/hrh/documents/14126-eLearningReport.pdf [accessed 2018-11-14] [WebCite Cache ID 73ua0wCm4]
9. Carr MM, Reznick RK, Brown DH. Comparison of computer-assisted instruction and seminar instruction to acquire psychomotor and cognitive knowledge of epistaxis management. Otolaryngol Head Neck Surg 1999 Oct;121(4):430-434. [doi: 10.1016/S0194-5998(99)70233-0] [Medline: 10504600]
10. Hudson JN. Computer-aided learning in the real world of medical education: does the quality of interaction with the computer affect student learning? Med Educ 2004 Aug;38(8):887-895. [doi: 10.1111/j.1365-2929.2004.01892.x] [Medline: 15271050]
11. Crisp N, Gawanas B, Sharp I, Task Force for Scaling Up Education and Training for Health Workers. Training the health workforce: scaling up, saving lives. The Lancet 2008 Feb;371(9613):689-691. [doi: 10.1016/s0140-6736(08)60309-8] [Medline: 18295028]
12. Brandt MG, Davies ET. Visual-spatial ability, learning modality and surgical knot tying. Can J Surg 2006 Dec;49(6):412-416 [FREE Full text] [Medline: 17234070]
13. Greenhalgh T. Computer assisted learning in undergraduate medical education. BMJ 2001 Jan 06;322(7277):40-44 [FREE Full text] [Medline: 11141156]
14. Glicksman JT, Brandt MG, Moukarbel RV, Rotenberg B, Fung K. Computer-assisted teaching of epistaxis management: a Randomized Controlled Trial. Laryngoscope 2009 Mar;119(3):466-472. [doi: 10.1002/lary.20083] [Medline: 19235749]

XSL•FO
RenderX

15. Hunt CE, Kallenberg GA, Whitcomb ME. Medical students' education in the ambulatory care setting: background paper 1 of the Medical School Objectives Project. Acad Med 1999 Mar;74(3):289-296. [Medline: 10099654]
16. Woo MK, Ng KH. A model for online interactive remote education for medical physics using the Internet. J Med Internet Res 2003;5(1):e3 [FREE Full text] [doi: 10.2196/jmir.5.1.e3] [Medline: 12746208]
17. Rasmussen K, Belisario JM, Wark PA, Molina JA, Loong SL, Cotic Z, et al. Offline eLearning for undergraduates in health professions: A systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010405 [FREE Full text] [doi: 10.7189/jogh.04.010405] [Medline: 24976964]
18. Hervatis V, Kyaw B, Semwal M, Dunleavy G, Tudor CL, Zary N. PROSPERO CRD42016045679. 2016 Apr 15. Offline and computer-based eLearning interventions for medical students' education [Cochrane Protocol] URL: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016045679 [accessed 2019-02-18] [WebCite Cache ID 76GXpzD6T]
19. 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 Dec 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
20. Posadzki P, Bala M, Kyaw B, Semwal M, Divakar U, Koperny M, et al. Offline digital education for post-registration health professions: a systematic review by the Digital Health Education collaboration. J Med Internet Res (forthcoming) 2018:a [FREE Full text]
21. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
22. Tudor Car L, Kyaw BM, Dunleavy G, Smart NA, Semwal M, Rotgans JI, et al. Digital Problem-Based Learning in Health Professions: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 28;21(2):e12945 (forthcoming) [FREE Full text] [doi: 10.2196/12945] [Medline: 30816846]
23. Gentry SV, Gauthier A, Ehrstrom BL, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: a systematic review by the Digital Health Education Collaboration. J Med Internet Res (forthcoming) 2018:a (forthcoming) [FREE Full text]
24. Dunleavy G, Nikolaou CK, Nifakos S, Atun R, Law GCY, Tudor Car L. Mobile Digital Education for Health Professions: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 12;21(2):e12937 [FREE Full text] [doi: 10.2196/12937] [Medline: 30747711]
25. Semwal M, Whiting P, Bajpai R, Bajpai S, Kyaw BM, Tudor Car L. Digital Education for Health Professions on Smoking Cessation Management: Systematic Review by the Digital Health Education Collaboration. J Med Internet Res 2019 Mar 04;21(3):e13000 [FREE Full text] [doi: 10.2196/13000] [Medline: 30829576]
26. Huang Z, Semwal M, Lee SY, Tee M, Ong W, Tan WS, et al. Digital Health Professions Education on Diabetes Management: Systematic Review by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 21;21(2):e12997 [FREE Full text] [doi: 10.2196/12997] [Medline: 30789348]
27. Barr H, Freeth D, Hammick M, Koppel I, Reeves S. CAIPE: Centre for the Advancement of Interprofessional Education. 2000. Evaluations of interprofessional education: a United Kingdom review for health and social care URL: https://www.caipe.org/resources/publications/barr-h-freethd-hammick-m-koppel-i-reeves-s-2000-evaluations-of-interprofessional-education [accessed 2019-03-07] [WebCite Cache ID 76hlNDi45]
28. Endnote (X.8.2). Philadelphia, PA: Clarivate Analytics; 2018. URL: https://endnote.com/downloads/available-updates/ [accessed 2019-03-01] [WebCite Cache ID 76XdZ69fO]
29. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. London: The Cochrane Collaboration; 2011. URL: http://crtha.iums.ac.ir/files/crtha/files/cochrane.pdf [accessed 2019-03-01] [WebCite Cache ID 76XcKQdGg]
30. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003 Oct 04;327(7418):785-789 [FREE Full text] [doi: 10.1136/bmj.327.7418.785] [Medline: 14525877]
31. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 07;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
32. RevMan. Review Manager 5.3. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2014. URL: https://community.cochrane.org/help/tools-and-software/revman-5 [accessed 2019-03-01] [WebCite Cache ID 76Xbuy3C8]
33. Schünemann HJ, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Presenting results and 'Summary of findings' tables (Chapter 11). In: Cochrane Handbook for Systematic Reviews of Interventions. London: The Cochrane Collaboration; 2008.
34. Chan RJ, Webster J, Marquart L. Information interventions for orienting patients and their carers to cancer care facilities. Cochrane Database Syst Rev 2011 Dec 07(12):CD008273. [doi: 10.1002/14651858.CD008273.pub2] [Medline: 22161425]
35. Ackermann O, Siemann H, Schwarting T, Ruchholtz S. [Effective skill training by means of E-learning in orthopaedic surgery]. Z Orthop Unfall 2010 May;148(3):348-352. [doi: 10.1055/s-0029-1240549] [Medline: 20135601]

XSL•FO
RenderX

36. Amesse LS, Callendar E, Pfaff-Amesse T, Duke J, Herbert WN. Evaluation of Computer-aided Strategies for Teaching Medical Students Prenatal Ultrasound Diagnostic Skills. Med Educ Online 2008 Sep 24;13:13 [FREE Full text] [doi: 10.3885/meo.2008.Res00275] [Medline: 20165541]
37. Armstrong P, Elliott T, Ronald J, Paterson B. Comparison of traditional and interactive teaching methods in a UK emergency department. Eur J Emerg Med 2009 Dec;16(6):327-329. [doi: 10.1097/MEJ.0b013e32832b6375] [Medline: 19421067]
38. Carrero E, Gomar C, Penzo W, Fábregas N, Valero R, Sánchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. Med Teach 2009 May;31(5):e189-e195. [doi: 10.1080/01421590802512896] [Medline: 19241215]
39. Cheng Y, Liu DR, Wang VJ. Teaching Splinting Techniques Using a Just-in-Time Training Instructional Video. Pediatr Emerg Care 2017 Mar;33(3):166-170. [doi: 10.1097/PEC.0000000000000390] [Medline: 25834963]
40. Davis J, Crabb S, Rogers E, Zamora J, Khan K. Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: a randomized controlled trial. Med Teach 2008;30(3):302-307. [doi: 10.1080/01421590701784349] [Medline: 18484458]
41. de Jong Z, van Nies JA, Peters SW, Vink S, Dekker FW, Scherpbier A. Interactive seminars or small group tutorials in preclinical medical education: results of a randomized controlled trial. BMC Med Educ 2010 Nov 13;10:79 [FREE Full text] [doi: 10.1186/1472-6920-10-79] [Medline: 21073744]
42. Desch LW, Esquivel MT, Anderson SK. Comparison of a computer tutorial with other methods for teaching well-newborn care. Am J Dis Child 1991 Nov;145(11):1255-1258. [Medline: 1951216]
43. Devitt P, Palmer E. Computer-aided learning: an overvalued educational resource? Med Educ 1999 Feb;33(2):136-139. [Medline: 10211264]
44. Elves AW, Ahmed M, Abrams P. Computer-assisted learning; experience at the Bristol Urological Institute in the teaching of urology. Br J Urol 1997 Nov;80 Suppl 3:59-62. [Medline: 9415087]
45. Fasce E, Ramírez L, Ibáñez P. [Evaluation of a computer-based independent study program applied to fourth year medical students]. Rev Med Chil 1995 Jun;123(6):700-705. [Medline: 8525222]
46. Finley JP, Sharratt GP, Nanton MA, Chen RP, Roy DL, Paterson G. Auscultation of the heart: a trial of classroom teaching versus computer-based independent learning. Med Educ 1998 Jul;32(4):357-361. [Medline: 9743795]
47. Gelb DJ. Is newer necessarily better?: assessment of a computer tutorial on neuroanatomical localization. Neurology 2001 Feb 13;56(3):421-422. [Medline: 11171920]
48. Green MJ, Levi BH. Teaching advance care planning to medical students with a computer-based decision aid. J Cancer Educ 2011 Mar;26(1):82-91 [FREE Full text] [doi: 10.1007/s13187-010-0146-2] [Medline: 20632222]
49. Hilger AE, Hamrick HJ, Denny FWJ. Computer instruction in learning concepts of streptococcal pharyngitis. Arch Pediatr Adolesc Med 1996 Jun;150(6):629-631. [Medline: 8646314]
50. Holt RI, Miklaszewicz P, Cranston IC, Russell-Jones D, Rees PJ, Sönksen PH. Computer assisted learning is an effective way of teaching endocrinology. Clin Endocrinol (Oxf) 2001 Oct;55(4):537-542. [Medline: 11678838]
51. Lee CSC, Rutecki GW, Whittier FC, Clarett MR, Jarjoura D. A comparison of interactive computerized medical education software with a more traditional teaching format. Teaching and Learning in Medicine 1997 Jan;9(2):111-115. [doi: 10.1080/10401339709539824]
52. MacFadyen JC, Brown JE, Schoenwald R, Feldman RD. The effectiveness of teaching clinical pharmacokinetics by computer. Clin Pharmacol Ther 1993 Jun;53(6):617-621. [Medline: 8513652]
53. Mangione S, Nieman LZ, Greenspon LW, Margulies H. A comparison of computer-assisted instruction and small-group teaching of cardiac auscultation to medical students. Med Educ 1991 Sep;25(5):389-395. [Medline: 1758315]
54. McDonough M, Marks IM. Teaching medical students exposure therapy for phobia/panic - randomized, controlled comparison of face-to-face tutorial in small groups vs. solo computer instruction. Med Educ 2002 May;36(5):412-417. [Medline: 12028390]
55. Mojtahedzadeh R, Mohammadi A, Emami AH, Rahmani S. Comparing live lecture, internet-based & computer-based instruction: A randomized controlled trial. Med J Islam Repub Iran 2014;28:136 [FREE Full text] [Medline: 25694994]
56. Nola M, Morović A, Dotlić S, Dominis M, Jukić S, Damjanov I. Croatian implementation of a computer-based teaching program from the University of Kansas, USA. Croat Med J 2005 Jun;46(3):343-347 [FREE Full text] [Medline: 15861510]
57. Perfeito JA, Forte V, Giudici R, Succi JE, Lee JM, Sigulem D. [Development and assessment of a multimedia computer program to teach pleural drainage techniques]. J Bras Pneumol 2008 Jul;34(7):437-444 [FREE Full text] [Medline: 18695787]
58. Pusic MV, Leblanc VR, Miller SZ. Linear versus web-style layout of computer tutorials for medical student learning of radiograph interpretation. Acad Radiol 2007 Jul;14(7):877-889. [doi: 10.1016/j.acra.2007.04.013] [Medline: 17574137]
59. Ram SP, Phua KK, Ang BS. The effectiveness of a computer-aided instruction courseware developed using interactive multimedia concepts for teaching Phase III MD students. Medical Teacher 2009 Jul 03;19(1):51-52. [doi: 10.3109/01421599709019348]
60. Santer DM, Michaelsen VE, Erkonen WE, Winter RJ, Woodhead JC, Gilmer JS, et al. A comparison of educational interventions. Multimedia textbook, standard lecture, and printed textbook. Arch Pediatr Adolesc Med 1995 Mar;149(3):297-302. [Medline: 7532074]

XSL•FO
RenderX

61. Seabra D, Srougi M, Baptista R, Nesrallah LJ, Ortiz V, Sigulem D. Computer aided learning versus standard lecture for undergraduate education in urology. J Urol 2004 Mar;171(3):1220-1222. [doi: 10.1097/01.ju.0000114303.17198.37] [Medline: 14767306]
62. Shomaker TS, Ricks DJ, Hale DC. A prospective, randomized controlled study of computer-assisted learning in parasitology. Acad Med 2002 May;77(5):446-449. [Medline: 12010707]
63. Solomon DJ, Ferenchick GS, Laird-Fick HS, Kavanaugh K. A randomized trial comparing digital and live lecture formats [ISRCTN40455708. BMC Med Educ 2004 Nov 29;4:27 [FREE Full text] [doi: 10.1186/1472-6920-4-27] [Medline: 15569389]
64. Stanford W, Erkonen WE, Cassell MD, Moran BD, Easley G, Carris RL, et al. Evaluation of a computer-based program for teaching cardiac anatomy. Invest Radiol 1994 Feb;29(2):248-252. [Medline: 8169104]
65. Summers A, Rinehart GC, Simpson D, Redlich PN. Acquisition of surgical skills: a randomized trial of didactic, videotape, and computer-based training. Surgery 1999 Aug;126(2):330-336. [Medline: 10455902]
66. Taveira-Gomes T, Prado-Costa R, Severo M, Ferreira MA. Characterization of medical students recall of factual knowledge using learning objects and repeated testing in a novel e-learning system. BMC Med Educ 2015 Jan 24;15:4 [FREE Full text] [doi: 10.1186/s12909-014-0275-0] [Medline: 25616353]
67. Vichitvejpaisal P, Sitthikongsak S, Preechakoon B, Kraiprasit K, Parakkamodom S, Manon C, et al. Does computer-assisted instruction really help to improve the learning process? Med Educ 2001 Oct;35(10):983-989. [Medline: 11564203]
68. Vivekananda-Schmidt P, Lewis M, Hassell AB, ARC Virtual Rheumatology CAL Research Group. Cluster randomized controlled trial of the impact of a computer-assisted learning package on the learning of musculoskeletal examination skills by undergraduate medical students. Arthritis Rheum 2005 Oct 15;53(5):764-771 [FREE Full text] [doi: 10.1002/art.21438] [Medline: 16208642]
69. Vivekananda-Schmidt P, Lewis M, Hassell AB, ARC Virtual Rheumatology CAL Research Group. Cluster randomized controlled trial of the impact of a computer-assisted learning package on the learning of musculoskeletal examination skills by undergraduate medical students. Arthritis Rheum 2005 Oct 15;53(5):764-771 [FREE Full text] [doi: 10.1002/art.21438] [Medline: 16208642]

Abbreviations

DVD: digital versatile disc
USB: universal serial bus
RCT: randomized control trial
cRCT: cluster randomized control trial
SMD: standardized mean difference
GRADE: Grading of Recommendations, Assessment, Development, and Evaluations

Edited by A Marusic; submitted 22.01.19; peer-reviewed by D Adeloye, D Rotimi, B Benzon; comments to author 08.02.19; revised version received 19.02.19; accepted 21.02.19; published 25.03.19.

Please cite as:
Kyaw BM, Posadzki P, Dunleavy G, Semwal M, Divakar U, Hervatis V, Tudor Car L
Offline Digital Education for Medical Students: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration
J Med Internet Res 2019;21(3):e13165
URL: http://www.jmir.org/2019/3/e13165/
doi:10.2196/13165
PMID:30907731

©Bhone Myint Kyaw, Pawel Posadzki, Gerard Dunleavy, Monika Semwal, Ushashree Divakar, Vasilis Hervatis, Lorainne Tudor Car. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 25.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Serious Gaming and Gamification Education in Health Professions: Systematic Review

Sarah Victoria Gentry[1,2], BMBS, MPhil; Andrea Gauthier[3], PhD; Beatrice L'Estrade Ehrstrom[4], MD; David Wortley[5], FRSA; Anneliese Lilienthal[4], MS; Lorainne Tudor Car[6], MD, PhD; Shoko Dauwels-Okutsu[7], PhD; Charoula K Nikolaou[8], PhD; Nabil Zary[4,9,10], MD, PhD; James Campbell[11], MPH, MSc; Josip Car[7], MD, PhD

[1]Norwich Medical School, University of East Anglia, Norwich, United Kingdom
[2]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom
[3]Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[4]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden
[5]Royal Society of Arts, London, United Kingdom
[6]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[7]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[8]Centre de Philosophie du Droit (Cellule Biogov), Universite Catholique de Louvain, Louvain-la-Neuve, Belgium
[9]Games for Health Innovations Centre, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[10]Emerging Technologies Lab, Mohammed VI University of Health Sciences, Casablanca, Morocco
[11]Health Workforce Department, World Health Organization, Geneva, Switzerland

Corresponding Author:
Sarah Victoria Gentry, BMBS, MPhil
Norwich Medical School
University of East Anglia
Norwich Research Park
Norwich, NR4 7TJ
United Kingdom
Phone: 44 1603 456161
Email: sarah.gentry@doctors.org.uk

Abstract

Background: There is a worldwide shortage of health workers, and this issue requires innovative education solutions. Serious gaming and gamification education have the potential to provide a quality, cost-effective, novel approach that is flexible, portable, and enjoyable and allow interaction with tutors and peers.

Objective: The aim of this systematic review was to evaluate the effectiveness of serious gaming/gamification for health professions education compared with traditional learning, other types of digital education, or other serious gaming/gamification interventions in terms of patient outcomes, knowledge, skills, professional attitudes, and satisfaction (primary outcomes) as well as economic outcomes of education and adverse events (secondary outcomes).

Methods: A comprehensive search of MEDLINE, EMBASE, Web of Knowledge, Educational Resources Information Centre, Cochrane Central Register of Controlled Trials, PsycINFO, and Cumulative Index to Nursing and Allied Health Literature was conducted from 1990 to August 2017. Randomized controlled trials (RCTs) and cluster RCTs were eligible for inclusion. Two reviewers independently searched, screened, and assessed the study quality and extracted data. A meta-analysis was not deemed appropriate due to the heterogeneity of populations, interventions, comparisons, and outcomes. Therefore, a narrative synthesis is presented.

Results: A total of 27 RCTs and 3 cluster RCTs with 3634 participants were included. Two studies evaluated gamification interventions, and the remaining evaluated serious gaming interventions. One study reported a small statistically significant difference between serious gaming and digital education of primary care physicians in the time to control blood pressure in a subgroup of their patients already taking antihypertensive medications. There was evidence of a moderate-to-large magnitude of effect from five studies evaluating individually delivered interventions for objectively measured knowledge compared with traditional learning. There was also evidence of a small-to-large magnitude of effect from 10 studies for improved skills compared with traditional learning. Two and four studies suggested equivalence between interventions and controls for knowledge and

XSL•FO
RenderX

skills, respectively. Evidence suggested that serious gaming was at least as effective as other digital education modalities for these outcomes. There was insufficient evidence to conclude whether one type of serious gaming/gamification intervention is more effective than any other. There was limited evidence for the effects of serious gaming/gamification on professional attitudes. Serious gaming/gamification may improve satisfaction, but the evidence was limited. Evidence was of low or very low quality for all outcomes. Quality of evidence was downgraded due to the imprecision, inconsistency, and limitations of the study.

Conclusions: Serious gaming/gamification appears to be at least as effective as controls, and in many studies, more effective for improving knowledge, skills, and satisfaction. However, the available evidence is mostly of low quality and calls for further rigorous, theory-driven research.

(J Med Internet Res 2019;21(3):e12994) doi:10.2196/12994

KEYWORDS

video games; education, professional; review

Introduction

Innovative approaches and modalities for education in health professions education are constantly sought to improve teaching and learning and ultimately patient care and outcomes. Digital education may be one such innovation. This review focuses on serious gaming and gamification education.

For the purposes of this review, we have used the terminology defined by Alvarez [1,2]. The term “serious game” was used to refer only to games designed specifically for the “serious” purpose of providing health professions education via a digital device. The term “serious diverting” was used to refer to the use of games originally designed primarily for entertainment used without modification, as part of health professions education delivered via a digital device. “Serious gaming” was used to refer to any use of digital games for health professions education, thereby encompassing “serious games” and “serious diverting.”

A related but separate concept—“gamification”—can be defined as “the application of the characteristics and benefits of games to real world processes or problems” [3]. Gamification differs from serious games in terms of the design intention, with gamification interventions involving the application of game elements with a utilitarian purpose and serious games designed as full-fledged games for a purpose other than just entertainment [4]. Wortley suggests that both may be experienced by the user as a complete game, although typically, gamification involves the use of game components outside a game setting, such as rewarding users completing an electronic learning (e-learning) module with badges or points. Gamification has the potential to allow for greater involvement of the user in setting his/her own objectives or outcomes, personalization of the intervention, and cost-effectiveness [3]. Most, but not all, uses of the term refer to interventions involving the use of enabling digital technologies.

Serious gaming/gamification has the potential to provide learners with opportunities to be part of active learning, solving clinical problems, and gaining experience in risk-free surroundings [5], without needing to involve patients. Learners may have the opportunity to develop analytical skills, strategic thinking, knowledge, multitasking, decision making, communication, and psychomotor skills [6], with multiplayer functions providing opportunities for collaborative learning [7]. The motivational properties of gaming have the potential to be harnessed for educational purposes [8]. Serious gaming/gamification can be used at a time and place that suits the learner. The reusable nature of serious gaming/gamification may allow more frequent or longer interactions, free up lecturer time, and provide monetary savings [9]. However, this could lead to reduced opportunities to ask questions, hold discussions, and spend time with patients. Use of such interventions within small groups, with lecturer support, could allow for discussion and interaction but would likely increase lecturer time needed as compared to traditional learning. Serious gaming/gamification, like other kinds of e-learning, may ease the process of updating materials, as modifications to content can be made continuously, unlike with a text book.

Although serious gaming and gamification interventions appear to have much potential, rigorous evaluation is required to assess whether they can lead to effective learning. There is a potential for the game or game elements to become a distraction rather than a facilitator of learning, with the method “more memorable than the message” [10]; therefore, the quality of learning must be the focus, as opposed to the capabilities of the technology used [11].

This systematic review is one of a series of reviews evaluating the scope for implementation and potential impact of a wide range of digital health education interventions for pre- and postregistration health professions. This review was conducted in collaboration with the World Health Organization’s Health Workforce Department. The objective of this work is to compare the effectiveness of serious gaming and gamification education versus various controls in improving learners’ knowledge, skills, professional attitudes, and satisfaction as well as patient outcomes.

Methods

While conducting and reporting the review, we adhered to the gold-standard systematic review methods recommended by the Cochrane Collaboration [12]. For a detailed description of the methodology, please refer to our previous paper [13].

Search Strategy and Data Sources

We comprehensively searched the following databases between 1990 and August 2017: MEDLINE (Ovid), EMBASE (Elsevier), Web of Science, Educational Resource Information Centre

XSL•FO
RenderX

(Ovid), Cochrane Central Register of Controlled Trials (CENTRAL), (The Cochrane Library), PsycINFO (Ovid), and Cumulative Index to Nursing and Allied Health Literature (EBSCO). The search strategy for MEDLINE is presented in the Multimedia Appendix 1. We searched for papers in English but considered eligible studies in any language. We also searched two trials registries, reference lists of all included studies, and relevant systematic reviews and contacted the relevant investigators for further information.

Eligibility Criteria

We included randomized controlled trials (RCTs) and cluster RCTs (cRCTs) of pre- and postregistration health professions using serious gaming/gamification with any type of controls (traditional learning, digital education, or another type of serious gaming/gamification intervention), which measured patient outcomes, knowledge, skills (cognitive and psychomotor), professional attitudes, and satisfaction (primary outcomes) and adverse effects or costs (secondary outcomes). We excluded crossover trials due to the high likelihood of carry-over effect. Participants were not excluded on the basis of sociodemographic characteristics such as age, gender, ethnicity, or any other related characteristics. Outcome definitions are available in the associated paper [13].

Data Selection, Extraction, and Management

The search results from different electronic databases were combined in a single EndNote library (X 8.2; Clarivate Analytics, Philadelphia, PA), and duplicate records were removed. Two reviewers independently screened titles and abstracts to identify studies that potentially met the inclusion criteria. The full texts of these articles were retrieved and read. Two review authors independently assessed these articles against the eligibility criteria (SG, AG, and BE). At least two reviewers independently extracted the data for each of the included studies using a structured data-extraction form. We extracted all relevant data on the characteristics of participants, interventions, controls, and outcomes measures. For continuous data, we reported standardized mean differences and SDs. None of the studies reported dichotomous data. Disagreements were resolved through discussion.

Assessment of Risk of Bias

Two reviewers independently assessed the risk of bias of the included studies using the Cochrane Collaboration's "Risk of bias" tool [12]. Studies were assessed for the risk of bias in the following domains: random sequence generation, allocation concealment, blinding (participants and personnel), blinding (outcome assessment), completeness of outcome data (attrition bias), selective outcome reporting (reporting bias), and other risk of bias. For cRCTs, we also assessed recruitment bias, baseline imbalances, loss of clusters, and incorrect analysis. Judgements concerning the risk of bias for each study were classified as high, low, or unclear.

Data Synthesis

Data were synthesized using Review Manager (Version 5.3; The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark). Included studies were insufficiently homogenous in terms of population, inclusion criteria, interventions, and outcomes for meta-analysis. The decision not to perform a meta-analysis was made by a consensus of review authors. We present a narrative synthesis of findings, with effect sizes calculated for outcomes where there were sufficient data. Where possible, we assessed the quality of studies and size of effect. Results are presented by outcome and separately for each comparison (serious gaming/gamification vs traditional learning, serious gaming/gamification vs digital health education, and serious gaming/gamification vs serious gaming/gamification).

Assessment of Evidence Quality

The results for comparisons between serious gaming/gamification and traditional learning as well as serious gaming/gamification and digital education are presented in the narrative summary of findings tables (Tables 7 and 8). Two authors (SG and AG) rated the overall quality of the evidence as implemented and described in GRADEprofiler (GRADEproGDT online version; Evidence Prime, Inc, Hamilton, ON, Canada) and chapter 11 of the Cochrane Handbook for Systematic Reviews of Interventions [12]. We considered the following criteria to assess the quality of the evidence: limitations of studies (risk of bias), inconsistency of results, indirectness of the evidence, imprecision, and publication bias. We also downgraded the quality, where appropriate. This was done for all primary outcomes reported in the review.

Results

Our searches yielded a total of 30,532 citations and 30 studies (27 RCTs and 3 cRCTs) including 3634 participants (Figure 1).

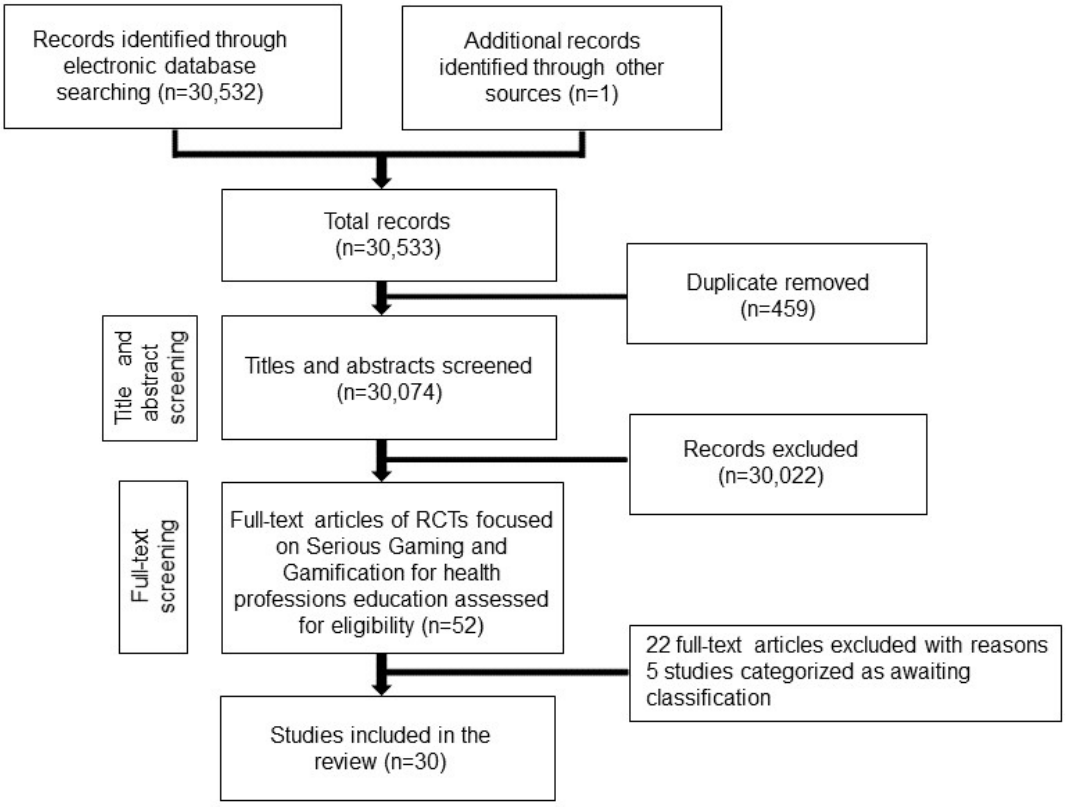
Included Studies

Study Designs and Populations

Sample sizes ranged from 14 [14] to 1470 [15] participants. Almost half the included studies had sample sizes below 50.

XSL•FO
RenderX

Figure 1. PRISMA flow chart. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT: randomized controlled trial.



Fourteen studies were conducted in Europe [16-30], and 11 studies were conducted in North America [14,15,31-40], one of which recruited participants from 63 countries via the internet [15]. One study was conducted in Singapore [41]. Four studies were conducted in middle-income countries, three of which were conducted in Brazil [42-44]. One study was conducted by authors based in China and Taiwan, but it was unclear where the study itself was carried out [45]. None of the included studies were conducted in low-income countries. Details of study designs and population for each trial are compared in Tables 1-3 and a summary is given below.

Eleven studies included only medical students [14,16,17,19,20,24-26,39,40,42,45]. Five studies included only nursing students [22,27,29,37,41] and four included only surgical residents [18,30,31,38]. The remaining studies included primary care doctors (n=2) [36,43]; dental students (n=2) [21,32]; anesthesiology residents (n=1) [35]; urologists (n=1) [15]; speech and language science students (n=1) [44]; participants of the Major Incident Medical Management and Support course, which typically includes doctors, nurses, and paramedics with an interest in prehospital care (n=1) [23]; nursing and medical students (n=1) [28]; and medical students, residents, and specialists in Obstetrics and Gynecology (n=1, reported in one article and one conference abstract) [33,34].

XSL•FO
RenderX

Table 1. Study designs and populations of the included studies comparing serious gaming/gamification and traditional learning.

Study	Study type	Population (n)	Country	Field of study
Adams et al 2012 [31]	RCT[a]	Surgical residents (31)	United States	General surgery
Boada et al 2015 [27]	RCT	Nursing students (109)	Spain	Cardiopulmonary resuscitation skills
Boeker et al 2013 [20]	RCT	Medical students (145)	Germany	Urology
Cook 2012 et al [22]	RCT	Nursing students (34)	United Kingdom	Intermediate life support
De Araujo et al 2016 [42]	RCT	Medical students (20)	Brazil	Surgical skills
Del Blanco et al 2017 [28]	RCT	Nursing and medical students (132)	Spain	Preparation for going into the operating theatre
Diehl et al 2017 [43]	RCT	Primary care physician (134)	Brazil	Insulin management in primary care
Foss et al 2014 [29]	RCT	Nursing students (201)	Norway	Medication calculation
Giannotti et al 2013 [30]	RCT	Surgical residents (42)	Italy	Surgical skills
Graafland et al 2017 [18]	RCT	Surgical residents (31)	Netherlands	Minimally invasive surgery
Hannig et al 2013 [21]	RCT	Dental students (55)	Germany	Alginate mixing skills
Katz et al 2017 [35]	RCT	Anesthesiology residents (44)	United States	Liver transplant anesthesiology
Knight et al 2010 [23]	cRCT[b]	Health professionals on a Major Incident Management Course (91, 2 clusters)	United Kingdom	Major incident management and support
Lagro et al 2014 [19]	cRCT	Medical students (145, 5 clusters)	Netherlands	Geriatrics
LeFlore et al 2012 [37]	RCT	Nursing students (106)	United States	Pediatric respiratory disease
Li et al 2015 [45]	RCT	Medical students (97)	China/Taiwan	Cardiopulmonary resuscitation
Plerhoples et al 2011 [38]	RCT	Surgical residents (40)	United States	Surgical skills
Rondon et al 2013 [44]	RCT	Speech-language and hearing science students (29)	Brazil	Anatomy and physiology
Tan et al 2016 [41]	cRCT	Nursing students (103, 7 clusters)	Singapore	Blood transfusion administration

[a]RCT: randomized controlled trial.

[b]cRCT: cluster randomized controlled trial.

Table 2. Study designs and populations of the included studies comparing serious gaming/gamification and other digital education interventions.

Study	Study type	Population (n)	Country	Field of study
Amer et al 2011 [32]	RCT[a]	Dental students (80)	United States	Operative dentistry
Chien et al 2013 [14]	RCT	Medical students (14)	United States	Laparoscopic surgical tasks
Dankbaar et al 2016 [16]	RCT	Medical students (79)	Netherlands	Approach to acutely unwell patients
Dankbaar et al 2017 [17]	RCT	Medical students (66)	Netherlands	Patient safety and stress management
Gauthier et al 2015 [40]	RCT	Medical students (44)	Canada	Vascular anatomy
Kerfoot et al 2014 [36]	RCT	Primary care physician (111)	United States	Management of blood pressure in primary care
Sward et al 2008 [39]	RCT	Medical students (100)	United States	Pediatrics

[a]RCT: randomized controlled trial.

XSL•FO
RenderX

Table 3. Study designs and populations of included studies comparing serious gaming/gamification and another type of serious gaming/gamification intervention.

Study	Study type	Population	Country	Field of study
Adams et al 2012 [31]	RCT[a]	Surgical residents (31)	United States	General surgery
De Araujo et al 2016 [42]	RCT	Medical students (20)	Brazil	Surgical skills
Hedman et al 2013 [24] and Kolga et al	RCT	Medical students (30)	Sweden	Surgical skills
Ju et al 2011 [34] and Ju et al 2012 [33]	RCT	Medical students, residents and attendings (42)	United States	Surgical skills
Kerfoot et al 2012 [15]	RCT	Urologists (1470)	United States (participants recruited online from 63 countries)	Urology guideline knowledge
Kolga et al 2008 [26]	RCT	Medical students (22)	Sweden	Surgical skills

[a]RCT: randomized controlled trial.

Interventions

Characteristics of the interventions included are compared in Tables 4-6 and a summary is given below.

Two of the included studies evaluated "gamification" interventions [15,36]. The remainder evaluated serious gaming interventions. Two of these were group interventions, in which a serious gaming intervention was projected to a traditional classroom of learners who played together [39,44].

A total of 22 interventions had original design purposes other than entertainment, of which 11 interventions were classified as "Message Broadcasting - Educative" [15,17,20,28,36,37,39-41,43,44]; four were classified as "Training - Mental" [19,22,23,29], three were classified as "Training - Physical" [18,21,32], and four were classified as both "Training - Mental" and "Training - Physical" [16,27,35,45]. All of the interventions with design purposes other than entertainment were classified under "Education" for "Scope."

The remaining eight interventions were commercial off-the-shelf games designed only for the purpose of entertainment but used for training of motor skills as part of "Serious Diverting" interventions. These were all classified as "Training - Physical" for "Purpose" and as "Entertainment" and "General Public" for "Scope" [14,24-26,30,31,33,34,38,42].

XSL•FO
RenderX

Table 4. Characteristics of included interventions in studies comparing serious gaming/gamification and traditional learning.

Study	Intervention type	Intervention duration	Intervention frequency	Intervention intensity	Control
Adams et al 2012 [31]	First-person shooter, commercial-off the-shelf intervention	6 weeks	Weekly	Mean of 5.7 (SD 1.3) hours	Box trainer
Boada et al 2015 [27]	Life support–simulation activities	Access for 1 week	__[a]	All did >50% of the tasks	Usual education
Boeker et al 2013 [20]	Electronic adventure game "Uro-Island"	Access for 1 week	—	—	Written script
Cook et al 2012 [22]	Platform for undergraduate life support education game	2 weeks	—	Unlimited access	Usual learning
De Araujo et al 2016 [42]	Surgical commercial-off-the shelf intervention (SurgG)	Access for 3 weeks	—	Mean of 647 minutes per week	Usual learning (ContG)
Del Blanco et al 2017 [28]	Videogaming intervention	Access for 1 day	Once	Variable	Usual learning
Diehl et al 2017 [43]	"InsuOnline" game	Access for 21 days	—	Mean of 4 hours	Onsite learning activity
Foss et al 2014 [29]	"The Medication Game" online training	Access for 4.5 weeks	—	—	Standard education
Giannotti et al 2013 [30]	Nintendo Wii training	4 weeks	5 days per week	60 minutes	Usual training
Graafland et al 2017 [18]	Game enhanced curriculum (Dr Game, Surgeon Trouble)	—	Two sessions	30 minutes	Usual training
Hannig et al 2013 [21]	Skills-O-Mat interactive game	60 minutes	Once	—	Teacher-catered workshop
Katz et al 2017 [35]	"OCT trainer" game where players work through the steps in liver transplant anesthesiology	30 days	81% self- reported playing 1-3 times per week	—	Usual training
Knight et al 2010 [23]	"Triage Trainer" computer game	60 minutes	Once	—	Card-sorting exercise
Lagro et al 2014 [19]	Geriatrics game in which players must balance patient-oriented goals and preferences, appropriateness of medical care, and costs	60-90 minutes	Once	—	Standard educational activity
LeFlore et al 2012 [37]	"Virtual Patient Trainer" game	2-3 hours	Once	—	Traditional lecture
Li et al 2015 [45]	3D cardiopulmonary resuscitation game	3 months (with 2-week extension possible)	—	—	Reminders to refresh their skills sent frequently
Plerhoples et al 2011 [38]	Commercial off-the-shelf intervention	10 minutes	Once	—	Standard educational activity
Rondon et al 2013 [44]	Computer game-based learning played as a group on a projector	9 weeks	Once per week	1 hour	Short scientific texts
Tan et al 2016 [41]	Videogame simulating blood transfusion–administration challenges and minigames	30 minutes	Once	—	Usual education

[a]Not available.

XSL•FO
RenderX

Table 5. Characteristics of included interventions in studies comparing serious gaming/gamification and other digital education interventions.

Study	Intervention type	Intervention duration	Intervention frequency	Intervention intensity	Control
Amer et al 2011 [32]	Interactive dental videogame	Up to 20 minutes	Once	__[a]	3-minute video on resin bonding
Chien et al 2013 [14]	3D tennis game	40 minutes	Once	—	Virtual simulator training platform
Dankbaar et al 2016 [16]	Computer-based simulation game "abcdeSIM"	Access for 4 weeks	—	Estimated to take 2-4 hours to complete; mean logged game time 90 (SD 49) minutes	Electronic module
Dankbaar et al 2017 [17]	"Air-Medic Sky-1" game	1 week	—	3-4 hours	Digital education module
Gauthier et al 2015 [40]	"Vascular Invaders" game	Access for 35 days	—	—	Vascular anatomy study aid (online)
Kerfoot et al 2014 [36]	Online spaced-education game (question emailed every 3 days; resent 12 or 24 days later if answered incorrectly or correctly, respectively; retired after answered correctly on >two consecutive attempts)	Access for 52 weeks	—	Mean of 38 (SD 7) weeks to complete the cycle of questions	Online posting
Sward et al 2008 [39]	Web-based pediatric board game	4 weeks	One per week	1 hour	Self-study Web flash cards

[a]Not available.

Table 6. Characteristics of included interventions in studies comparing serious gaming/gamification and another type of serious gaming/gamification intervention.

Study	Intervention type	Intervention duration	Intervention frequency	Intervention intensity	Control
Adams et al 2012 [31]	FPS[a] COTS[b] intervention	6 weeks	Weekly	Mean of 5.7 (SD 1.3) hours	Non-FPS COTS intervention
De Araujo et al 2016 [42]	Surgical COTS intervention (SurgG)	Access for 3 weeks	__[c]	Mean of 647 minutes per week	Usual learning (ContG), FPS COTS (ShotG), Racing COTS (RaceG) interventions
Hedman et al 2013 [24] and Kolga et al 2009 [25]	Systematic video game training with FPS COTS intervention	5 weeks	5 days per week	30-60 minutes	Non-FPS COTS intervention
Ju 2011 et al [34]) and Ju et al 2012 [33]	Wii COTS intervention	30 minutes	Once	—	Play Station 2 COTS intervention
Kerfoot et al 2012 [15]	Online spaced-education game - 4 questions every 4 days	8-42 days	—	—	Spaced-education game – 2 questions every 2 days
Kolga et al 2008 [26]	FPS COTS intervention	5 weeks	5 days per week	30 minutes	2D non-FPS COTS intervention

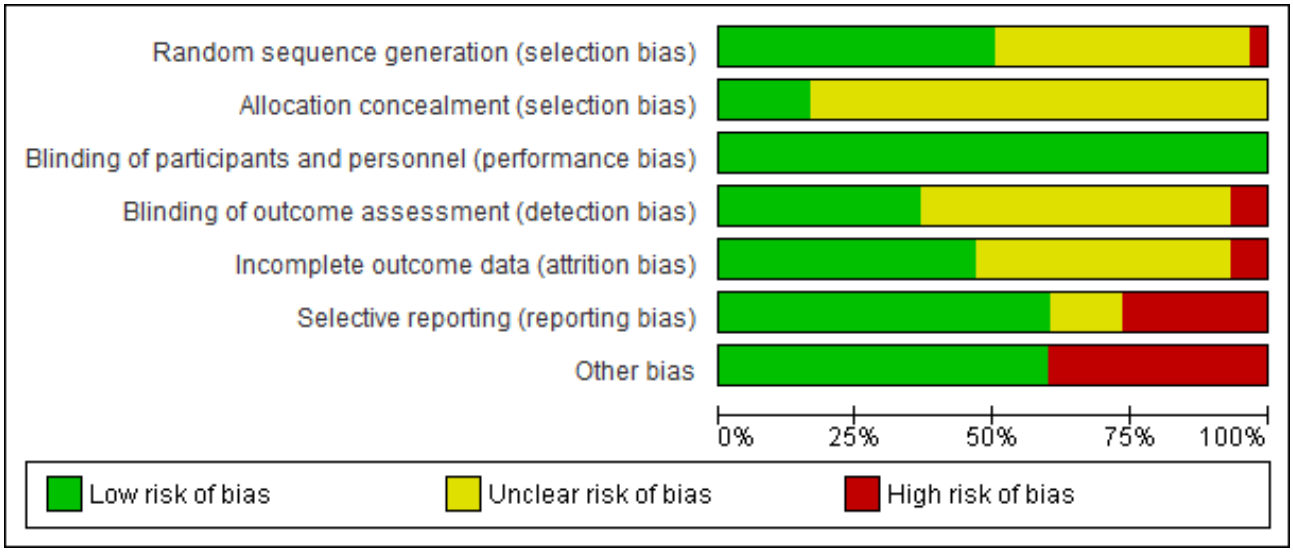
[a]FPS: first-person shooter.

[b]COTS: commercial off the shelf.

[c]Not available.

XSL•FO
RenderX

Figure 2. Risk-of-bias graph.



Comparisons and Outcomes

Serious gaming/gamification was compared with traditional learning in 19 studies [18-23,27-31,35,37,38,41-45], with digital education in 7 studies [14,16,17,32,36,39,40], and with other serious gaming/gamification interventions in 6 studies [15,24-26,31,33,42].

One study addressed patient outcomes [36]. Fourteen studies assessed knowledge [15,17,19,20,28,32,36,37,39-41,43,44]. Twenty-three studies addressed outcomes relating to skills [14,16-19,21-35,37-39,41,42,45]. Four studies assessed outcomes related to attitudes [17,24,25,28,43]. Sixteen studies addressed participant satisfaction [15-17,19-22,26,27,29,32,37,39,42,43,45].

Risk of Bias in Included Studies

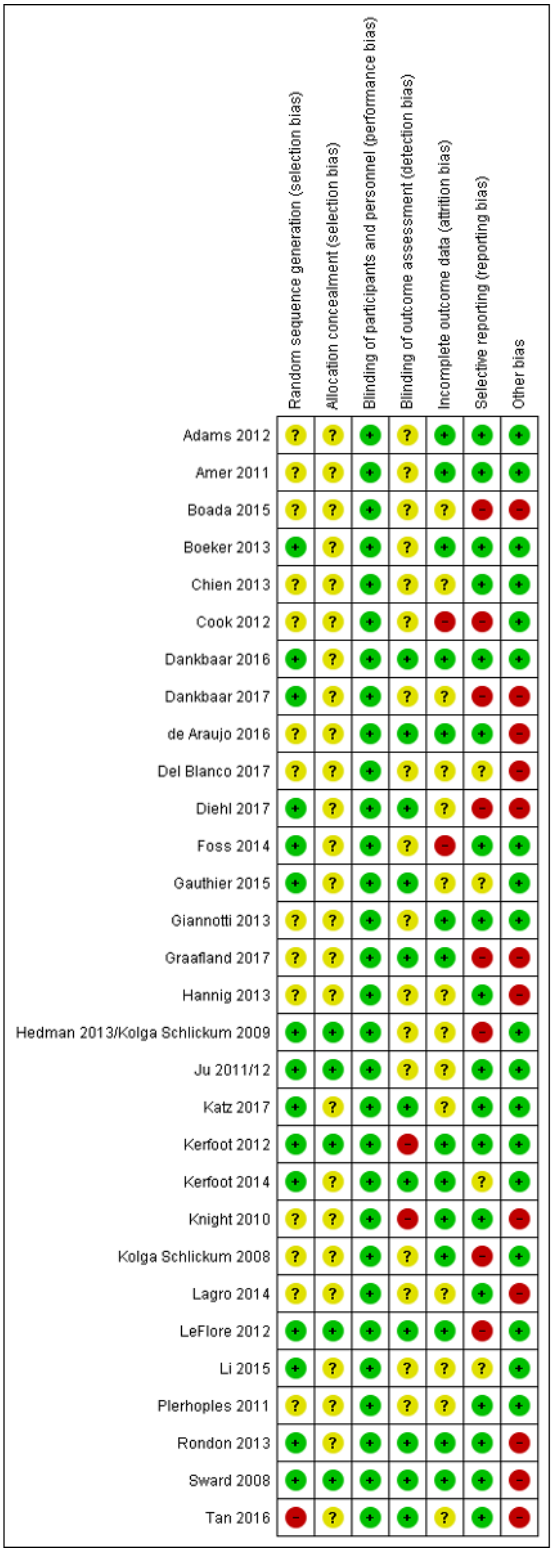
Figures 2 and 3 summarize the risk-of-bias assessments for the included studies. A total of 25 of the included studies were considered to be at high risk of bias [14,16-23,26-32,35,36,38,40-45] according to Cochrane guidelines, because they had a high or unclear risk of bias for either the sequence generation or allocation concealment domains [12]. All three cRCTs were rated high for incorrect analysis, as none accounted for clustering in the analysis.

Effects of Interventions

Effects of the interventions are compared in Multimedia Appendices 2-4 and a summary is given below.

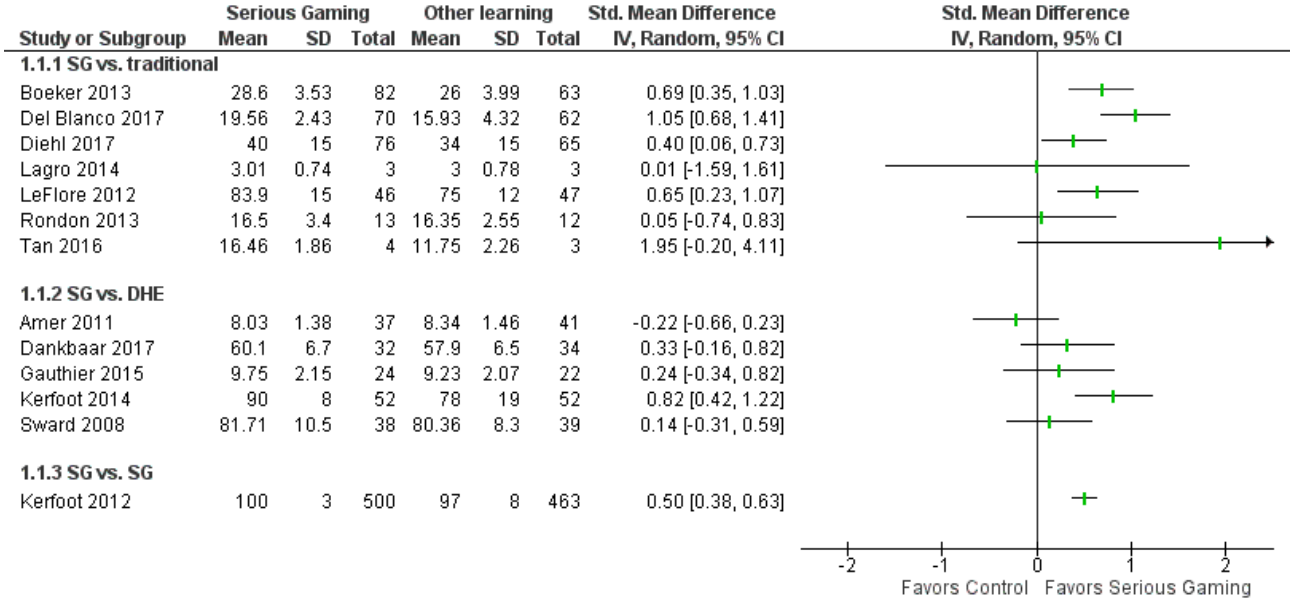
XSL•FO
RenderX

Figure 3. Risk-of-bias summary.



XSL•FO
RenderX

Figure 4. Forest plot for knowledge outcomes. IV: inverse variance; SG: serious games; DHE: digital health education.



Primary Outcomes

Patient-Related Outcomes

One study measured patient-related outcomes [36]. This study compared serious gaming/gamification with an online posting intervention for primary care physicians and reported significantly shorter time to control blood pressure in the intervention group for only a subgroup of participants whose patients were already on antihypertensive medication at the start of the study (P=.02), although this may not be clinically significant (117 vs 125 days). Data were insufficient for calculation of standardized mean differences (SMD). The quality rating assessed using Grading of Recommendations, Assessment, Development and Evaluations (GRADE) was low.

Knowledge

Figure 4 summarizes the results of studies reporting knowledge outcomes.

Serious Gaming/Gamification Versus Traditional Learning

Four RCTs reported higher posttest scores in a serious gaming/gamification group than in a traditional learning group, with a mostly moderate magnitude of effect [20,28,37,43]. Interventions included a videogame to prepare students to enter the operating room [28] (SMD 1.05, 95% CI: 0.68-1.41), a urology educational adventure game [20] (SMD: 0.69, 95% CI: 0.35-1.03), pediatric respiratory disease-assessment game for nurses [37] (SMD: 0.65, 95% CI 0.23-1.07), and the InsuOnline serious game [43] (SMD: 0.40, 95% CI: 0.06-0.73). Comparisons were made between usual learning, written script, and traditional lectures.

One RCT of a group intervention, where speech and language science students played a serious game together in a classroom via a projector, found no difference in anatomy and physiology knowledge compared to a self-study control [44] (SMD: 0.05, 95% CI: –0.74 to 0.83).

Two cRCTs were also included [19,41]. One showed evidence of a large magnitude of effect for a blood transfusion serious game as compared to usual education, although the effect may not have been statistically significant (SMD: 1.95, 95% CI: –0.20 to 4.11) [41]. The second study showed no evidence of effect for a geriatric game compared with standard education, although this study measured perceived knowledge rather than an objective measure (SMD: 0.01, 95% CI: –1.50 to 1.61) [19].

All the individually played games with an objective assessment of knowledge suggested that serious gaming/gamification was superior to traditional learning. The quality rating assessed using GRADE was low for this outcome and comparison (Table 7).

Serious Gaming/Gamification Versus Other Modalities of Digital Education

Five studies found no evidence of a difference. Studies included comparison of serious gaming on dentin bonding and an online lecture control [32] and serious gaming with digital education on patient safety [17] and Web-based vascular anatomy study aids with and without game elements [40]. One study of a group serious gaming intervention found no difference in pediatric knowledge between groups who played a projected board game in teams in a conference room with Web-based pediatric flashcards [39]. One study found that compared with an online educational posting, serious gaming (an online spaced-education game) may improve knowledge (large magnitude of effect) [36] (SMD: 0.82, 95% CI: 0.42-1.22). The quality rating assessed using GRADE was low (Table 8).

XSL•FO
RenderX

Table 7. Summary of findings for serious gaming versus traditional learning. Patient or population: various health professionals, settings: high- and middle-income countries, intervention: serious gaming and gamification, comparison: traditional learning.

Outcomes	Number of participants (number of studies)	Quality of evidence (GRADE[a])	Comments
Knowledge (measures include multiple-choice questions, clinical scenario–based questions, and self-assessment; follow-up mostly immediately after the intervention, longest follow-up of 52 weeks)	769 (7)	Low[b,c,d]	All the individually played games with an objective assessment of knowledge suggested serious gaming/gamification was superior to traditional learning. Four RCTs[e] and one cRCT[f] reported higher postintervention knowledge scores between the serious gaming and control groups, with moderate-to-large effect sizes, although the result for the cRCT may not have been statistically significant[g]. An RCT of a serious gaming intervention reported no difference between groups. A cRCT assessing perceived knowledge reported no difference between groups.
Skills (measures include performance metrics on a simulator, practical examinations, OSCEs[h] and self-evaluation; most studies followed up until immediately after the intervention only)	1195 (14)	Low	Six RCTs reported higher postintervention skill scores on all measures of skills employed in that study in the serious gaming group, with small-to-large effect sizes. A further cRCT suggested higher skill scores of small magnitude but may not have been statistically significant[g]. Three RCTs measured skill outcomes using multiple measures (and no summary measure) and reported higher postintervention scores for some of these measures and no difference for others. Two RCTs and one cRCT reported no difference in postintervention skill scores between groups. One cRCT suggested serious gaming may be inferior to traditional learning, but this result may not have been statistically significant[g].
Attitudes (measured with participant-completed rating scales; follow-up immediately after the test)	369 (3)	Very low[b,c,i,j]	One RCT reported higher postintervention attitude scores in the serious gaming group (small effect size) and one RCT reported no difference between groups. One reported higher scores in the intervention groups, but this result may not have been statistically significant[g].
Satisfaction (3 questions on attitudes toward learning experience measured on a 4-point Likert scale; follow-up immediately after the intervention)	144 (1)	Low	One study reported higher postintervention satisfaction scores in the serious gaming group compared with the control.

[a]GRADE: Grading of Recommendations, Assessment, Development and Evaluations.

[b]Rated down one level for study limitations: The risk of bias was unclear for multiple domains.

[c]Rated down one level for imprecision: All included studies assessing this comparison and outcome had fewer than 400 participants.

[d]Low quality (+ + – –): Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

[e]RCT: randomized controlled trial.

[f]cRCT: cluster randomized controlled trial.

[g]None of the 3 included cRCTs accounted for clustering in their analyses. They were therefore reanalyzed using the number of clusters as the sample sizes and were likely significantly underpowered.

[h]OSCE: objective structured clinical examination.

[i]Rated down one level for inconsistency: There was considerable heterogeneity in the results without a clear explanation.

[j]Very low quality (+ – – –): We are uncertain about the estimate.

XSL•FO
RenderX

Table 8. Summary of findings for serious gaming versus other modalities of digital education. Patient or population: health professionals in education, settings: high-income countries, intervention: serious gaming and gamification, comparison: other modalities of digital education.

Outcomes	Number of participants (number of studies)	Quality of evidence (GRADE[a])	Comments
Patient outcomes (blood pressure)	111 (1)	Low[b,c,d]	One study reported better scores for blood pressure in some subgroups. Effect sizes could not be estimated due to missing data.
Knowledge (measures include multiple-choice questions and clinical scenario–based questions; follow-up mostly immediately after the intervention)	403 (5)	Low	One study reported higher scores in the serious gaming group with a large magnitude of effect. Four studies reported no difference.
Skills (measures include performance metrics on a simulator, practical examinations, OSCEs[e], and self-evaluation; most studies followed up until immediately after the intervention only)	290 (5)	Low	One study reported superior scores in the virtual reality control group compared with the serious gaming intervention group. Two studies reported no difference. Two studies reported insufficient data for calculation of effect sizes.
Attitudes (measured with participant-completed rating scales; follow-up immediately after the test)	66 (1)	Low	One study reported no difference in postintervention attitudes scores between groups.
Satisfaction (measured with participant-completed rating scales; follow-up immediately after the test)	245 (3)	Low	Three studies reported higher satisfaction scores in the serious gaming group than groups of other modalities of digital education.

[a]GRADE: Grading of Recommendations, Assessment, Development, and Evaluations.

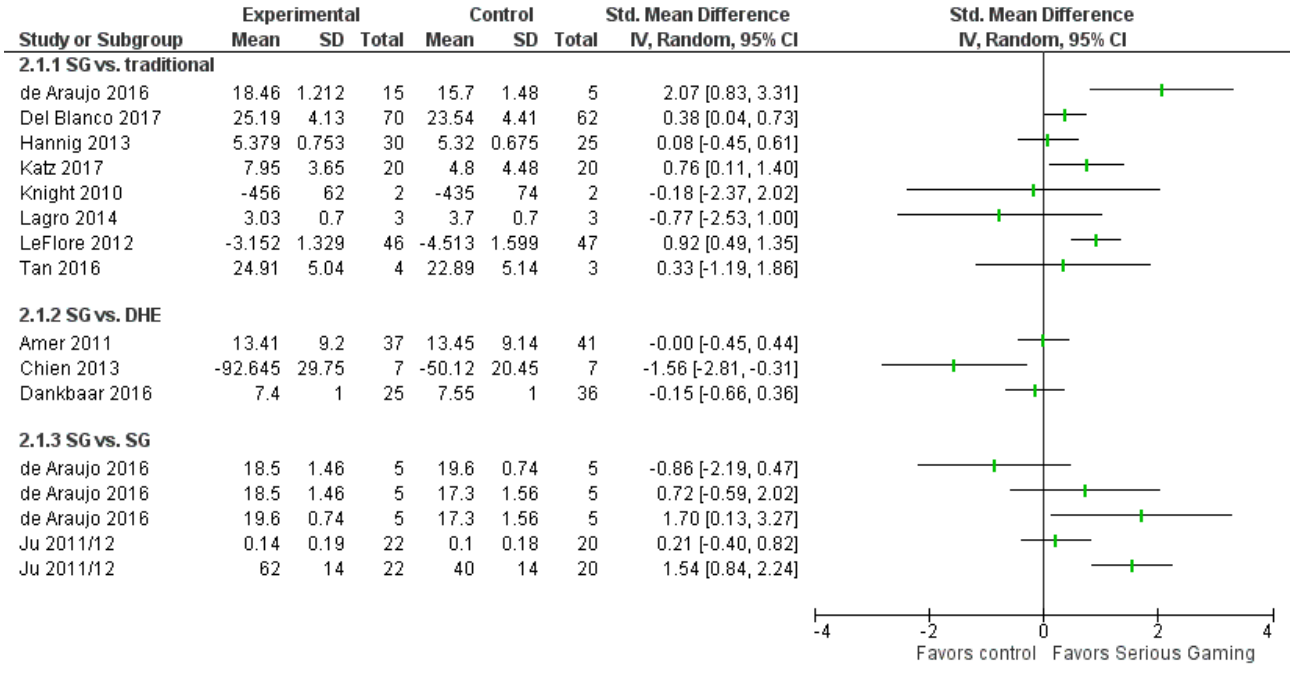
[b]Rated down one level for imprecision: All included studies assessing this comparison and outcome had fewer than 400 participants.

[c]Rated down one level for inconsistency: There was considerable heterogeneity in the results without a clear explanation.

[d]Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

[e]OSCE: objective structured clinical examination.

Figure 5. Forest plot for skills outcomes. IV: inverse variance. SG: serious games; DHE: digital health education.



Serious Gaming/Gamification Versus Serious Gaming/Gamification

One study of a spaced-education game found that interventions with greater question spacing (four questions every 4 days rather than two questions every 2 days) resulted in higher posttest scores, with a moderate magnitude of effect [15] (SMD: 0.50, 95% CI: 0.38-0.64). The quality rating assessed using GRADE was moderate, as the one included study had a low risk of bias in all but one domain.

Skills

A total of 24 studies addressed skill outcomes. Figure 5 summarizes the results of studies reporting skill outcomes.

XSL•FO
RenderX

Serious Gaming/Gamification Versus Traditional Learning

Twelve RCTs [21,22,27-31,35,37,38,42,45] and three cRCTs [19,23,41] compared serious gaming/gamification to traditional learning in this outcome category. The results were inconsistent, and studies were generally of low quality, making it difficult to draw conclusions about the efficacy of these interventions.

Six studies reported significant differences between groups for overall skill assessments in favor of serious gaming [27,28,35,37,42,45], with a magnitude of effect ranging from small to large. However, SMDs for two of these studies could not be calculated due to missing data [27,45]. Interventions included games with scenarios simulating clinical environments [27,28,35,37] and serious diverting interventions for improving practical skills [42].

Three studies comparing serious gaming/gamification with traditional learning used multiple measures for assessing skill outcomes; differences in favor of serious gaming/gamification were observed for some, but not all, of these skill measures, and the studies did not present an overall estimate of the effect [22,30,38]. Effect sizes could not be estimated, as SDs were not reported and attempts to contact the authors for further data were unsuccessful.

Two studies reported no significant difference in skill outcomes when comparing serious gaming/gamification and traditional learning and another reported no differences in pre- and posttest scores in either group [31].

Three cRCTs were also included [19,23,41]. One showed evidence of an effect of small magnitude, favoring a blood transfusion game group [41] (SMD: 0.33, 95% CI: –1.19 to 1.86); the second study found evidence of a moderate magnitude of effect, favoring the standard educational activity group, although skill measures were self-perceived as opposed to objective [19] (SMD: –0.77, 95% CI: –2.53 to 1.00); and the third showed no evidence of effect for a triage trainer game [23] (SMD: –0.18, 95% CI: –2.37 to 2.02). Each of these results may not be statistically significant.

There is some evidence that serious gaming/gamification interventions are more effective for improving skills than traditional learning. The quality rating assessed using GRADE was low, as the risk of bias was unclear for multiple domains and all the included studies had fewer than 400 participants.

Serious Gaming/Gamification Versus Other Modalities Of Digital Education

Five studies comparing skill outcomes for serious gaming/gamification and other modalities of digital education found no evidence of a difference in outcomes between groups [16,17,22,31,32]. In these studies, serious gaming was compared with an online video on dentin bonding [32] and with an electronic module (e-module) on patient safety [17] and management of an acutely unwell patient [16]. Another study reported higher postintervention skill score in a virtual reality control group than a commercial off-the-shelf intervention, with a large magnitude of effect for the time taken to complete surgical skill tasks (peg transfer and bimanual carrying; SMD: –1.56, 95% CI: –0.31 to –2.81), but reported no difference for distance travelled with surgical instruments when completing these tasks [14]. The quality rating assessed using GRADE was low.

Serious Gaming/Gamification Versus Serious Gaming/Gamification

We are uncertain whether any particular type of serious gaming/gamification is more effective than the other for improving skills. In three of the five studies comparing two serious gaming/gamification interventions, games involving motor skills, visuospatial skills, and manual dexterity may be more effective than interventions involving cognitive skills for improving laparoscopic surgical skills [24-26,33,34], but the quality of available evidence is very low.

Professional Attitudes

Summary

Figure 6 summarizes the results of studies including professional attitudes outcomes.

Two RCTs compared a serious gaming/gamification intervention with traditional learning and measured outcomes related to professional attitudes. There was some evidence of a small magnitude of effect for a serious game, preparing students to go into the operating theatre for the first time, compared with traditional learning (SMD: 0.49, 95% CI: 0.14-0.84) [28]. A study comparing an insulin-prescribing game with an onsite learning activity for primary care reported insufficient data for comparisons between groups [43].

One cRCT was also included [41]. When reanalyzed with the number of clusters as the sample size to account for clustering in the analysis, there was evidence of intervention effectiveness, but this may not have been statistically significant and the analysis was likely underpowered (SMD: 1.23, 95% CI: –0.55 to 3.02). The quality of evidence for this outcome and comparison was rated very low according to the GRADE assessment.

Serious Gaming/Gamification Versus Other Modalities Of Digital Education

One study compared a serious game and an e-module on patient safety and reported no difference between groups in perceived patient safety behavior or reported stress [17].

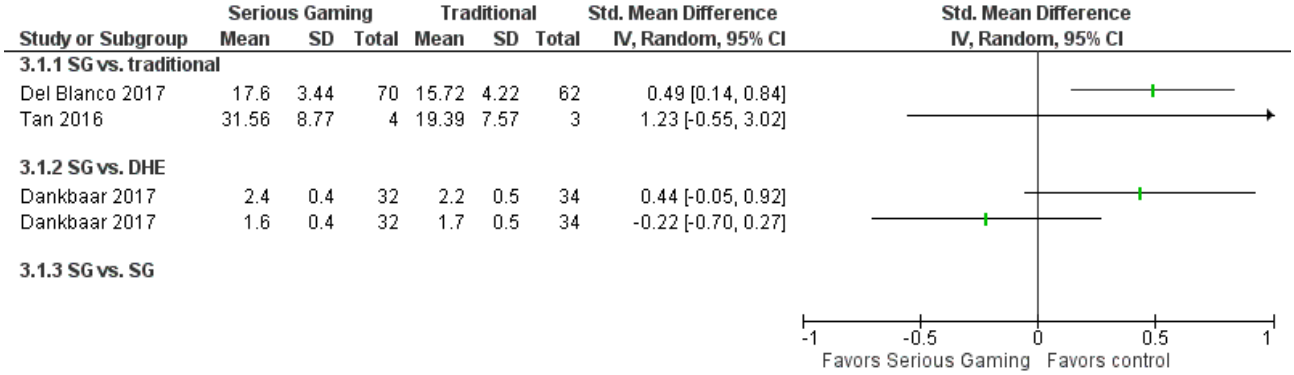
The quality of evidence was rated low according to the GRADE assessment.

Serious Gaming/Gamification Versus Serious Gaming/Gamification

One study (reported in two papers) compared two serious diverting interventions, one was a first-person shooter (FPS) and one was a non-FPS, and reported no significant differences in self-efficacy or positive engagement modes [24,25]. Data were insufficient for calculation of effect sizes. The quality of the evidence was very low according to GRADE assessment.

XSL•FO
RenderX

Figure 6. Forest plot for attitudes outcomes. IV: inverse variance; SG: serious games; DHE: digital health education.



Satisfaction

Summary

Eleven RCTs [15,16,20-22,26,27,32,33,39,45] and two cRCTs measured outcomes relating to satisfaction [19,41]. Seven studies did not measure satisfaction in a comparison group [19,21,22,27,32,35,45], Diehl et al measured satisfaction with the intervention and comparison group using different scales [43], and Kerfoot et al reported results for participants in both groups combined [15]; therefore, these studies did not meet the inclusion criteria for this review. The remaining studies showed mixed evidence and are compared below.

Serious Gaming/Gamification Versus Traditional Learning

One study reported significantly better attitudes toward learning among a serious gaming group (a urology adventure game group) compared with a written script [20]. The quality rating assessed using GRADE was low.

Serious Gaming/Gamification Versus Other Modalities of Digital Education

Three studies reported higher satisfaction scores for serious gaming/gamification on managing acutely unwell patients [16], patient safety [17], and training during a pediatric clerkship [39] compared with an e-module [16,17] or Web-based flashcards [39] covering the same topics. The quality rating assessed using GRADE was low.

Serious Gaming/Gamification Versus Serious Gaming/Gamification

Results of a participant survey [26] suggested that more participants in the FPS gaming group than in the non-FPS gaming group found the intervention beneficial for their performance on a surgical simulator. No significance test was reported. The quality rating assessed using GRADE was very low.

Secondary Outcomes

No studies measured economic outcomes of education or adverse effects of the intervention.

Discussion

Overview

The objective of this systematic review was to evaluate the effectiveness of serious gaming and gamification interventions for delivering pre- and postregistration health professions education. A total of 30 studies, most at high risk of bias according to Higgins [12], were identified, with high levels of heterogeneity in terms of populations and outcomes.

Serious gaming/gamification has the potential to reach a global audience and hence has been identified as a possible educational strategy that could contribute to transformation of health professions education. Results from our review show that serious gaming/gamification in pre- and post- registration health professions education could result in increased knowledge, skills, and satisfaction when compared to traditional education and, perhaps, other modalities of digital education.

Most of the current literature on the effectiveness of serious gaming/gamification has been performed in high-income countries, which limits the applicability of this review's findings to low- and middle-income countries. This is a key gap in the evidence, as low-and middle-income countries are most affected by the worldwide shortage of trained health workers [46]. Other limitations of the evidence base include the lack of studies assessing patient outcomes, or clinician behavior.

The cost of serious gaming devices might be a barrier for use compared with traditional lectures or text books. For example, some of the included studies used game consoles, which many health care workers, particularly in low- and middle-income settings, may not have access to. Other included studies used lower-cost modes of delivery, such as projecting a serious game to a group of students who played together. However, none of the eligible studies provided any information about economic outcomes of education or adverse or unintended effects of the intervention, which limits our understanding of the feasibility of implementing these interventions in practice and our understanding of the applicability of serious gaming/gamification as a cost-effective solution.

Considering the types of interventions that may be effective, based on classification of interventions by original design intention, there were no clear patterns suggesting differing effectiveness between custom designed games and commercial

XSL•FO
RenderX

off-the-shelf games for skill outcomes. Only custom-designed interventions were used to improve knowledge.

There was considerable heterogeneity in the results, particularly for skill outcomes, which we were unable to explain by systematic consideration of the types of intervention, population, and comparison group.

As serious gaming/gamification is an emerging field in the education sector, there are few previous reviews of the literature on its role in health professions education. Wang et al [47] conducted a systematic review of serious games for training health care professionals focused on game development and evaluation methodologies and reported a growing number of interventions and diversity of game genres over time [47]. Similar to our review, they found that study designs and methodological quality were heterogeneous and that best practices for development, evaluation, and use of such interventions are still being defined. A scoping review of serious gaming/gamification in health professions highlighted the need for economic evaluation of interventions, particularly when studies show no difference in efficacy between a serious game and traditional learning [48]. Our review contributes to the literature by providing an up-to-date summary of the evidence, focused on intervention effectiveness with a comprehensive systematic search. This is the first systematic review of the evidence indicating that serious gaming/gamification may improve participant knowledge compared with traditional learning.

There is a broad range of literature on serious gaming beyond health professions education. Meta-analyses have suggested that these interventions could significantly enhance learning among school students [49], adult workforce trainees [50], and mixed-age groups with regard to cognitive and attitudinal outcomes [51,52] and knowledge acquisition [53]. These reviews also suggested that games were more effective if they were supplemented with other methods of instruction, had multiple sessions, and involved active rather than passive learning. It was unclear whether playing as a group or alone was more effective. Systematic reviews have also suggested that serious gaming may have a role in the management of various medical conditions such as depression [54] and chronic conditions in young people [55] and in improving health outcomes [56]. The body of evidence on gamification interventions for education is smaller, with a systematic mapping study suggesting that most studies focused on the role of such interventions in student engagement and were published only as conference papers rather than full peer-reviewed articles [57]. A systematic review identified some evidence that gamification can be beneficial for health behavior change and well-being [58].

This review suggests that serious gaming may have the potential to advance education by improving knowledge, and possibly skill, outcomes for health professions compared with traditional learning. It may be able to provide educational interventions that are of equivalent educational value to other kinds of digital education, but with improved learner satisfaction. If this approach is equivalent to other kinds of education in terms of outcome but more cost-effective or able to offer greater access, it may provide further reasons to recommend serious gaming/gamification interventions, but no studies assessing these factors were identified.

Only two studies assessed gamification interventions. One suggested that the intervention was more effective than an online posting in improving knowledge by a large magnitude. The other suggested greater improvements in patient outcomes for questions spaced with four questions every 4 days rather than two questions every 2 days. These findings suggest that it may be worthwhile to incorporate gamification techniques into education, where possible, particularly for interventions aimed at improving knowledge, although further evidence is needed to establish the effectiveness among different groups of health professions for a wider range of patient outcomes and skill- and attitudes-related outcomes.

Strengths and Limitations

This review adopted a detailed and comprehensive search strategy without language limitations, followed by robust screening, data extraction, and risk-of-bias assessments, adhering to the Cochrane guidelines [12]. Thirty studies were found to be eligible, but most of them were at high risk of bias according to Higgins [12], with high levels of heterogeneity in terms of populations and outcomes. This heterogeneity of the included studies made it inappropriate to perform meta-analysis for any outcomes. Evidence for the majority of the outcomes and comparisons in the review was considered of low quality. Many studies have small sample sizes that were unlikely to provide sufficient power to detect an effect, provided insufficient detail for complete risk of bias assessment, and did not report all data for all outcomes assessed; in addition, statistical analysis was often not performed appropriately for the data (eg, not accounting for clustering), reducing confidence in the results (Figures 2 and 3). Only two studies of gamification interventions were identified.

Future Research

Serious gaming has the potential to contribute to the field of health professions education, but given that most studies to date are of low quality and carried out in high-income countries, future research should seek to use an RCT or cRCT design following a published protocol; evaluate interventions with a robust theoretical underpinning; be adequately powered; involve participants from low- and middle-income countries; appropriately randomize participants and blind outcome assessors, where possible; use validated outcome-assessment tools, facilitating comparability between interventions and studies; compare both serious gaming and gamification interventions with each other and with controls (other types of digital health education or traditional learning); and assess patient outcomes, participant behavior, attitudes, economic outcomes of education, and adverse events.

Conclusions

There is some evidence that serious gaming/gamification may improve health professionals' knowledge after the intervention compared with traditional education. In addition, some low-quality evidence shows that serious gaming/gamification may improve or be equivalent to traditional education for skills and to other modalities of digital education for knowledge and

XSL•FO
RenderX

skills. Future research should evaluate theory-grounded interventions and assess patient outcomes, economic outcomes of education, and adverse events.

Acknowledgments

This review was conducted in collaboration with the Health Workforce Department at the World Health Organization (WHO). We would like to thank Mr Carl Gornitzki, Ms GunBrit Knutssön, and Mr Klas Moberg from the University Library, Karolinska Institutet, Sweden, for developing the search strategy. We gratefully acknowledge funding from the Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore (e-learning for health professionals' education grant). We would like to thank Julian Alvarez, CIREL Laboratory of the Educational Science Department, University of Lille, France, and Jurriaan van Rijswijk, Games for Health Europe, Helmond, the Netherlands, for their content expertise and invaluable guidance in developing the review protocol.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MEDLINE (Ovid) search strategy.

[PDF File (Adobe PDF File), 93KB - jmir_v21i3e12994_app1.pdf]

Multimedia Appendix 2

Outcome and results of included studies comparing serious gaming.

[PDF File (Adobe PDF File), 96KB - jmir_v21i3e12994_app2.pdf]

Multimedia Appendix 3

Outcome and results of included studies comparing serious gaming/gamification and other digital education approaches.

[PDF File (Adobe PDF File), 137KB - jmir_v21i3e12994_app3.pdf]

Multimedia Appendix 4

Outcome and results of included studies comparing serious gaming/gamification and another type of serious gaming/gamification intervention.

[PDF File (Adobe PDF File), 99KB - jmir_v21i3e12994_app4.pdf]

References

1. Alvarez J, Djaouti D. Serious Games: An Introduction. Limoges, France: Questions Théoriques; 2012.
2. Alvarez J. From Videogame to Serious Game: the concept of Serious diverting and Serious Modding. 2015 Presented at: GameDev Days; 8-9 April; Tallinn, Estonia.
3. Wortley, David. SlideShare. 2013. Gamification - Concepts and Applications URL: https://www.slideshare.net/dwortley/dropbox-file [accessed 2018-10-16] [WebCite Cache ID 73DhKDUWL]
4. Deterding S, Dixon D, Khaled R, Nacke L. Gamification: toward a definition. 2011 Presented at: The ACM CHI Conference on Human Factors in Computing; May 7-12; Vancouver, Canada.
5. Akl EA, Kairouz VF, Sackett KM, Erdley WS, Mustafa RA, Fiander M, et al. Educational games for health professionals. Cochrane Database Syst Rev 2013 Mar 28(3):CD006411. [doi: 10.1002/14651858.CD006411.pub4] [Medline: 23543543]
6. Susi T, Johannesson M, Backlund P. Serious Games - An Overview. 2007. URL: https://www.diva-portal.org/smash/get/diva2:2416/FULLTEXT01.pdf [accessed 2019-02-25] [WebCite Cache ID 76SEmyVq4]
7. Prensky M. Digital game-based learning. Comput Entertain 2003 Oct 01;1(1):21-25. [doi: 10.1145/950566.950596]
8. Garris R, Ahlers R, Driskell J. Games, Motivation, and Learning: A Research and Practice Model. Simulation & Gaming 2016 Aug 18;33(4):441-467. [doi: 10.1177/1046878102238607]
9. Ruiz JG, Mintzer MJ, Issenberg SB. Learning objects in medical education. Med Teach 2006 Nov;28(7):599-605. [doi: 10.1080/01421590601039893] [Medline: 17594550]
10. Allery LA. Educational games and structured experiences. Med Teach 2004 Sep;26(6):504-505. [doi: 10.1080/01421590412331285423] [Medline: 15763824]

XSL•FO
RenderX

11. Vogel M, Wood DF. Love it or hate it? Medical students' attitudes to computer-assisted learning. Med Educ 2002 Mar;36(3):214-215. [Medline: 11879510]
12. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. 2011. URL: https://training.cochrane.org/handbook [accessed 2019-02-25] [WebCite Cache ID 76SFC3vIW]
13. 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 Dec 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
14. Chien JH, Suh IH, Park S, Mukherjee M, Oleynikov D, Siu K. Enhancing fundamental robot-assisted surgical proficiency by using a portable virtual simulator. Surg Innov 2013 Apr;20(2):198-203. [doi: 10.1177/1553350612458545] [Medline: 22956399]
15. Kerfoot BP, Baker H. An online spaced-education game for global continuing medical education: a randomized trial. Ann Surg 2012 Jul;256(1):33-38. [doi: 10.1097/SLA.0b013e31825b3912] [Medline: 22664558]
16. Dankbaar MEW, Alsma J, Jansen EEH, van Merrienboer JJG, van Saase JLCM, Schuit SCE. An experimental study on the effects of a simulation game on students' clinical cognitive skills and motivation. Adv Health Sci Educ Theory Pract 2016 Aug;21(3):505-521 [FREE Full text] [doi: 10.1007/s10459-015-9641-x] [Medline: 26433730]
17. Dankbaar MEW, Richters O, Kalkman CJ, Prins G, Ten Cate OTJ, van Merrienboer JJG, et al. Comparative effectiveness of a serious game and an e-module to support patient safety knowledge and awareness. BMC Med Educ 2017 Feb 02;17(1):30 [FREE Full text] [doi: 10.1186/s12909-016-0836-5] [Medline: 28148296]
18. Graafland M, Bemelman WA, Schijven MP. Game-based training improves the surgeon's situational awareness in the operation room: a randomized controlled trial. Surg Endosc 2017 Dec;31(10):4093-4101 [FREE Full text] [doi: 10.1007/s00464-017-5456-6] [Medline: 28281117]
19. Lagro J, van de Pol MHJ, Laan A, Huijbregts-Verheyden FJ, Fluit LCR, Olde Rikkert MGM. A randomized controlled trial on teaching geriatric medical decision making and cost consciousness with the serious game GeriatriX. J Am Med Dir Assoc 2014 Dec;15(12):957.e1-957.e6. [doi: 10.1016/j.jamda.2014.04.011] [Medline: 24913210]
20. Boeker M, Andel P, Vach W, Frankenschmidt A. Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. PLoS One 2013;8(12):e82328 [FREE Full text] [doi: 10.1371/journal.pone.0082328] [Medline: 24349257]
21. Hannig A, Lemos M, Spreckelsen C, Ohnesorge-Radtke U, Rafai N. Skills-O-Mat: Computer Supported Interactive Motion- and Game-Based Training in Mixing Alginate in Dental Education. Journal of Educational Computing Research 2013 Jul 09;48(3):315-343. [doi: 10.2190/EC.48.3.c]
22. Cook NF, McAloon T, O'Neill P, Beggs R. Impact of a web based interactive simulation game (PULSE) on nursing students' experience and performance in life support training--a pilot study. Nurse Educ Today 2012 Aug;32(6):714-720. [doi: 10.1016/j.nedt.2011.09.013] [Medline: 22082881]
23. Knight JF, Carley S, Tregunna B, Jarvis S, Smithies R, de Freitas S, et al. Serious gaming technology in major incident triage training: a pragmatic controlled trial. Resuscitation 2010 Sep;81(9):1175-1179. [doi: 10.1016/j.resuscitation.2010.03.042] [Medline: 20732609]
24. Hedman L, Schlickum M, Felländer-Tsai L. Surgical novices randomized to train in two video games become more motivated during training in MIST-VR and GI Mentor II than students with no video game training. Stud Health Technol Inform 2013;184:189-194. [Medline: 23400154]
25. Schlickum MK, Hedman L, Enochsson L, Kjellin A, Felländer-Tsai L. Systematic video game training in surgical novices improves performance in virtual reality endoscopic surgical simulators: a prospective randomized study. World J Surg 2009 Nov;33(11):2360-2367. [doi: 10.1007/s00268-009-0151-y] [Medline: 19649553]
26. Kolga SM, Hedman L, Enochsson L, Kjellin A, Felländer-Tsai L. Transfer of systematic computer game training in surgical novices on performance in virtual reality image guided surgical simulators. Stud Health Technol Inform 2008;132:210-215. [Medline: 18391288]
27. Boada I, Rodriguez-Benitez A, Garcia-Gonzalez JM, Olivet J, Carreras V, Sbert M. Using a serious game to complement CPR instruction in a nurse faculty. Comput Methods Programs Biomed 2015 Nov;122(2):282-291. [doi: 10.1016/j.cmpb.2015.08.006] [Medline: 26319184]
28. Del Blanco A, Torrente J, Fernández-Manjón B, Ruiz P, Giner M. Using a videogame to facilitate nursing and medical students' first visit to the operating theatre. A randomized controlled trial. Nurse Educ Today 2017 Aug;55:45-53. [doi: 10.1016/j.nedt.2017.04.026] [Medline: 28521249]
29. Foss B, Lokken A, Leland A, Stordalen J, Mordt P, Oftedal BJ. Digital Game-Based Learning: a supplement for medication calculation drills in nurse education. E-Learning and Digital Media 2014;11(4):342-349.
30. Giannotti D, Patrizi G, Di Rocco G, Vestri AR, Semproni CP, Fiengo L, et al. Play to become a surgeon: impact of Nintendo Wii training on laparoscopic skills. PLoS One 2013;8(2):e57372 [FREE Full text] [doi: 10.1371/journal.pone.0057372] [Medline: 23460845]
31. Adams BJ, Margaron F, Kaplan BJ. Comparing video games and laparoscopic simulators in the development of laparoscopic skills in surgical residents. J Surg Educ 2012;69(6):714-717. [doi: 10.1016/j.jsurg.2012.06.006] [Medline: 23111035]

XSL•FO
RenderX

32. Amer RS, Denehy GE, Cobb DS, Dawson DV, Cunningham-Ford MA, Bergeron C. Development and evaluation of an interactive dental video game to teach dentin bonding. J Dent Educ 2011 Jun;75(6):823-831 [FREE Full text] [Medline: 21642529]
33. Ju R, Chang PL, Buckley AP, Wang KC. Comparison of Nintendo Wii and PlayStation2 for enhancing laparoscopic skills. JSLS 2012;16(4):612-618 [FREE Full text] [doi: 10.4293/108680812X13462882737294] [Medline: 23484573]
34. Ju R, Chang PL, Buckley AP, Wang KV. Is Nintendo Wii a More Suitable Video Game Platform Than Playstation 2 for Enhancing Laparoscopic Skills? Journal of Minimally Invasive Gynecology 2011 Nov;18(6):S65-S66. [doi: 10.1016/j.jmig.2011.08.226]
35. Katz D, Zerillo J, Kim S, Hill B, Wang R, Goldberg A, et al. Serious gaming for orthotopic liver transplant anesthesiology: A randomized control trial. Liver Transpl 2017 Dec;23(4):430-439 [FREE Full text] [doi: 10.1002/lt.24732] [Medline: 28133947]
36. Kerfoot BP, Turchin A, Breydo E, Gagnon D, Conlin PR. An online spaced-education game among clinicians improves their patients' time to blood pressure control: a randomized controlled trial. Circ Cardiovasc Qual Outcomes 2014 May;7(3):468-474 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.113.000814] [Medline: 24847084]
37. LeFlore J, Anderson M, Zielke MA, Nelson KA, Thomas PE, Hardee G, et al. Can a virtual patient trainer teach student nurses how to save lives--teaching nursing students about pediatric respiratory diseases. Simul Healthc 2012 Feb;7(1):10-17. [doi: 10.1097/SIH.0b013e31823652de] [Medline: 22228285]
38. Plerhoples TA, Zak Y, Hernandez-Boussard T, Lau J. Another use of the mobile device: warm-up for laparoscopic surgery. J Surg Res 2011 Oct;170(2):185-188. [doi: 10.1016/j.jss.2011.03.015] [Medline: 21529831]
39. Sward KA, Richardson S, Kendrick J, Maloney C. Use of a Web-based game to teach pediatric content to medical students. Ambul Pediatr 2008;8(6):354-359. [doi: 10.1016/j.ambp.2008.07.007] [Medline: 19084784]
40. Gauthier A, Corrin M, Jenkinson J. Exploring the influence of game design on learning and voluntary use in an online vascular anatomy study aid. Computers & Education 2015 Sep;87:24-34. [doi: 10.1016/j.compedu.2015.03.017]
41. Tan AJQ, Lee CCS, Lin PY, Cooper S, Lau LST, Chua WL, et al. Designing and evaluating the effectiveness of a serious game for safe administration of blood transfusion: A randomized controlled trial. Nurse Educ Today 2017 Aug;55:38-44. [doi: 10.1016/j.nedt.2017.04.027] [Medline: 28521248]
42. de Araujo TB, Silveira FR, Souza DLS, Strey YTM, Flores CD, Webster RS. Impact of video game genre on surgical skills development: a feasibility study. J Surg Res 2016 Mar;201(1):235-243. [doi: 10.1016/j.jss.2015.07.035] [Medline: 26850208]
43. Diehl LA, Souza RM, Gordan PA, Esteves RZ, Coelho ICM. InsuOnline, an Electronic Game for Medical Education on Insulin Therapy: A Randomized Controlled Trial With Primary Care Physicians. J Med Internet Res 2017 Dec 09;19(3):e72 [FREE Full text] [doi: 10.2196/jmir.6944] [Medline: 28279950]
44. Rondon S, Sassi FC, Furquim de Andrade CR. Computer game-based and traditional learning method: a comparison regarding students' knowledge retention. BMC Med Educ 2013 Feb 25;13:30 [FREE Full text] [doi: 10.1186/1472-6920-13-30] [Medline: 23442203]
45. Li J, Xu Y, Xu Y, Yue P, Sun L, Guo M, et al. 3D CPR Game Can Improve CPR Skill Retention. Stud Health Technol Inform 2015;216:974. [Medline: 26262276]
46. World Health Organization. 2018. Global health workforce shortage to reach 12.9 million in coming decades URL: https://www.who.int/mediacentre/news/releases/2013/health-workforce-shortage/en/ [accessed 2018-10-16] [WebCite Cache ID 73Dl2vrdN]
47. Wang R, DeMaria S, Goldberg A, Katz D. A Systematic Review of Serious Games in Training Health Care Professionals. Simul Healthc 2016 Feb;11(1):41-51. [doi: 10.1097/SIH.0000000000000118] [Medline: 26536340]
48. Ricciardi F, De Paolis LT. A Comprehensive Review of Serious Games in Health Professions. International Journal of Computer Games Technology 2014;2014:1-11. [doi: 10.1155/2014/787968]
49. Clark D, Tanner-Smith E, Killingsworth S. Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. Rev Educ Res 2016 Mar;86(1):79-122 [FREE Full text] [doi: 10.3102/0034654315582065] [Medline: 26937054]
50. Sitzmann T. A meta-analytic examination of the instructional effectiveness of computer-based simulation games. Personnel Psychology 2011;64:489-528. [doi: 10.1111/j.1744-6570.2011.01190.x]
51. Vogel J, Vogel D, Cannon-Bowers J, Bowers C, Muse K, Wright M. Computer Gaming and Interactive Simulations for Learning: A Meta-Analysis. Journal of Educational Computing Research 2016 Jul 22;34(3):229-243. [doi: 10.2190/FLHV-K4WA-WPVQ-H0YM]
52. Wouters P, van Nimwegen C, van Oostendorp H, van der Spek E. A meta-analysis of the cognitive and motivational effects of serious games. Journal of Educational Psychology 2013;105(2):249-265. [doi: 10.1037/a0031311]
53. Boyle E, Hainey T, Connolly T, Gray G, Earp J, Ott M, et al. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Computers & Education 2016 Mar;94:178-192. [doi: 10.1016/j.compedu.2015.11.003]
54. Fleming T, Cheek C, Merry S, Thabrew H, Bridgman H, Stasiak K, et al. Serious games for the treatment or prevention of depression: a systematic review. Revista de Psicopatologia y Psicologia Clinica 2014;19(3):a.

XSL•FO
RenderX

55. Charlier N, Zupancic N, Fieuws S, Denhaerynck K, Zaman B, Moons P. Serious games for improving knowledge and self-management in young people with chronic conditions: a systematic review and meta-analysis. J Am Med Inform Assoc 2016 Jan;23(1):230-239. [doi: 10.1093/jamia/ocv100] [Medline: 26186934]
56. Primack B, Carroll M, McNamara M, Klem M, King B, Rich M, et al. Role of video games in improving health-related outcomes: a systematic review. Am J Prev Med 2012 Jun;42(6):630-638 [FREE Full text] [doi: 10.1016/j.amepre.2012.02.023] [Medline: 22608382]
57. De Sousa Borges S, Durelli V, Reis H, Isotani S. A systematic mapping on gamification applied to education. 2014 Presented at: Proceedings of the 29th Annual ACM Symposium on Applied Computing, ACM; March 24-28; Gyeongju, Republic of Korea p. 216-222.
58. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: A systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]

Abbreviations

COTS: commercial off the shelf
cRCT: cluster randomized controlled trials
e-learning: electronic learning
e-module: electronic module
FPS: first-person shooter
OSCE: objective structured clinical examination
RCT: randomized controlled trials
SMD: standardized mean differences

Edited by A Marusic; submitted 30.11.18; peer-reviewed by V Katavic, R Alkoudmani, K Blondon; comments to author 21.12.18; revised version received 21.01.19; accepted 05.02.19; published 28.03.19.

Please cite as:
Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, Dauwels-Okutsu S, Nikolaou CK, Zary N, Campbell J, Car J
Serious Gaming and Gamification Education in Health Professions: Systematic Review
J Med Internet Res 2019;21(3):e12994
URL: http://www.jmir.org/2019/3/e12994/
doi:10.2196/12994
PMID:30920375

©Sarah Victoria Gentry, Andrea Gauthier, Beatrice L'Estrade Ehrstrom, David Wortley, Anneliese Lilienthal, Lorainne Tudor Car, Shoko Dauwels-Okutsu, Charoula K Nikolaou, Nabil Zary, James Campbell, Josip Car. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

The Case for a Hippocratic Oath for Connected Medical Devices: Viewpoint

Beau Woods[1,2,3*], BS; Andrea Coravos[1,4,5,6*], MBA; Joshua David Corman[1,3,7,8*], BA

[1]I Am The Cavalry, Washington, DC, United States
[2]Biohacking Village, New York, NY, United States
[3]Atlantic Council, Washington, DC, United States
[4]Elektra Labs, Boston, MA, United States
[5]Harvard-MIT Center for Regulatory Science, Harvard University, Boston, MA, United States
[6]Digital Medicine (DiMe) Society, New York, NY, United States
[7]Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA, United States
[8]PTC, Needham, MA, United States
[*]all authors contributed equally

Corresponding Author:
Andrea Coravos, MBA
Elektra Labs
Harvard Innovation Labs
125 Western Ave
Boston, MA, 02163
United States
Phone: 1 9788421240
Email: andrea@elektralabs.com

Abstract

Prior to graduating from medical school, soon-to-be physicians take the Hippocratic Oath, a symbolic declaration to provide care in the best interest of patients. As the medical community increasingly deploys connected devices to deliver patient care, a critical question emerges: should the manufacturers and adopters of these connected technologies be governed by the symbolic spirit of the Hippocratic Oath? In 2016, I Am The Cavalry, a grassroots initiative from the cybersecurity research community, published the first Hippocratic Oath for Connected Medical Devices (HOCMD), containing 5 principles. Over the past three years, the HOCMD has gained broad support and influenced regulatory policy. We introduce 5 case studies of the HOCMD in practice, illustrating how the 5 principles can lead to a safer and more effective adoption of connected medical technologies.

(J Med Internet Res 2019;21(3):e12568) doi:10.2196/12568

KEYWORDS

ethics; cybersecurity; information technology; delivery of health care; connected devices

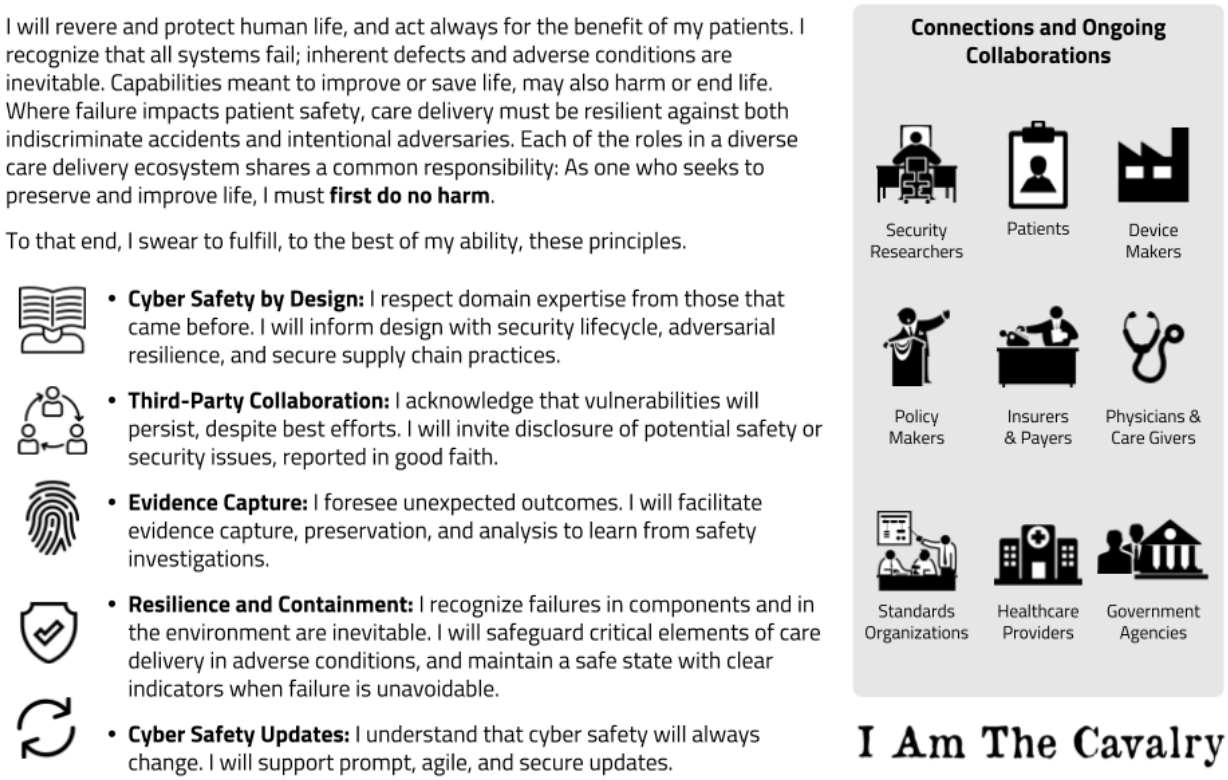
Introduction

Prior to graduating from medical school, soon-to-be physicians take the Hippocratic Oath [1], a symbolic declaration to provide care in the best interest of patients. As the medical community increasingly deploys connected devices to deliver patient care, a critical question emerges: should the manufacturers and adopters of these connected technologies be governed by the symbolic spirit of the Hippocratic Oath?

Five years ago, Joshua David Corman and Nicholas Percoco founded *I Am The Cavalry*, a grassroots initiative from the cybersecurity research community. Its mission is "to ensure technologies with the potential to impact public safety and human life are worthy of our trust" [2]. In 2016, along with a diverse stakeholder group across the health care ecosystem, including payers, providers, patients, policy makers, and physicians, I Am The Cavalry drafted the first version of a Hippocratic Oath for Connected Medical Devices (HOCMD) (see Figure 1) and published an open letter on its website [3]. Similar to graduating medical students who pledge to protect the patient and hospital systems who pledge to protect patients' data and information, the HOCMD outlines five guiding ethical principles for manufacturers, organizations, and individuals delivering care through connected medical devices.

XSL•FO
RenderX

Figure 1. Hippocratic Oath for Connected Medical Devices (HOCMD).



Adopting these principles not only builds trust in the connected product for patients, clinicians, hospitals, and regulators, but also serves a business function. Similar to a drug that is pulled off the market because of unintended side effects, connected products with security vulnerabilities can sharply impact a company's share price and valuation. For medical device investors who want to avoid security issues in their portfolio companies, these principles enable them to evaluate a connected medical product's risk profile, safety, and effectiveness.

Although all systems eventually fail, not all failures have to cause harm. Organizations that follow these five principles tend to build safer products with fewer malfunctions and deliver better care to patients. To illustrate the principles, we have outlined five case studies based on true events; we have blinded some examples at the request of the parties involved.

Five Case Studies based on the 5 Oath Principles

Oath Principle #1: Cyber Safety by Design

Riley, a hospital procurement analyst, reviews several medical devices, struggling to distinguish security features between similar products and the potential impact on cost and risk to her organization. Riley and her peers often rely on sales literature and expensive internal testing to understand how to buy the best devices to fit their clinical environments. Products with better documentation make her job easier and increase the organization's confidence that the devices can be deployed and used safely and effectively, without undue cost or risk.

Riley prefers to procure devices from manufacturers who follow practices laid out in *cyber safety by design*, because she can more easily anticipate and avoid product failures. For example, preferred devices would have instructions for safe and secure implementation, design assumptions, a *software bill of materials* (SBOM), and other documentation outlining how cybersecurity is designed into the product's lifecycle. Health care organizations like the Mayo Clinic have started requiring manufacturers to disclose more information about their security capabilities [4]. These requirements are being widely adopted by manufacturers like Philips, Siemens, and Becton Dickinson, and are becoming standard practice through industry-led initiatives like the Manufacturer Disclosure Statement for Medical Device Security (MDS2) disclosures [5].

This past year, the US Food and Drug Administration (FDA) incorporated the *cyber safety by design* principles in their Medical Device Safety Action Plan [6] and updated Premarket Guidance [7], which proposes requiring an SBOM or *cybersecurity bill of materials* (CBOM), in which manufacturers have to clearly communicate choices and capabilities that impact cyber safety and other practices from this first principle.

XSL•FO
RenderX

Oath Principle #2: Third-Party Collaboration

In early 2017, Jay Radcliffe, a researcher with the security firm Rapid7, discovered a flaw in how the Johnson & Johnson Animas insulin pumps handle commands from the unit's remote control. Exploiting the flaw could potentially lead to unauthorized access to the pump through its unencrypted radio frequency communication system.

Undisclosed flaws represent potential harm to patients through accidents or adverse events. While it may seem inconceivable, many manufacturers have threatened security researchers with legal repercussions [8] for identifying these kinds of issues. Although immediate public disclosure often catalyzes a prompt fix, releasing the disclosure often starts a footrace between defenders and adversaries that may put patients at greater risk. Jay had a difficult decision to make: should he disclose the vulnerability and, if so, how?

Thankfully, Johnson & Johnson had recently drafted a coordinated disclosure policy that invited third-party collaboration. After investigation, the issue was considered to be a low safety risk due to existing mitigations. Jay and Johnson & Johnson codeveloped an effective approach to addressing the issue, communicated with patients and physicians, kept the FDA and other organizations informed, and improved internal processes. This was a win for all involved and decreased potential harm to patients.

Many throughout the health care ecosystem, including regulators like the FDA, are voicing their support for coordinated vulnerability disclosure. The FDA recently outlined this principle in their guidance for Post-Market Management of Cybersecurity in Medical Devices [9] and published a collaborative report with the Medical Device Innovation Consortium (MDIC) that advances the concept of coordinated disclosure [10]. Perhaps untraditionally, the FDA has also started to participate at hacker conferences like DEF CON [11] to encourage greater collaboration when surfacing vulnerabilities.

As of this writing, nearly 20 companies have published programs [12] to receive and handle reports of security vulnerabilities. Programs based around standards (ie, ISO 29147, ISO 30111, and National Telecommunications and Information Administration [NTIA] Early Stage Template [13]) frequently leverage existing mechanisms internally that are well-tested and focus on incentives, aligning researchers and device makers toward safer products.

Oath Principle #3: Evidence Capture

Two years ago, a hospital sent a life-critical device to the manufacturer for investigation, suspecting that it was tampered with by a physician who was being sued for malpractice. After conducting a review, the manufacturer discovered that evidence of patient care had been wiped from the device, but it was unclear who did this, when, or why. The connected device lacked forensically sound evidence-capture tools.

When doctors are not sure why an adverse event happened, they perform tests or, in the case of death, an autopsy. When medical devices are involved, evidence from the device can and should support postmortem investigations as part of that autopsy. For this reason, the third principle of the Oath, *evidence capture*, alerts manufacturers to the need for built-in evidence-capture capabilities and tamper-evident logging to protect patients.

For instance, a secure system could store a fixed record of safety- and security-related data, such as software integrity checks and activity logs. This information could then be reviewed by analysts at the health care organization or the manufacturer to support investigations into malicious tampering, accidental harm, or cybersecurity issues. In addition, this capability lays the foundation for future capabilities that allow for more prompt and agile responses.

The FDA's recent draft premarket guidance [7] calls for a forensically sound evidence-capture capability, despite relatively little interest from industry and regulatory bodies. This capability would be invaluable; most devices currently lack activity logs and other integrity checks, indicating that it is possible that there have been adverse events related to device malfunction of which manufacturers are unaware. The FDA's updated guidance will likely incentivize more device makers to include this capability in their products.

Oath Principle #4: Resilience and Containment

In 2017, several German journalists sat in a room learning how to hack medical devices at a press event held by Dräger. First, they learned how to use tools to penetrate a basic installation of Windows XP, an operating system that had been unsupported for three years. After gaining complete control of the ventilators, the journalists tried similar techniques against a version of Windows XP that Dräger uses in some of its current products. None of the journalists, nor the subsequent professional penetration testers, were able to hack the device over the network. The company had hardened the system by disabling nonessential functions, encrypting data in transit, and isolating the network-connected elements from those that deliver patient care.

Medical devices are increasingly connected to the rest of the hospital and to the Internet. A traditional stethoscope or ultrasound are not connected to each other, let alone exposed to technically savvy adversaries thousands of miles away. Practices in the *resilience and containment* principle can contain the reach of adversaries by reducing risk from cascading cybersecurity failures, making failures evident, and ensuring devices default to a safe mode when they must fail.

Building on the *resilience and containment* principle, the FDA's premarket guidance outlines best practices, such as improving device isolation using firewalls, reducing elective exposure by disabling network-connected components, and protecting patient record integrity by encrypting data while in transit and on the device. The FDA has also issued two safety notifications related to failure to segregate safety-critical from noncritical components of Hospira infusion pumps [14,15].

Oath Principle #5: Cyber Safety Updates

In 2016, a vulnerability was discovered in St. Jude Medical's pacemakers, implantable cardioverter defibrillators, and bedside monitors. The chances that an adversary would exploit this vulnerability to cause harm seemed low, but the attack could

XSL•FO
RenderX

be carried out stealthily, triggering heart conditions in patients. An estimated 465,000 patients would be at risk before the manufacturer could build replacements and doctors could implant them [16].

Fortunately, St. Jude Medical had a better way to address the issue. A series of three software updates greatly reduced cost, increased adoption, hastened correction timelines, and minimized side effects compared to extreme solutions, such as replacing the device through explantation. A study showed that 25% of patients had the software update applied at their next clinic visit, with no pacing failures observed [17].

It should be common practice to update software routinely. The FDA has repeatedly stated that manufacturers are responsible for delivering security updates, which are considered routine and therefore not often subject to additional reviews. However, security and promptness of these updates still lag for most device makers and communication is difficult with many stakeholders involved with each device.

Accelerating Adoption of the Hippocratic Oath for Connected Medical Devices

While most doctors practicing in the United States take the Hippocratic Oath before going into clinical practice, developers of software-connected devices traditionally do not share a similar gatekeeper. As of today, individuals and organizations can publicly commit to the HOCMD via the I Am The Cavalry website [3]. Though this lacks the resonance of repeating the Oath aloud at a hard-earned graduation, the digital pledge carries its own gravitas. Digital statements are accessible and searchable. An online pledge holds leadership, scientists, and engineers accountable for their work by the public and signals an unwavering faith in the quality of the products they have created. The Oath can also be taken by organizations, like hospital systems, who pledge to act not only in the best interest of the corporal patient, but also in the best interest of the patient's information—the patient's *digital specimen*.

Oaths can serve two important purposes: first, to establish a set of standards and second, to remind an oath-taking community of their promised commitments. Although oaths have a storied history, evidence does not necessarily suggest that taking an oath leads to better behavior in practice even within medical professions [18]. Thus, as the HOCMD matures, its principles may be deployed more effectively in other formats. For instance, a checklist may be a more effective format than an oath for putting the principles into routine practice [19]. Atul Gawande popularized practical checklists in health care when he published The Checklist Manifesto, a short book on how not to make big mistakes, and outlined how a checklist format can increase efficiency, consistency, and safety [20].

The HOCMD's influence on regulatory policy is another example of its principles moving from theory to practice. Although the FDA draft guidance for Management of Cybersecurity in Medical Devices, published in October 2018, does not call out the Oath by name, it references four Oath principles in its premarket guidance [7] and references the fifth in its postmarket guidance [9]. FDA officials have shared in speeches [11] and tweets [21] their learnings from the security and hacker communities as they draft modern policies to bring safe, connected, medical products to market.

We envision a world where all medical device manufacturers stand proudly behind their work by publicly committing to upholding these principles of cyber safety; we also envision a world where patients, hospitals, doctors, insurers, and regulators strongly favor more secure medical devices built by manufacturers who align with the Oath. As software-driven connected products drive more care delivery, we hope that those who manufacture and adopt life-critical products will commit to carrying similar values as the physicians who have attended to patients for centuries.

Acknowledgments

For suggesting compelling examples of Oath principles and reviews of the case studies, we gratefully acknowledge Jay Radcliffe, MS; Chris Tyberg; Hannes Molsen; Colin Morgan; Ginny Fahs; Ana Sofia Warner, MD; Stacey Higginbotham; Adam Conner-Simons; Janine Medina, MS; and the Biohacking Village at DEF CON. AC acknowledges support from Harvard-MIT Center for Regulatory Science. We also want to recognize the I Am The Cavalry working group, who drafted the original Hippocratic Oath for Connected Medical Devices.

Conflicts of Interest

The authors are members of I Am The Cavalry, the grassroots organization that drafted the Hippocratic Oath for Connected Medical Devices. I Am The Cavalry is an entirely volunteer effort without legal incorporation status.

References

1. Tyson P. PBS. 2001 Mar 27. The Hippocratic Oath today URL: https://www.pbs.org/wgbh/nova/article/hippocratic-oath-today/ [accessed 2018-10-22] [WebCite Cache ID 73Lh06EQQ]
2. I Am The Cavalry. URL: https://www.iamthecavalry.org/ [accessed 2019-01-02] [WebCite Cache ID 757vwSlS2]
3. I Am The Cavalry. Hippocratic Oath for Connected Medical Devices URL: https://www.iamthecavalry.org/domains/medical/oath/ [accessed 2018-10-20] [WebCite Cache ID 73Ij9CNM9]

XSL•FO
RenderX

4. Medical and Research Device Risk Assessment Vendor Packet Instructions. Rochester, MN: Mayo Clinic; 2018. URL: https://www.mayoclinic.org/documents/medical-device-vendor-instructions/doc-20389647 [accessed 2018-10-20] [WebCite Cache ID 73IjdfDIl]
5. National Electrical Manufacturers Association (NEMA). 2013 Oct 07. Manufacturer disclosure statement for medical device security URL: https://www.nema.org/Standards/Pages/Manufacturer-Disclosure-Statement-for-Medical-Device-Security.aspx [accessed 2018-10-21] [WebCite Cache ID 73KmoQ4Ao]
6. Medical Device Safety Action Plan: Protecting Patients, Promoting Public Health. Silver Spring, MD: US Food and Drug Administration; 2018. URL: https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM604690.pdf [accessed 2018-10-21] [WebCite Cache ID 73KvY8HRa]
7. Content of Premarket Submissions for Management of Cybersecurity in Medical Devices: Draft Guidance for Industry and Food and Drug Administration Staff. Silver Spring, MD: Center for Devices and Radiological Health, US Food and Drug Administration; 2018 Oct 18. URL: https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM623529.pdf [accessed 2018-10-21] [WebCite Cache ID 73WNSLF4s]
8. Whittaker Z. ZDNet. 2018 Feb 19. Lawsuits threaten infosec research — just when we need it most URL: https://www.zdnet.com/article/chilling-effect-lawsuits-threaten-security-research-need-it-most/ [accessed 2018-10-21] [WebCite Cache ID 73KwFicUC]
9. Postmarket Management of Cybersecurity in Medical Devices: Guidance for Industry and Food and Drug Administration Staff. Silver Spring, MD: Center for Devices and Radiological Health, US Food and Drug Administration; 2016 Dec 28. URL: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm482022.pdf [accessed 2018-10-21] [WebCite Cache ID 73KwPnl1c]
10. Medical Device Cybersecurity Report: Advancing Coordinated Vulnerability Disclosure. Arlington, VA: Medical Device Innovation Consortium (MDIC); 2018. URL: http://mdic.org/wp-content/uploads/2018/10/MDIC-CybersecurityReport.pdf [accessed 2018-10-21] [WebCite Cache ID 73KwojnAP]
11. Gottlieb S. US Food and Drug Administration. 2018 Oct 01. Statement from FDA Commissioner Scott Gottlieb, MD on FDA's efforts to strengthen the agency's medical device cybersecurity program as part of its mission to protect patients URL: https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/UCM622074.htm [accessed 2018-10-22] [WebCite Cache ID 73LhhBvmf]
12. I Am The Cavalry. Vulnerability disclosure programs URL: https://www.iamthecavalry.org/resources/disclosure-programs/ [accessed 2018-10-21] [WebCite Cache ID 73KxNazAS]
13. "Early Stage" Coordinated Vulnerability Disclosure Template. Version 1.1. Washington, DC: NTIA Safety Working Group; 2016 Dec 15. URL: https://webcache.googleusercontent.com/search?q=cache:mbJ9mUcs-_IJ:https://www.ntia.doc.gov/files/ntia/publications/ntia_vuln_disclosure_early_stage_template.pdf+&cd=1&hl=en&ct=clnk&gl=us [accessed 2019-02-25] [WebCite Cache ID 76SW6QYuN]
14. Finkle J. Reuters. 2015 Jul 31. FDA warns of security flaw in Hospira infusion pumps URL: https://www.reuters.com/article/us-hospira-fda-cybersecurity/fda-warns-of-security-flaw-in-hospira-infusion-pumps-idUSKCN0Q52GJ20150731 [accessed 2018-10-22] [WebCite Cache ID 73Lij10Tw]
15. The Security Ledger. 2015 May 14. FDA safety advisory warns of cyber risk of drug pumps URL: https://securityledger.com/2015/05/fda-safety-advisory-warns-of-cyber-risk-of-drug-pumps/ [accessed 2018-10-22] [WebCite Cache ID 73Lj58ZRQ]
16. US Food and Drug Administration. 2017 Aug 29. Firmware update to address cybersecurity vulnerabilities identified in Abbott's (formerly St. Jude Medical's) implantable cardiac pacemakers: FDA safety communication URL: https://www.fda.gov/medicaldevices/safety/alertsandnotices/ucm573669.htm [accessed 2019-02-19] [WebCite Cache ID 76J0ovzrJ]
17. Saxon LA, Varma N, Epstein LM, Ganz LI, Epstein AE. Factors influencing the decision to proceed to firmware upgrades to implanted pacemakers for cybersecurity risk mitigation. Circulation 2018 Sep 18;138(12):1274-1276. [doi: 10.1161/CIRCULATIONAHA.118.034781] [Medline: 29748188]
18. Sierles F, Kushner B, Krause P. A controlled experiment with a medical student honor system. J Med Educ 1988 Sep;63(9):705-712. [Medline: 3418674]
19. Loukides M, Mason H, Patil DJ. Ethics and Data Science. Sebastopol, CA: O'Reilly Media, Inc; Jul 2018.
20. Gawande A. The Checklist Manifesto: How to Get Things Right. New York, NY: Metropolitan Books; 2010.
21. Gottlieb S. Twitter. 2018 Oct 07. Twitter post URL: https://twitter.com/SGottliebFDA/status/1048987791596642310 [accessed 2019-02-19] [WebCite Cache ID 76J0PZkHp]

Abbreviations

CBOM: cybersecurity bill of materials
FDA: US Food and Drug Administration
HOCMD: Hippocratic Oath for Connected Medical Devices
MDIC: Medical Device Innovation Consortium

XSL•FO
RenderX

MDS2: Manufacturer Disclosure Statement for Medical Device Security
NTIA: National Telecommunications and Information Administration
SBOM: software bill of materials

Edited by G Eysenbach, E Perakslis; submitted 01.11.18; peer-reviewed by K McDonald, P Lynch, J Suleder; comments to author 28.11.18; revised version received 23.01.19; accepted 28.01.19; published 19.03.19.

Please cite as:
Woods B, Coravos A, Corman JD
The Case for a Hippocratic Oath for Connected Medical Devices: Viewpoint
J Med Internet Res 2019;21(3):e12568
URL: http://www.jmir.org/2019/3/e12568/
doi:10.2196/12568
PMID:30888323

©Beau Woods, Andrea Coravos, Joshua David Corman. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Author Contribution Correction: An Integrated Influenza Surveillance Framework Based on National Influenza-Like Illness Incidence and Multiple Hospital Electronic Medical Records for Early Prediction of Influenza Epidemics: Design and Evaluation

Cheng-Yi Yang[1*], PhD; Ray-Jade Chen[2,3*], MSc, MD; Wan-Lin Chou[3], MS; Yuarn-Jang Lee[4], MD; Yu-Sheng Lo[1], PhD

[1]Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei, Taiwan

[2]Department of Surgery, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

[3]Taipei Medical University Hospital, Taipei, Taiwan

[4]Division of Infectious Disease, Department of Internal Medicine, Taipei Medical University Hospital, Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Yu-Sheng Lo, PhD
Graduate Institute of Biomedical Informatics
Taipei Medical University
250 Wuxing Street
Taipei, 11031
Taiwan
Phone: 886 939588576
Email: Loyusen@tmu.edu.tw

Related Article:

Correction of: http://www.jmir.org/2019/2/e12341/

(J Med Internet Res 2019;21(3):e13699) doi:10.2196/13699

The authors of "An Integrated Influenza Surveillance Framework Based on National Influenza-Like Illness Incidence and Multiple Hospital Electronic Medical Records for Early Prediction of Influenza Epidemics: Design and Evaluation" (J Med Internet Res 2019;21(2):e12341) inadvertently marked Yu-Sheng Lo as an equal contributor when that designation should have only been applied to Cheng-Yi Yang and Ray-Jade Chen. The asterisk denoting equal contribution has now been removed from Yu-Sheng Lo.

The correction will appear in the online version of the paper on the JMIR website on March 11, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Edited by G Eysenbach; submitted 14.02.19; this is a non–peer-reviewed article;accepted 06.03.19; published 12.03.19.

Please cite as:
Yang CY, Chen RJ, Chou WL, Lee YJ, Lo YS
Author Contribution Correction: An Integrated Influenza Surveillance Framework Based on National Influenza-Like Illness Incidence and Multiple Hospital Electronic Medical Records for Early Prediction of Influenza Epidemics: Design and Evaluation
J Med Internet Res 2019;21(3):e13699
URL: http://www.jmir.org/2019/3/e13699/
doi:10.2196/13699
PMID:30860974

XSL•FO
RenderX

©Cheng-Yi Yang, Ray-Jade Chen, Wan-Lin Chou, Yuarn-Jang Lee, Yu-Sheng Lo. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX