

Journal of Medical Internet Research

Impact Factor (2018): 4.945 - ranked #1 medical informatics journal by Impact Factor
Volume 19 (2017), Issue 4 ISSN: 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Evidence of Benefit of Telerehabilitation After Orthopedic Surgery: A Systematic Review (e142) Jose Pastora-Bernal, Rocio Martín-Valero, Francisco Barón-López, María Estebanez-Pérez.	5
Effect of a Nine-Month Web- and App-Based Workplace Intervention to Promote Healthy Lifestyle and Weight Loss for Employees in the Social Welfare and Health Care Sector: A Randomized Controlled Trial (e108) Nina Balk-Møller, Sanne Poulsen, Thomas Larsen.	18
Web-Based Intervention for Physical Activity and Fruit and Vegetable Intake Among Chinese University Students: A Randomized Controlled Trial (e106) Yan Duan, Julian Wienert, Chun Hu, Gang Si, Sonia Lippke.	31
Mobile Phone-Delivered Cognitive Behavioral Therapy for Insomnia: A Randomized Waitlist Controlled Trial (e70) Corine Horsch, Jaap Lancee, Fiemke Griffioen-Both, Sandor Spruit, Siska Fitrianie, Mark Neerincx, Robbert Beun, Willem-Paul Brinkman.	4 6
Using Persuasive Technology to Increase Physical Activity in People With Chronic Obstructive Pulmonary Disease by Encouraging Regular Walking: A Mixed-Methods Study Exploring Opinions and Preferences (e124) Yvonne Bartlett, Thomas Webb, Mark Hawley.	67
Therapist-Supported Internet-Based Cognitive Behavior Therapy for Stress, Anxiety, and Depressive Symptoms Among Postpartum Women: A Systematic Review and Meta-Analysis (e138) Ying Lau, Tha Htun, Suei Wong, Wai Tam, Piyanee Klainin-Yobas.	81
The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder (e105) Jennifer Nicholas, Andrea Fogarty, Katherine Boydell, Helen Christensen.	99
Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles (e118) Sohrab Saeb, Thaddeus Cybulski, Stephen Schueller, Konrad Kording, David Mohr.	112
Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach (e125) Jennifer Carroll, Anne Moorhead, Raymond Bond, William LeBlanc, Robert Petrella, Kevin Fiscella.	131

Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis (e132)	
Cheng Wen, Stefan Schneider, Arthur Stone, Donna Spruijt-Metz.	140
#WhyWeTweetMH: Understanding Why People Use Twitter to Discuss Mental Health Problems (e107)	
Natalie Berry, Fiona Lobban, Maksim Belousov, Richard Emsley, Goran Nenadic, Sandra Bucci.	154
Understanding Health Care Social Media Use From Different Stakeholder Perspectives: A Content Analysis of an Online Health Community (e109)	
Yingjie Lu, Yang Wu, Jingfang Liu, Jia Li, Pengzhu Zhang.	167
Pregnant Women's Views on the Feasibility and Acceptability of Web-Based Mental Health E-Screening Versus Paper-Based Screening: A Randomized Controlled Trial (e88)	
Dawn Kingston, Marie-Paule Austin, Sander Veldhuyzen van Zanten, Paula Harvalik, Rebecca Giallo, Sarah McDonald, Glenda MacQueen, Lydia Vermeijden, Gerri Lasiuk, Wendy Sword, Anne Biringier.	186
Seeking Web-Based Information About Attention Deficit Hyperactivity Disorder: Where, What, and When (e126)	
Sara Rosenblum, Elad Yom-Tov.	204
Usability, Acceptability, and Adherence to an Electronic Self-Monitoring System in Patients With Major Depression Discharged From Inpatient Wards (e123)	
Lise Lauritsen, Louise Andersen, Emilia Olsson, Stine Søndergaard, Lasse Nørregaard, Philip Løventoft, Signe Svendsen, Erik Frøkjær, Hans Jensen, Ida Hageman, Lars Kessing, Klaus Martiny.	219
Gender-Specific Determinants and Patterns of Online Health Information Seeking: Results From a Representative German Health Survey (e92)	
Eva Baumann, Fabian Czerwinski, Doreen Reifegerste.	231
Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey (e101)	
Clemens Ernsting, Stephan Dombrowski, Monika Oedekoven, Julie O'Sullivan, Melanie Kanzler, Adelheid Kuhlmeier, Paul Gellert.	247
Social Health Inequalities and eHealth: A Literature Review With Qualitative Synthesis of Theoretical and Empirical Studies (e136)	
Karine Latulippe, Christine Hamel, Dominique Giroux.	259
Influence of Telecommunication Modality, Internet Transmission Quality, and Accessories on Speech Perception in Cochlear Implant Users (e135)	
Georgios Mantokoudis, Roger Koller, Jérémie Guignard, Marco Caversaccio, Martin Kompis, Pascal Senn.	273
Importance-Performance Analysis of Personal Health Records in Taiwan: A Web-Based Survey (e131)	
Hsiao-Hsien Rau, Yi-Syuan Wu, Chi-Ming Chu, Fu-Chung Wang, Min-Huei Hsu, Chi-Wen Chang, Kang-Hua Chen, Yen-Liang Lee, Senyeong Kao, Yu-Lung Chiu, Hsyien-Chia Wen, Anis Fuad, Chien-Yeh Hsu, Hung-Wen Chiu.	285
The Multimodal Assessment of Adult Attachment Security: Developing the Biometric Attachment Test (e100)	
Federico Parra, Raphaële Miljkovitch, Gwenaëlle Persiaux, Michelle Morales, Stefan Scherer.	302
An Analysis of 2.3 Million Participations in the Continuing Medical Education Program of a General Medical Journal: Suitability, User Characteristics, and Evaluation by Readers (e49)	
Hildegard Christ, Jeremy Franklin, Reinhard Griebenow, Christopher Baethge.	319

Pedestrian Inattention Blindness While Playing Pokémon Go as an Emerging Health-Risk Behavior: A Case Report (e86)	
Stefania Barbieri, Gianna Vettore, Vincenzo Pietrantonio, Rossella Snenghi, Alberto Tredese, Mauro Bergamini, Sara Previato, Armando Stefanati, Rosa Gaudio, Paolo Feltracco.	331
Toward Game-Based Digital Mental Health Interventions: Player Habits and Preferences (e128)	
Regan Mandryk, Max Birk.	340
Analyzing and Predicting User Participations in Online Health Communities: A Social Support Perspective (e130)	
Xi Wang, Kang Zhao, Nick Street.	356
Use of Machine Learning Classifiers and Sensor Data to Detect Neurological Deficit in Stroke Patients (e120)	
Eunjeong Park, Hyuk-Jae Chang, Hyo Nam.	371
Ethics and Privacy Implications of Using the Internet and Social Media to Recruit Participants for Health Research: A Privacy-by-Design Framework for Online Recruitment (e104)	
Jacqueline Bender, Alaina Cyr, Luk Arbuckle, Lorraine Ferris.	384
A Content Analysis of Health and Safety Communications Among Internet-Based Sex Work Advertisements: Important Information for Public Health (e111)	
Julie Kille, Vicky Bungay, John Olliffe, Chris Atchison.	398
eHealth Search Patterns: A Comparison of Private and Public Health Care Markets Using Online Panel Data (e117)	
Janina Schneider, Christopher Holland.	410
A Psychometric Analysis of the Italian Version of the eHealth Literacy Scale Using Item Response and Classical Test Theory Methods (e114)	
Nicola Diviani, Alexandra Dima, Peter Schulz.	425
Impact of Information and Communication Technologies on Nursing Care: Results of an Overview of Systematic Reviews (e122)	
Geneviève Rouleau, Marie-Pierre Gagnon, José Côté, Julie Payne-Gagnon, Emilie Hudson, Carl-Ardy Dubois.	439
Web-Based Medical Appointment Systems: A Systematic Review (e134)	
Peng Zhao, Ilhoi Yoo, Jaie Lavoie, Beau Lavoie, Eduardo Simoes.	460
Timely Digital Patient-Clinician Communication in Specialist Clinical Services for Young People: A Mixed-Methods Study (The LYNC Study) (e102)	
Frances Griffiths, Carol Bryce, Jonathan Cave, Melina Dritsaki, Joseph Fraser, Kathryn Hamilton, Caroline Huxley, Agnieszka Ignatowicz, Sung Kim, Peter Kimani, Jason Madan, Anne-Marie Slowther, Mark Sujan, Jackie Sturt.	469
Is Provider Secure Messaging Associated With Patient Messaging Behavior? Evidence From the US Army (e103)	
Vickee Wolcott, Ritu Agarwal, D. Nelson.	484
Connecting Families to Their Health Record and Care Team: The Use, Utility, and Impact of a Client/Family Health Portal at a Children's Rehabilitation Hospital (e97)	
Gillian King, Joanne Maxwell, Amir Karmali, Simon Hagens, Madhu Pinto, Laura Williams, Keith Adamson.	493
Young People's Preferences for an Asthma Self-Management App Highlight Psychological Needs: A Participatory Study (e113)	
Dorian Peters, Sharon Davis, Rafael Calvo, Susan Sawyer, Lorraine Smith, Juliet Foster.	507

Corrigenda and Addenda

Authorship Correction: Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles ([e143](#))

Sohrab Saeb, Thaddeus Cybulski, Stephen Schueller, Konrad Kording, David Mohr. 520

Original Paper

Evidence of Benefit of Telerehabilitation After Orthopedic Surgery: A Systematic Review

Jose Manuel Pastora-Bernal^{1*}, MSc, B Physiotherapy, MBA; Rocio Martín-Valero^{2*}, PhD; Francisco Javier Barón-López^{1*}, PhD; María José Estebanez-Pérez^{3*}, B Physiotherapy

¹Faculty of Health Sciences, Physiotherapy, University of Málaga, Malaga, Spain

²Faculty of Nursery and Physiotherapy, Department of Nursery and Physiotherapy, University of Cadiz, Cadiz, Spain

³Rehabilitación Costa del Sol, Fisioterapia, Private Centre, Fuengirola, Spain

* all authors contributed equally

Corresponding Author:

Jose Manuel Pastora-Bernal, MSc, B Physiotherapy, MBA

Faculty of Health Sciences

Physiotherapy

University of Málaga

C/ Arquitecto Francisco Peñalosa Ampliación del C Teatinos 29071 Malaga

Ampliación del C Teatinos

Malaga, 29071

Spain

Phone: 34 658561622 ext 00

Fax: 34 952468989

Email: jmpastora@uma.es

Abstract

Background: In addition to traditional physiotherapy, studies based on telerehabilitation programs have published the results of effectiveness, validity, noninferiority, and important advantages in some neurological, cognitive, and musculoskeletal disorders, providing an opportunity to define new social policies and interventions.

Objectives: The aim of this systematic review is to investigate the effects of telerehabilitation after surgical procedures on orthopedic conditions as well as to describe how interventions are designed and to determine whether telerehabilitation is comparable with conventional methods of delivery. This systematic review summarizes the levels of evidence and grades of recommendation regarding telerehabilitation intervention (synchronous or asynchronous provided via the telerehabilitation medium, either in conjunction with, or in isolation of, other treatment interventions) after surgical procedures on orthopedic conditions.

Methods: Study quality was assessed using the Physiotherapy Evidence Database (PEDro) scores and grade of recommendation following the recommendation of the Oxford Centre for Evidence-Based Medicine.

Results: We found 3 studies with PEDro scores between 6 and 8, which is considered as level 1 evidence (good; 20% [3/15]), 4 studies with a score of 5, which is considered as level 2 evidence (acceptable; 27% [4/15]), and the remaining 8 studies had scores of 4 or less, which is considered (poor; 53% [8/15]). A total of 1316 participants received telerehabilitation intervention in the selected studies, where knee and hip replacement were 75% of all the studies. Strong and moderate grades of evidence (grade of recommendation A–B) were found in knee and hip replacement interventions. Studies on the upper limb were 25% of the studies, but only 1 study presented a moderate grade of evidence (grade of recommendation B) and the rest were of poor methodological quality with weak evidence (grade of recommendation C).

Conclusions: Conclusive evidence on the efficacy of telerehabilitation for treatment after an orthopedic surgery, regardless of pathology, was not obtained. We found strong evidence in favor of telerehabilitation in patients following total knee and hip arthroplasty and limited evidence in the upper limb interventions (moderate and weak evidence). Future research needs to be more extensive and conclusive. To the best of the authors' knowledge, this is the first attempt at evaluating the quality of telerehabilitation intervention research after surgical procedures on orthopedic conditions in a systematic review. Clinical messages and future research recommendations are included in the review.

KEYWORDS

telerehabilitation; orthopedic surgery; musculoskeletal disorders; systematic review; telemedicine; mobile health; mHealth; telehealth; physiotherapy

Introduction

The increasing availability of low-cost Internet and communication technologies has boosted the opportunity to apply technology-based solutions to provide health services during hospitalization and after discharge from hospital [1]. With the general ageing of the population—at least in industrialized countries—and the limited resources devoted to public health, the development of new rehabilitation models and practices seems mandatory in order to cope with the change in population needs [2,3]. Telerehabilitation, one of the emerging fields of telemedicine, is defined as a set of tools, procedures, and protocols to deliver the rehabilitation process remotely [4]. It will be increasingly important to develop scalable and sustainable telerehabilitation programs [5].

The growing demand for rehabilitation can result in increased costs and longer waiting lists, becoming a threat to the sustainability of health care services [6]. Telerehabilitation can help with this issue by discharging patients from points of care while improving their adherence to treatment [6]. Recent telerehabilitation studies have addressed different approaches: feasibility and efficacy, patients' and professionals' satisfaction, and cost analysis studies. This includes neurological diseases [7-10], stroke patients [11-13], intensive care unit [14], breast cancer [15], COPD [16], and musculoskeletal disorders [17-22]. Although trials supporting the role of telerehabilitation continue to emerge, implementation has been slow [23].

The scientific community believes that telerehabilitation will play an important role in improving, or at least maintaining, the continuity of rehabilitation care and services as they are reorganized, as it is able to increase the efficiency of the services while containing costs [24].

Early systematic reviews have pointed out that, despite the growing number of telerehabilitation experiences worldwide, evidence of clinical and economic effectiveness is still lacking [2,25]. In addition, these reviews have commented on the lack of methodological rigor and variability of approaches used in telerehabilitation studies [4]. With the growing number of telerehabilitation programs and with emerging databases providing potentially useful and reliable data on clinical outcomes [26], it is reasonable to expect that the number of scientific publications will rapidly increase [27], along with the number of systematic reviews on the topic [4]. Systematic reviews that included telerehabilitation interventions in musculoskeletal conditions yielded a few eligible trials, so efficacy remains inconclusive [1,28]. Telerehabilitation shows promise in many fields, but compelling evidence of the benefit for, and impact on, routine rehabilitation programs is still limited. There is a need for detailed, better-quality studies and for studies on the use of telerehabilitation in routine care [26].

Musculoskeletal injuries are frequent events in routine care and are the most common source of chronic pain and disability [29]. Orthopedic surgeries are experiencing some of the greatest growth rates in developed nations across the world. A study from 2014 found that the most common inpatient operating room procedure in the United States involved the musculoskeletal system [30]. Therapeutic exercises are commonly prescribed following a surgery in an attempt to maximize functional outcome [31]. Physical therapists have been utilizing therapeutic exercises since the conception of the profession, and they have been demonstrated to be fundamental in improving function, performance, and disability [32]. Telerehabilitation offers the possibility to develop therapeutic exercise at a distance, among others.

Recently, a systematic review concluded that telerehabilitation is promising and highlighted the fact that for those individuals who are unable to attend traditional face-to-face services, particularly following elective orthopedic surgical procedures [33], telerehabilitation should be considered as a viable option in the holistic management of their musculoskeletal condition [33].

Despite the existence of systematic reviews on telerehabilitation interventions, none of them has explored post-surgical rehabilitation in musculoskeletal injuries. Therefore, the aim of this paper is to investigate the effects of telerehabilitation after surgical procedures on orthopedic conditions, as well as to describe how interventions are designed and to determine whether telerehabilitation is comparable with conventional methods of delivery within this population. We have considered all forms of interventions that use telecommunications technology to telerehabilitation interventions.

Methods

Search Strategy and Eligibility Criteria

This review has been carried out following the PRISMA 2009 guidelines [34]. The review protocol was registered with an international registration database [PROSPERO, Registration Number: CRD42016047846].

As with most of the recent systematic reviews on the topic [2,26,27,33,35], the following literature searches were performed to identify all possible studies that could help answer the research question. MEDLINE, Physiotherapy Evidence Database (PEDro), Scopus, PsycINFO, Web of Knowledge, CINAHL, SPORTDiscus, Directory of Open Access Journals (DOAJ), Cochrane, Embase, Academic Search Complete, Fuente Académica, and Consejo Superior de Investigaciones Científicas (CSIC) were searched. In addition, the search was performed in a relevant bibliographic database from the University of Málaga.

The initial search was carried out in June 2016, and was completed with a new search to update the review in October 2016. The following combinations of keywords were used: telerehabilitation, telerehabilitation physiotherapy, post-surgery telerehabilitation, musculoskeletal disorders, systematic review, telemedicine (mobile health or health, mobile or mHealth or telehealth or eHealth), telemedicine physiotherapy program. The limits of searches were studies published between 2000 and 2016.

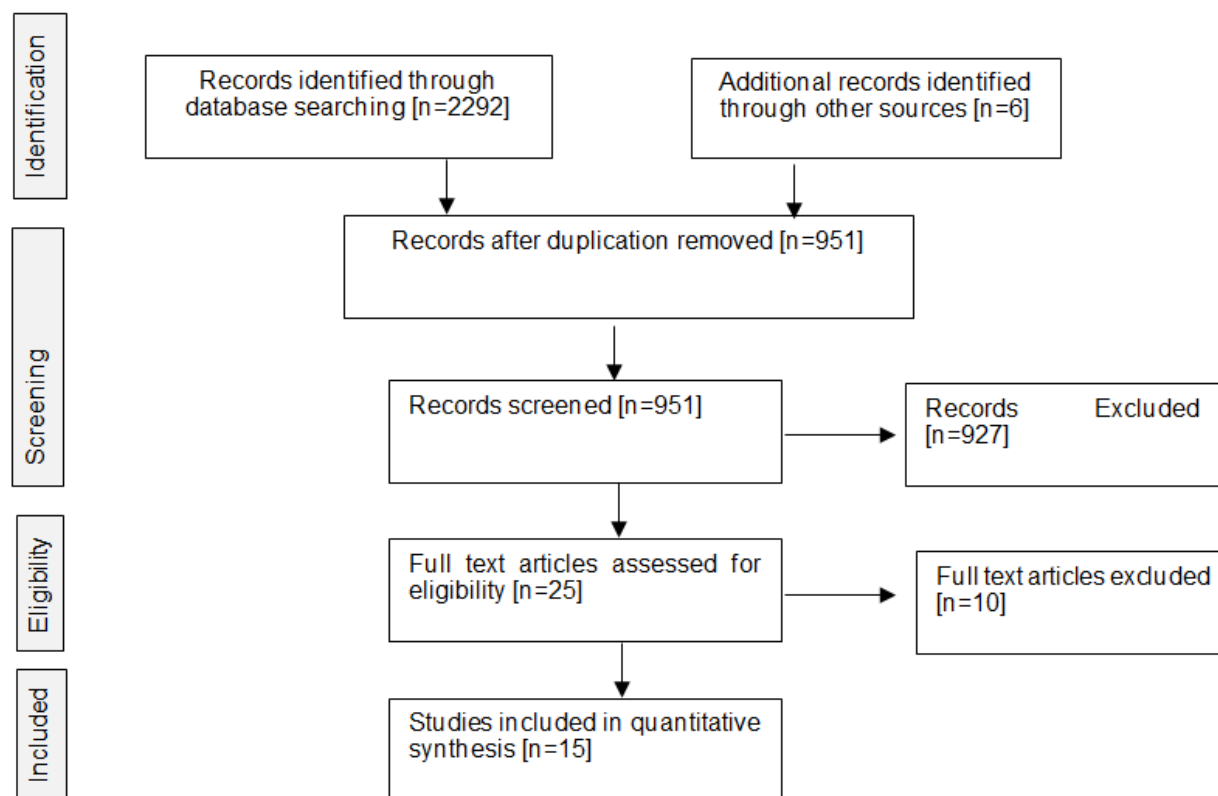
The main steps related to the search phase are reported in [Figure 1](#) using the PRISMA flow diagram [36]. After the application of the selected keywords, the entire set of records was analyzed to identify eventual duplication of articles retrieved from

different sources; the remaining articles were then assessed in full text for eligibility so as to identify all those matching the inclusion criteria (see [Multimedia Appendix 1](#)).

Two authors (JMPB and RMV) independently screened the titles and abstracts of all records retrieved using the database search strategy. The full text was obtained if further information was required to determine eligibility, or if uncertainty prevailed between authors. For trials published in a language other than English or Spanish, a translated version of the abstract was sourced to determine eligibility.

Disagreements between authors were initially resolved via discussion, and then by consultation with a third reviewer (FJBL).

Figure 1. Flowchart.



Eligibility Criteria: Inclusion and Exclusion Criteria

Eligibility criteria were based on the PICOS framework [37], as follows:

Participants

Adults (≥ 18 years) with telerehabilitation services after surgical procedures as a result of a primary orthopedic condition. Trials in which the participant's condition was secondary to a diagnosed health condition that was not primarily musculoskeletal in nature (eg, hand or shoulder dysfunction following stroke) were excluded.

Intervention

Any treatment intervention, synchronous or asynchronous, provided via a telerehabilitation medium, phone counseling, interactive virtual system, or gaming, either in conjunction with, or in isolation of, other treatment interventions was included.

Comparison

All trials were required to have a comparison group (of the same condition), where options included (but were not restricted to) face-to-face treatment or usual care. The comparison group could not be an alternative form of telerehabilitation. A pilot clinical trial without a comparison group was included if a telerehabilitation intervention had been carried out among participants.

Outcomes

Any clinical outcome, including measurements based on pain, quality of life, disability or function (physical, social, or psychological), was analyzed. Economic and cost-utility outcomes were not analyzed, nor were patient and clinician satisfaction or those outcomes measuring adherence to, or compliance with, rehabilitation programs.

Study Design

All types of study designs were considered: randomized clinical trials (RCT), clinical trials (CT), case reports, controlled clinical trial, and pilot study. Articles that were limited to describing the feasibility and fit-out of telerehabilitation interventions were excluded.

For all eligible trials, data extraction was independently completed by 2 authors (JMPB and MJEP), and was cross-checked for consistency by a third author (RMV). The primary authors of eligible trials were contacted when information was considered to be missing for either the quality assessment or data extraction process.

Evaluation of Methodological Quality, Level of Evidence, and Grade of Recommendation

An important step in conducting a systematic review is to assess the methodological quality of each included trial. In addition, reporting methodological quality provides clinicians with information about whether the results of clinical trials should influence their clinical practice. A valid way of assessing the methodological quality of clinical trials is therefore essential [38].

Two independent reviewers [JMPB and RMV] completed the checklist based on the PEDro scale. The methodological quality and risk of bias were evaluated using the PEDro scale [38,39]

based on the Delphi list [40]. It is considered a useful tool for carrying out the assessment methodology in scientific research.

The PEDro scale scores 10 items: random allocation, concealed allocation, similarity at baseline, subject blinding, therapist blinding, assessor blinding, >85% follow-up for at least one key outcome, intention-to-treat analysis, between-group statistical comparison for at least one key outcome, and point and variability measures for at least one key outcome. Items are scored as either present (1) or absent (0) and a score out of 10 is obtained by summation [38]. The scale includes an additional item (eligibility criteria) to evaluate the external validity, but the score is not counted.

According to Moseley et al, studies with a PEDro score ≥ 5 will be considered at low risk of bias and high methodological quality [41].

A study with a PEDro score of ≥ 6 is considered to have level 1 evidence (6-8=good, 9-10=excellent) and a study with a score of ≤ 5 is considered to have level 2 evidence (4-5=acceptable, <4=poor) [42].

Levels of evidence help us target the search at the type of evidence that is most likely to provide a reliable answer. They have been designed so that they can be used as a shortcut for busy clinicians, researchers, or patients to find the likely best evidence [43]. Grades of recommendation describe the strength and therefore value of the evidence relative to how rigorous the study was (see Tables 1 and 2) [44].

Table 1. Based on evidence-based medicine working group [44].

Grades of recommendation		Strength of evidence
A	Strong Evidence	A preponderance of level I and/or level II studies support the recommendation. This must include at least 1 level I study.
B	Moderate Evidence	A single high-quality randomized controlled trial or a preponderance of level II studies support the recommendation
C	Weak Evidence	A single level II study or a preponderance of level III and IV studies including statements of consensus by content experts support the recommendation
D	Conflicting Evidence	Higher-quality studies conducted on this topic disagree with respect to their conclusions. The recommendation is based on these conflicting studies
E	Theoretical/Foundational Evidence	A preponderance of evidence from animal or cadaver studies, from conceptual models/principles, or from basic sciences/bench research support this conclusion
F	Expert Opinion	Best practice based on the clinical Experience of the guidelines development team

Table 2. Based on grades of recommendation and levels of evidence for therapy or prevention. Material adapted from the recommendations at the center for evidence-based medicine in oxford [43].

Level of evidence	Strength of evidence
1a	Systematic review of (homogeneous) randomized controlled trials
1b	Individual randomized controlled trials (with narrow CIs)
2a	Systematic review of (homogeneous) cohort studies of “exposed” and “unexposed” subject
2b	Individual cohort study / Low-quality randomized controlled trials
3a	Systematic review of (homogeneous) case-control studies
3b	Individual case-control studies
4	Case Series, low-quality cohort or case-control studies
5	Expert opinion based on non systematic reviews of results or mechanistic studies

Results

The main findings of this review are presented in [Table 3](#), an evaluation of the methodological quality of the 15 studies selected according to the PEDro scale. Characteristics of the included studies are listed in [Multimedia Appendix 1](#) showing the grades of recommendation, regarding the effectiveness, results, and effect size in the different outcomes of telerehabilitation services. A subgroup analysis by population and intervention is presented in [Tables 4](#) and [5](#).

Studies included in the review had PEDro scores of 2-8, as shown in [Table 3](#). Studies were considered of high enough methodological quality if they had a score of at least 5. This was based on the fact that studies with a score closer to 4 did not use a triple-blind methodology (subject, evaluator, and treatment provider) [45].

We found 3 studies [20,46,47] with PEDro scores between 6 and 8, which is considered level 1 evidence (good; 20% [3/15]), 4 studies [48-50,51] with a score of 5, which is considered level 2 evidence (acceptable; 27% [4/15]), and the remaining 8 studies [52-56,21,57,58] had scores of 4 or less, which is considered (poor; 53% [8/15]).

A total of 1316 participants received telerehabilitation intervention. Strong and moderate grades of evidence (grade of recommendation A-B) were found in knee and hip replacement interventions (80% of all the studies). Studies in the upper limb were 20% of the studies included but only 1 study presented a moderate grade of evidence (grade or recommendation B) and the rest were of poor methodological quality with weak evidence (grade of recommendation C).

The subgroup analysis by population shows us 8 articles focused on a total knee replacement population (50% of articles), 4 on total hip replacement (25% of articles), 1 on shoulder joint replacement (6% of articles), 1 on proximal humerus fractures

(6% of articles), 1 on carpal tunnel release surgery (6% of articles), and 1 on rotator cuff tear (6% of articles). Another 3 telerehabilitation publications in the upper limbs (hand transplantation [case study]) [59], and hand surgery (preclinical trials and descriptive study) [60,61], have been identified in addition, but as no intervention on the subject has been published, they do not match the inclusion criteria for this review (See [Multimedia Appendix 2](#)).

In the subgroup analysis by telerehabilitation intervention ([Table 5](#)), we found strong evidence regardless of the intervention (videoconferencing, asynchronous, phone counseling, interactive virtual system and gaming, telerehabilitation system). Therefore, we may interpret that evidence level and degree of recommendation are related to studies design, and are not related to intervention.

We found great heterogeneity among the included studies. Sample size ranged between 5 [56] and 237 [53]. Telerehabilitation interventions included videoconferencing sessions [20,21,46,50,55], phone counseling [49,53], video games [62,63], asynchronous exercise videos, and interactive virtual systems [19,47,52,54-57].

Compared interventions also range between clinical protocol face-to-face physiotherapy [19,46,50,52,55], home physiotherapy visits and usual care at home [20,53,54], and a physiotherapy session followed by gaming [63].

Intervention duration ranged from 2 weeks [19] to 26 weeks [53] with follow-up periods from 13 sessions [56] to 9 months [49].

We found some homogeneous aspects in clinical outcomes primarily in the areas of function, quality of life, and specific daily life activities [20,46,54,63], and less assessment of disability (passive and active ranges of motion, balance, and muscle strength) and pain.

Table 3. Evaluation of methodological quality of the 15 selected studies.

PEDro scale criteria	Moffet et al [20]	Rus-sell et al [46]	Bini et al [47]	Pi-queras et al [51]	Tousig-nant et al [50]	Hør-dam et al [49]	Fung et al [48]	Li et al [53]	Rus-sell et al [54]	Eriks-son et al [55]	Eiser-mann et al [52]	Antón et al [58]	Tousig-nant et al [64]	Heuser et al [56]	Ma-cias et al [57]
Eligibility criteria ^a	Y ^b	Y	Y	Y	Y	Y	Y	Y	N ^c	Y	Y	Y	Y	Y	Y
Randomiza-tion	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	N	N	N
Allocation concealed	Y	Y	Y	N	Y	Y	N	N	N	N	N	N	N	N	N
Baseline comparabili-ty	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	N	N	N
Subject blinding	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Therapist blinding	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Evaluator blinding	Y	Y	N	Y	N	N	Y	N	Y	N	N	N	N	N	N
Appropriate continuation	Y	Y	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Intention to treat	Y	Y	Y	N	N	N	N	N	N	Y	N	N	N	N	N
Comparison between groups	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N
Specific measure-ments and variability	Y	Y	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	Y
Total PEDro Score	8	8	7	5	5	5	5	4	4	4	3	2	2	2	2

^aThe eligibility criteria do not contribute to the total score.^bY is Yes.^cN is No.

Table 4. Subgroup analysis by population.

Population	Authors and reference	Number of articles	Participants (n)	Articles with grade of recommendation A Level of evidence 1 (% of total articles)	Articles with grade of recommendation B Level of evidence 2 or 3 (% of total articles)	Articles with grade of recommendation C and D Level of evidence >3 (% of total articles)
Total knee arthroplasty	Moffet et al 2015 [20], Russell et al 2011 [46], Bini et al 2016 [47], Piqueras et al 2013 [19], Tousignant et al 2011 [50], Eisermann et al 2004 [52], Fung et al 2012 [48], Russell et al 2003 [54]	8	718	3 (19)	5 (31)	0
Total Hip Replacement	Hørdam et al 2009 [49], Eisermann et al 2004 [52], Li et al 2014 [53], Antón et al 2016 [58]	4	543	3 (19)	0	1 (6.25)
Shoulder joint replacement	Eriksson et al 2009 [55]	1	22	0	1 (6.25)	0
Proximal humerus fractures	Tousignant et al 2015 [64]	1	17	0	0	1 (6.25)
Carpal tunnel release surgery	Heuser et al 2007 [56]	1	5	0	0	1 (6.25)
Rotator Cuff Tear	Macías-Hernández et al 2016 [57]	1	11	0	0	1 (6.25)
Total %		16 (Eisermann et al included knee and hip population)	1316	38.00	37.25	25

Table 5. Subgroup analysis by intervention.

Intervention	Authors and reference	Number of articles	Participants (n)	Articles with level of evidence 1 and grade of recommendation A (% of total articles)	Articles with level of evidence 2 or 3 and grade of recommendation B (% of total articles)	Articles with level of evidence >3 and grade of recommendation C and D (% of total articles)
Videoconferencing (real-time)	Moffet et al 2015 [20], Russell et al 2011 [46], Tousignant et al 2011 [50], Eriksson et al 2009 [55], Tousignant et al 2015 [64]	5	357	3 (21.4)	1 (6.6)	1 (6.6)
Asynchronous videos program	Bini et al 2016 [47], Eisermann et al 2004 [52], Macías-Hernández et al 2016 [57]	3	336	2 (14.3)	0	1 (6.6)
Education sessions by telephone	Hørdam et al 2009 [49], Li et al 2014 [53]	2	398	2 (14.3)	0	0
Interactive virtual TR system & gaming	Piqueras et al 2013 [19], Fung et al 2012 [48], Russell et al 2003 [54], Antón et al 2016 [58], Heuser et al 2007 [56]	5	225	2 (14.3)	1 (6.6)	2 (13.3)
Total %		15	1316	60	13.33	26.67

Discussion

Principal Findings

This review confirms the strong evidence in favor of telerehabilitation among patients undergoing total knee and hip arthroplasty and the limited evidence in the upper limb (moderate and weak evidence).

To the best of our knowledge, this is the first review focused on telerehabilitation research after surgical procedures on orthopedic conditions. This systematic review applied a qualitative evaluation to provide a wider picture of currently available evidence.

First, we will discuss the contributions of the first systematic reviews on the topic. Second, we discuss the generalizations, previous results, and future recommendations. Third, we discuss if results are extrapolated to the upper limb (the results of this review show poor-quality methodology and moderate and weak evidence). Finally, we discuss about the inherent difficulties in conducting telerehabilitation research and future research recommendations.

Regarding the first aspect noted in the discussion, the first systematic reviews, contributions concluded that better-quality studies are needed as well as studies on the use of telerehabilitation in routine care. Telerehabilitation research is generally not very good and there are many reviews that criticize this [2,26,27,33,35]. In our review, 60% of the included studies are of poor methodological quality with weak evidence.

Regarding the second aspect noted in the discussion, the most recent systematic reviews provide statements on the effects of telerehabilitation interventions. Two recent studies provide statements such as, “there is a strong positive effect for patients following orthopedic surgery” [1] and “there is unequivocal evidence that the management of musculoskeletal conditions via real-time telerehabilitation is effective in improving physical function, disability, and pain” [33]. We agree with these statements, but only and exclusively for some pathologies. Our systematic review shows that there is still insufficient evidence on upper limb surgeries’ telerehabilitation interventions.

Regarding the third aspect noted in the discussion: does this statement transfer to upper limb such as shoulder arthroscopy, carpal tunnel release surgery, hand surgery, or shoulder arthroplasty? Could this be extrapolated to the rehabilitation process in fractures or surgery interventions in upper and lower limbs? These unresolved clinical questions reaffirm the need to identify the available evidence in post-surgical rehabilitation with telerehabilitation interventions.

For this systematic review, we seek to find evidence of post-surgical telerehabilitation programs, with special emphasis on programs that can be integrated into clinical practice.

In our review, telerehabilitation research in the upper limb (shoulder joint replacement, proximal humerus fractures, carpal tunnel release, and cuff rotator tears) presents moderate and weak levels of evidence. Notable is the judgment that none of the telerehabilitation studies in the upper limb included in this review present a high level of evidence and recommendation.

There is still a very small database for telerehabilitation studies after a musculoskeletal surgery that provides useful data on clinical outcomes, especially in conditions other than the replacement of joints in the lower limb. Therefore, conclusive evidence on the efficacy of telerehabilitation for treatment after an orthopedic surgery, regardless of pathology, was not obtained.

Research background has been used to discuss the strengths of telerehabilitation and the opportunities for future interventions and policies. Regarding the final aspect noted in the discussion, what are the inherent difficulties in conducting telerehabilitation research?

During the search, we observed a number of studies that provide descriptions of telerehabilitation interventions of low methodological quality. No validated clinical outcomes, too small a sample size, and a lack of comparison group are frequently found. Moreover, differences in telerehabilitation interventions, treatment period, and follow-up, create doubts in identifying whether the telerehabilitation gives comparable or better results.

A frequent problem in studies of telerehabilitation is the lack of blinding of therapists and patients. There is evidence that in clinical trials where allocation is not concealed and assessors, therapists, and participants are not blinded, a larger effect of intervention is reported than in higher quality trials with adequate blinding procedures [65].

It may be that there are good-quality studies the publication of which has been delayed; however, our findings are aligned and consistent with the most recent revisions regarding the need for future research needs to have stronger and more solid studies.

One of the biases identified is that telerehabilitation groups have more frequent contact with health professionals and with the intervention (especially in videoconferencing and phone contact), so they are likely to receive additional services. This creates biases whether the positive results are related to a more elaborate program than really with the interventions method.

How could this be addressed in future research? As blinding of patients and therapists is not possible in telerehabilitation interventions, several methodological aspects are fundamental for future research.

Telerehabilitation interventions should be conceptualized, coded, classified, and grouped in a similar way to physiotherapy technique codes, enabling identification in detail when the effect is due to the type of intervention. Comparison group must be the actual best evidence treatment for the same condition that allows identification of whether telerehabilitation offers better or comparable outcomes. Telerehabilitation frequency must be the same as the control group to avoid biases related to a more elaborate program. Greater homogeneous is needed especially in terms of type, duration, and intervention follow-up for each specific pathologies. Studies that show negative results should be published, avoiding publication biases. Large sample size and improvement in study quality (allocated and evaluator blinding) must be addressed. Orthopedic conditions and musculoskeletal injuries different to replacement joints in lower limbs need quality research.

Clinical Messages

High-methodological-quality studies should be conducted to confirm that telerehabilitation shows clinically relevant outcomes after surgery in orthopedic and musculoskeletal injuries, especially in upper limbs. Telerehabilitation appears to be an effective alternative to face-to-face service delivery after hospital discharge of patients following total knee arthroplasty and hip replacement. Clinical outcomes are comparable and not inferior. Despite some limitations, there seem to be clear benefits from physiotherapy at a distance

regardless with the telerehabilitation technique it offers (videoconferencing, phone intervention, asynchronous video exercise programs, or gaming). Future challenges include identifying whether positive results are due to the type of intervention or the increased frequency and intensity that telerehabilitation allow.

Future research recommendations for telerehabilitation should include high-quality studies with clear conclusions and statements that could improve health interventions and health policies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of the included studies.

[[PDF File \(Adobe PDF File\), 45KB - jmir_v19i4e142_app1.pdf](#)]

Multimedia Appendix 2

Excluded articles.

[[PDF File \(Adobe PDF File\), 45KB - jmir_v19i4e142_app2.pdf](#)]

References

1. Agostini M, Moja L, Banzi R, Pistotti V, Tonin P, Venneri A, et al. Telerehabilitation and recovery of motor function: a systematic review and meta-analysis. *J Telemed Telecare* 2015 Jun;21(4):202-213. [doi: [10.1177/1357633X15572201](#)] [Medline: [25712109](#)]
2. Rogante M, Grigioni M, Cordella DG, Giacomozzi C. Ten years of telerehabilitation: a literature overview of technologies and clinical applications. *NeuroRehabilitation* 2010;27(4):287-304. [doi: [10.3233/NRE-2010-0612](#)] [Medline: [21160118](#)]
3. Zampolini M, Todeschini E, Bernabeu GM, Hermens H, Ilsbrouckx S, Macellari V, et al. Tele-rehabilitation: present and future. *Ann Ist Super Sanita* 2008;44(2):125-134 [FREE Full text] [Medline: [18660562](#)]
4. Rogante M, Kairy D, Giacomozzi C, Grigioni M. A quality assessment of systematic reviews on telerehabilitation: what does the evidence tell us? *Ann Ist Super Sanita* 2015;51(1):11-18 [FREE Full text] [doi: [10.4415/ANN_15_01_04](#)] [Medline: [25857379](#)]
5. Dinesen B, Nonnecke B, Lindeman D, Toft E, Kidholm K, Jethwani K, et al. Personalized telehealth in the future: a global research agenda. *J Med Internet Res* 2016 Mar 01;18(3):e53 [FREE Full text] [doi: [10.2196/jmir.5257](#)] [Medline: [26932229](#)]
6. Ruiz-Fernandez D, Marín-Alonso O, Soriano-Paya A, García-Pérez JD. eFisioTrack: a telerehabilitation environment based on motion recognition using accelerometry. *ScientificWorldJournal* 2014;2014:495391 [FREE Full text] [doi: [10.1155/2014/495391](#)] [Medline: [24526898](#)]
7. Houlihan BV, Jette A, Friedman RH, Paasche-Orlow M, Ni P, Wierbicky J, et al. A pilot study of a telehealth intervention for persons with spinal cord dysfunction. *Spinal Cord* 2013 Sep;51(9):715-720. [doi: [10.1038/sc.2013.45](#)] [Medline: [23752260](#)]
8. Jelcic N, Agostini M, Meneghello F, Bussè C, Parise S, Galano A, et al. Feasibility and efficacy of cognitive telerehabilitation in early Alzheimer's disease: a pilot study. *Clin Interv Aging* 2014;9:1605-1611 [FREE Full text] [doi: [10.2147/CIA.S68145](#)] [Medline: [25284993](#)]
9. Paul L, Coulter EH, Miller L, McFadyen A, Dorfman J, Mattison PG. Web-based physiotherapy for people moderately affected with Multiple Sclerosis: quantitative and qualitative data from a randomized, controlled pilot study. *Clin Rehabil* 2014 Sep;28(9):924-935. [doi: [10.1177/0269215514527995](#)] [Medline: [24691218](#)]
10. Hoffmann T, Russell T, Thompson L, Vincent A, Nelson M. Using the internet to assess activities of daily living and hand function in people with Parkinson's disease. *NeuroRehabilitation* 2008;23(3):253-261. [Medline: [18560142](#)]
11. Lin K, Chen C, Chen Y, Huang W, Lai J, Yu S, et al. Bidirectional and multi-user telerehabilitation system: clinical effect on balance, functional activity, and satisfaction in patients with chronic stroke living in long-term care facilities. *Sensors (Basel)* 2014 Jul 11;14(7):12451-12466 [FREE Full text] [doi: [10.3390/s140712451](#)] [Medline: [25019632](#)]

12. Langan J, Delave K, Phillips L, Pangilinan P, Brown SH. Home-based telerehabilitation shows improved upper limb function in adults with chronic stroke: a pilot study. *J Rehabil Med* 2013 Feb;45(2):217-220 [[FREE Full text](#)] [doi: [10.2340/16501977-1115](https://doi.org/10.2340/16501977-1115)] [Medline: [23319181](#)]
13. Deng H, Durfee WK, Nuckley DJ, Rheude BS, Severson AE, Skluzacek KM, et al. Complex versus simple ankle movement training in stroke using telerehabilitation: a randomized controlled trial. *Phys Ther* 2012 Feb;92(2):197-209 [[FREE Full text](#)] [doi: [10.2522/ptj.20110018](https://doi.org/10.2522/ptj.20110018)] [Medline: [22095209](#)]
14. Jackson JC, Ely EW, Morey MC, Anderson VM, Denne LB, Clune J, et al. Cognitive and physical rehabilitation of intensive care unit survivors: results of the RETURN randomized controlled pilot investigation. *Crit Care Med* 2012 Apr;40(4):1088-1097 [[FREE Full text](#)] [doi: [10.1097/CCM.0b013e3182373115](https://doi.org/10.1097/CCM.0b013e3182373115)] [Medline: [22080631](#)]
15. Galiano-Castillo N, Ariza-García A, Cantarero-Villanueva I, Fernández-Lao C, Díaz-Rodríguez L, Legerén-Alvarez M, et al. Telehealth system (e-CUIDATE) to improve quality of life in breast cancer survivors: rationale and study protocol for a randomized clinical trial. *Trials* 2013 Jun 22;14:187 [[FREE Full text](#)] [doi: [10.1186/1745-6215-14-187](https://doi.org/10.1186/1745-6215-14-187)] [Medline: [23799886](#)]
16. Bedra M, McNabney M, Stiassny D, Nicholas J, Finkelstein J. Defining patient-centered characteristics of a telerehabilitation system for patients with COPD. *Stud Health Technol Inform* 2013;190:24-26. [Medline: [23823363](#)]
17. Bennell KL, Rini C, Keefe F, French S, Nelligan R, Kasza J, et al. Effects of adding an internet-based pain coping skills training protocol to a standardized education and exercise program for people with persistent hip pain (HOPE trial): randomized controlled trial protocol. *Phys Ther* 2015 Oct;95(10):1408-1422 [[FREE Full text](#)] [doi: [10.2522/ptj.20150119](https://doi.org/10.2522/ptj.20150119)] [Medline: [26023213](#)]
18. Truter P, Russell T, Fary R. The validity of physical therapy assessment of low back pain via telerehabilitation in a clinical setting. *Telemed J E Health* 2014 Feb;20(2):161-167. [doi: [10.1089/tmj.2013.0088](https://doi.org/10.1089/tmj.2013.0088)] [Medline: [24283249](#)]
19. Piqueras M, Marco E, Coll M, Escalada F, Ballester A, Cinca C, et al. Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med* 2013 Apr;45(4):392-396 [[FREE Full text](#)] [doi: [10.2340/16501977-1119](https://doi.org/10.2340/16501977-1119)] [Medline: [23474735](#)]
20. Moffet H, Tousignant M, Nadeau S, Mérette C, Boissy P, Corriveau H, et al. In-Home telerehabilitation compared with face-to-face rehabilitation after total knee arthroplasty: a noninferiority randomized controlled trial. *J Bone Joint Surg Am* 2015 Jul 15;97(14):1129-1141. [doi: [10.2106/JBJS.N.01066](https://doi.org/10.2106/JBJS.N.01066)] [Medline: [26178888](#)]
21. Tousignant M, Giguère AM, Morin M, Pelletier J, Sheehy A, Cabana F. In-home telerehabilitation for proximal humerus fractures: a pilot study. *Int J Telerehabil* 2014;6(2):31-37 [[FREE Full text](#)] [doi: [10.5195/ijt.2014.6158](https://doi.org/10.5195/ijt.2014.6158)] [Medline: [25945227](#)]
22. Bedra M, Finkelstein J. Feasibility of post-acute hip fracture telerehabilitation in older adults. *Stud Health Technol Inform* 2015;210:469-473. [Medline: [25991191](#)]
23. Shulver W, Killington M, Morris C, Crotty M. 'Well, if the kids can do it, I can do it': older rehabilitation patients' experiences of telerehabilitation. *Health Expect* 2017 Feb;20(1):120-129. [doi: [10.1111/hex.12443](https://doi.org/10.1111/hex.12443)] [Medline: [26890025](#)]
24. Theodoros D, Russell T. Telerehabilitation: current perspectives. *Stud Health Technol Inform* 2008;131:191-209. [Medline: [18431862](#)]
25. Kairy D, Lehoux P, Vincent C, Visintin M. A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation. *Disabil Rehabil* 2009;31(6):427-447. [doi: [10.1080/09638280802062553](https://doi.org/10.1080/09638280802062553)] [Medline: [18720118](#)]
26. Hailey D, Roine R, Ohinmaa A, Dennett L. Evidence of benefit from telerehabilitation in routine care: a systematic review. *J Telemed Telecare* 2011;17(6):281-287. [doi: [10.1258/jtt.2011.101208](https://doi.org/10.1258/jtt.2011.101208)] [Medline: [21844172](#)]
27. Laver KE, Schoene D, Crotty M, George S, Lannin NA, Sherrington C. Telerehabilitation services for stroke. *Cochrane Database Syst Rev* 2013(12):CD010255. [doi: [10.1002/14651858.CD010255.pub2](https://doi.org/10.1002/14651858.CD010255.pub2)] [Medline: [24338496](#)]
28. Pietrzak E, Cotea C, Pullman S, Nasveld P. Self-management and rehabilitation in osteoarthritis: is there a place for internet-based interventions? *Telemed J E Health* 2013 Oct;19(10):800-805. [doi: [10.1089/tmj.2012.0288](https://doi.org/10.1089/tmj.2012.0288)] [Medline: [24003979](#)]
29. Russell TG, Blumke R, Richardson B, Truter P. Telerehabilitation mediated physiotherapy assessment of ankle disorders. *Physiother Res Int* 2010 Sep;15(3):167-175. [doi: [10.1002/pri.471](https://doi.org/10.1002/pri.471)] [Medline: [20812313](#)]
30. Fingar K, Stocks C, Weiss A, Steiner C. HCUP Statistical Brief # 186 Internet. Rockville, MD: Agency for Healthcare Research and Quality; 2014. Most frequent operating room procedures performed in U.S. hospitals, 2003-2012 URL: <https://www.hcup-us.ahrq.gov/reports/statbriefs/sb186-Operating-Room-Procedures-United-States-2012.pdf> [accessed 2017-04-18] [[WebCite Cache ID 6poJNmFkt](#)]
31. Eriksson L, Lindström B, Ekenberg L. Patients' experiences of telerehabilitation at home after shoulder joint replacement. *J Telemed Telecare* 2011;17(1):25-30. [doi: [10.1258/jtt.2010.100317](https://doi.org/10.1258/jtt.2010.100317)] [Medline: [21075802](#)]
32. Saunders DG. Therapeutic exercise. *Clin Tech Small Anim Pract* 2007 Nov;22(4):155-159. [Medline: [18198783](#)]
33. Cottrell MA, Galea OA, O'Leary SP, Hill AJ, Russell TG. Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: A systematic review and meta-analysis. *Clin Rehabil* 2016 May 2:- Epub ahead of print. [doi: [10.1177/0269215516645148](https://doi.org/10.1177/0269215516645148)] [Medline: [27141087](#)]

34. 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. *BMJ* 2009;339:b2700 [[FREE Full text](#)] [Medline: [19622552](#)]
35. Hailey D, Roine R, Ohinmaa A, Dennett L. The status of telerehabilitation in neurological applications. *J Telemed Telecare* 2013 Sep;19(6):307-310. [doi: [10.1177/1357633X13501775](#)] [Medline: [24163293](#)]
36. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015 Jan 02;349:g7647 [[FREE Full text](#)] [Medline: [25555855](#)]
37. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare. Laverthorpe, York: CRD, University of York; Jan 2009:1-281.
38. de Morton N. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Aust J Physiother* 2009;55(2):129-133 [[FREE Full text](#)] [doi: [10.1016/S0004-9514\(09\)70043-1](#)]
39. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 2003 Aug;83(8):713-721. [Medline: [12882612](#)]
40. Verhagen A, de Vet HC, de Bie RA, Kessels A, Boers M, Bouter L, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 1998 Dec;51(12):1235-1241. [Medline: [10086815](#)]
41. Moseley A, Herbert R, Sherrington C, Maher C. Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). *Aust J Physiother* 2002;48(1):43-49 [[FREE Full text](#)] [doi: [10.1016/S0004-9514\(14\)60281-6](#)]
42. Foley NC, Teasell RW, Bhogal SK, Speechley MR. Stroke rehabilitation evidence-based review: methodology. *Top Stroke Rehabil* 2003;10(1):1-7. [Medline: [12970828](#)]
43. CEBM. 2011. Levels of evidence: introductory document URL: <http://www.cebm.net/2011-oxford-cebm-levels-evidence-introductory-document/> [accessed 2017-04-18] [[WebCite Cache ID 6poLH7ZV6](#)]
44. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 1995 Dec 13;274(22):1800-1804. [Medline: [7500513](#)]
45. Martin-Valero R, De La Casa AM, Casuso-Holgado M, Heredia-Madrado A. Systematic review of inspiratory muscle training after cerebrovascular accident. *Respir Care* 2015;60(11):1652-1659. [doi: [10.4187/respcare.03981](#)]
46. Russell T, Buttrum P, Wootton R, Jull G. Internet-based outpatient telerehabilitation for patients following total knee arthroplasty: a randomized controlled trial. *J Bone Joint Surg Am* 2011 Jan 19;93(2):113-120. [doi: [10.2106/JBJS.I.01375](#)] [Medline: [21248209](#)]
47. Bini SA, Mahajan J. Clinical outcomes of remote asynchronous telerehabilitation are equivalent to traditional therapy following total knee arthroplasty: A randomized control study. *J Telemed Telecare* 2017 Feb;23(2):239-247. [doi: [10.1177/1357633X16634518](#)] [Medline: [26940798](#)]
48. Fung V, Ho A, Shaffer J, Chung E, Gomez M. Use of Nintendo Wii Fit™ in the rehabilitation of outpatients following total knee replacement: a preliminary randomised controlled trial. *Physiotherapy* 2012 Sep;98(3):183-188. [doi: [10.1016/j.physio.2012.04.001](#)] [Medline: [22898573](#)]
49. Hørdam B, Sabroe S, Pedersen PU, Mejdahl S, Søballe K. Nursing intervention by telephone interviews of patients aged over 65 years after total hip replacement improves health status: a randomised clinical trial. *Scand J Caring Sci* 2010 Mar;24(1):94-100. [doi: [10.1111/j.1471-6712.2009.00691.x](#)] [Medline: [19422632](#)]
50. Tousignant M, Moffet H, Boissy P, Corriveau H, Cabana F, Marquis F. A randomized controlled trial of home telerehabilitation for post-knee arthroplasty. *J Telemed Telecare* 2011;17(4):195-198. [doi: [10.1258/jtt.2010.100602](#)]
51. Piqueras M, Marco E, Coll M, Escalada F, Ballester A, Cinca C, et al. Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med* 2013 Apr;45(4):392-396 [[FREE Full text](#)] [doi: [10.2340/16501977-1119](#)] [Medline: [23474735](#)]
52. Eisermann U, Haase I, Kladny B. Computer-aided multimedia training in orthopedic rehabilitation. *Am J Phys Med Rehabil* 2004 Sep;83(9):670-680. [Medline: [15314531](#)]
53. Li L, Gan Y, Zhang L, Wang Y, Zhang F, Qi J. The effect of post-discharge telephone intervention on rehabilitation following total hip replacement surgery. *Int J Nurs Sci* 2014 Jun;1(2):207-211. [doi: [10.1016/j.ijnss.2014.05.005](#)]
54. Russell TG, Buttrum P, Wootton R, Jull GA. Low-bandwidth telerehabilitation for patients who have undergone total knee replacement: preliminary results. *J Telemed Telecare* 2003;9 Suppl 2:S44-S47. [doi: [10.1258/135763303322596246](#)] [Medline: [14728759](#)]
55. Eriksson L, Lindström B, Gard G, Lysholm J. Physiotherapy at a distance: a controlled study of rehabilitation at home after a shoulder joint operation. *J Telemed Telecare* 2009;15(5):215-220. [doi: [10.1258/jtt.2009.081003](#)] [Medline: [19590025](#)]
56. Heuser A, Kourtev H, Winter S, Fensterheim D, Burdea G, Hentz V, et al. Telerehabilitation using the Rutgers Master II glove following carpal tunnel release surgery: proof-of-concept. *IEEE Trans Neural Syst Rehabil Eng* 2007 Mar;15(1):43-49. [doi: [10.1109/TNSRE.2007.891393](#)] [Medline: [17436875](#)]

57. Macías-Hernández SI, Vásquez-Sotelo DS, Ferruzca-Navarro MV, Badillo SS, Gutiérrez-Martínez J, Núñez-Gaona MA, et al. Proposal and evaluation of a telerehabilitation platform designed for patients with partial rotator cuff tears: a preliminary study. *Ann Rehabil Med* 2016 Aug;40(4):710-717 [FREE Full text] [doi: [10.5535/arm.2016.40.4.710](https://doi.org/10.5535/arm.2016.40.4.710)] [Medline: [27606278](https://pubmed.ncbi.nlm.nih.gov/27606278/)]
58. Antón D, Nelson M, Russell T, Goñi A, Illarramendi A. Validation of a Kinect-based telerehabilitation system with total hip replacement patients. *J Telemed Telecare* 2016 Apr;22(3):192-197. [doi: [10.1177/1357633X15590019](https://doi.org/10.1177/1357633X15590019)] [Medline: [26130735](https://pubmed.ncbi.nlm.nih.gov/26130735/)]
59. Giansanti D, Morelli S, Maccioni G, Lanzetta M, Macellari V. Health technology assessment of a homecare device for telemonitoring and telerehabilitation for patients after hand transplantation. *Telemed J E Health* 2008;14(1):69-75. [doi: [10.1089/tmj.2007.0023](https://doi.org/10.1089/tmj.2007.0023)] [Medline: [18328027](https://pubmed.ncbi.nlm.nih.gov/18328027/)]
60. Placidi G. A smart virtual glove for the hand telerehabilitation. *Comput Biol Med* 2007;37(8):1100-1107. [doi: [10.1016/j.compbiomed.2006.09.011](https://doi.org/10.1016/j.compbiomed.2006.09.011)]
61. Burdea G, Popescu V, Hentz V, Colbert K. Virtual reality-based orthopedic telerehabilitation. *IEEE Trans Rehabil Eng* 2000 Sep;8(3):430-432. [Medline: [11001524](https://pubmed.ncbi.nlm.nih.gov/11001524/)]
62. Antón D, Nelson M, Russell T, Goñi A, Illarramendi A. Validation of a Kinect-based telerehabilitation system with total hip replacement patients. *J Telemed Telecare* 2016 Apr;22(3):192-197. [doi: [10.1177/1357633X15590019](https://doi.org/10.1177/1357633X15590019)] [Medline: [26130735](https://pubmed.ncbi.nlm.nih.gov/26130735/)]
63. Antón D, Nelson M, Russell T, Goñi A, Illarramendi A. Validation of a Kinect-based telerehabilitation system with total hip replacement patients. *J Telemed Telecare* 2016 Apr;22(3):192-197. [doi: [10.1177/1357633X15590019](https://doi.org/10.1177/1357633X15590019)] [Medline: [26130735](https://pubmed.ncbi.nlm.nih.gov/26130735/)]
64. Tousignant M, Giguère A, Morin M, Pelletier J, Sheehy A, Cabana F. In-home telerehabilitation for proximal humerus fractures: a pilot study. *Int J Telerehabil* 2014;6(2):31-37 [FREE Full text] [doi: [10.5195/ijt.2014.6158](https://doi.org/10.5195/ijt.2014.6158)] [Medline: [25945227](https://pubmed.ncbi.nlm.nih.gov/25945227/)]
65. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess* 2003;7(1):1-76 [FREE Full text] [Medline: [12583822](https://pubmed.ncbi.nlm.nih.gov/12583822/)]
66. Levy CE, Silverman E, Jia H, Geiss M, Omura D. Effects of physical therapy delivery via home video telerehabilitation on functional and health-related quality of life outcomes. *J Rehabil Res Dev* 2015;52(3):361-370 [FREE Full text] [doi: [10.1682/JRRD.2014.10.0239](https://doi.org/10.1682/JRRD.2014.10.0239)] [Medline: [26230650](https://pubmed.ncbi.nlm.nih.gov/26230650/)]
67. Spangehl M. Is it time for telerehabilitation to go mainstream? *J Bone Joint Surg Am* 2015 Jul 15;97(14):e55. [doi: [10.2106/JBJS.O.00540](https://doi.org/10.2106/JBJS.O.00540)] [Medline: [26178897](https://pubmed.ncbi.nlm.nih.gov/26178897/)]
68. Peterson C, Watzlaf V. Telerehabilitation store and forward applications: a review of applications and privacy considerations in physical and occupational therapy practice. *Int J Telerehabil* 2014;6(2):75-84 [FREE Full text] [doi: [10.5195/ijt.2014.6161](https://doi.org/10.5195/ijt.2014.6161)] [Medline: [25945231](https://pubmed.ncbi.nlm.nih.gov/25945231/)]
69. Sharareh B, Schwarzkopf R. Effectiveness of telemedical applications in postoperative follow-up after total joint arthroplasty. *J Arthroplasty* 2014 May;29(5):918-922.e1. [doi: [10.1016/j.arth.2013.09.019](https://doi.org/10.1016/j.arth.2013.09.019)] [Medline: [24342278](https://pubmed.ncbi.nlm.nih.gov/24342278/)]
70. Bendixen RM, Levy C, Lutz BJ, Horn KR, Chronister K, Mann WC. A telerehabilitation model for victims of polytrauma. *Rehabil Nurs* 2008;33(5):215-220. [Medline: [18767403](https://pubmed.ncbi.nlm.nih.gov/18767403/)]
71. Burdea G, Popescu V, Hentz V, Colbert K. Virtual reality-based orthopedic telerehabilitation. *IEEE Trans Rehabil Eng* 2000 Sep;8(3):430-432. [Medline: [11001524](https://pubmed.ncbi.nlm.nih.gov/11001524/)]
72. Vuorenmaa M, Ylinen J, Piitulainen K, Salo P, Kautiainen H, Pesola M, et al. Efficacy of a 12-month, monitored home exercise programme compared with normal care commencing 2 months after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med* 2014 Feb;46(2):166-172 [FREE Full text] [doi: [10.2340/16501977-1242](https://doi.org/10.2340/16501977-1242)] [Medline: [24241606](https://pubmed.ncbi.nlm.nih.gov/24241606/)]
73. Girone M, Burdea G, Bouzit M, Popescu V, Deutsch JE. Orthopedic rehabilitation using the "Rutgers ankle" interface. *Stud Health Technol Inform* 2000;70:89-95. [Medline: [10977590](https://pubmed.ncbi.nlm.nih.gov/10977590/)]

Abbreviations

CT: clinical trials

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 17.10.16; peer-reviewed by D Kairy, S Peterson; comments to author 01.12.16; revised version received 10.01.17; accepted 18.03.17; published 28.04.17.

Please cite as:

Pastora-Bernal JM, Martín-Valero R, Barón-López FJ, Estebanez-Pérez MJ

Evidence of Benefit of Telerehabilitation After Orthopedic Surgery: A Systematic Review

J Med Internet Res 2017;19(4):e142

URL: <http://www.jmir.org/2017/4/e142/>

doi: [10.2196/jmir.6836](https://doi.org/10.2196/jmir.6836)

PMID: [28455277](https://pubmed.ncbi.nlm.nih.gov/28455277/)

©Jose Manuel Pastora-Bernal, Rocio Martín-Valero, Francisco Javier Barón-López, María José Estebanez-Pérez. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Effect of a Nine-Month Web- and App-Based Workplace Intervention to Promote Healthy Lifestyle and Weight Loss for Employees in the Social Welfare and Health Care Sector: A Randomized Controlled Trial

Nina Charlotte Balk-Møller¹, MA (Sociology); Sanne Kellebjerg Poulsen¹, MSc, PhD; Thomas Meinert Larsen¹, MSc, PhD

Department of Nutrition, Exercise and Sports, Faculty of Science, Copenhagen University, Frederiksberg C, Denmark

Corresponding Author:

Nina Charlotte Balk-Møller, MA (Sociology)

Department of Nutrition, Exercise and Sports

Faculty of Science

Copenhagen University

Rolighedsvej 26

Frederiksberg C, 1958

Denmark

Phone: 45 27126320

Fax: 45 35320829

Email: nbm@nexs.ku.dk

Abstract

Background: General health promoting campaigns are often not targeted at the people who need them the most. Web- and app-based tools are a new way to reach, motivate, and help people with poor health status.

Objective: The aim of our study was to test a Web- and mobile app-based tool (“SoSu-life”) on employees in the social welfare and health care sector in Denmark.

Methods: A randomized controlled trial was carried out as a workplace intervention. The tool was designed to help users make healthy lifestyle changes such as losing weight, exercise more, and quit smoking. A team competition between the participating workplaces took place during the first 16 weeks of the intervention. Twenty nursing homes for elderly people in 6 municipalities in Denmark participated in the study. The employees at the nursing homes were randomized either 1:1 or 2:1 on a municipality level to use the SoSu-life tool or to serve as a control group with no intervention. All participants underwent baseline measurements including body weight, waist circumference, body fat percentage, blood pressure, and blood cholesterol level and they filled in a questionnaire covering various aspects of health. The participants were measured again after 16 and 38 weeks.

Results: A total of 566 (SoSu-life: n=355, control: n=211) participants were included in the study. At 16 weeks there were 369 participants still in the study (SoSu-life: n=227, control: n=142) and 269 participants completed the 38 week intervention (SoSu-life: n=152, control: n=117). At 38 weeks, the SoSu-life group had a larger decrease in body weight (−1.01 kg, $P=.03$), body fat percentage (−0.8%, $P=.03$), and waist circumference (−1.8 cm, $P=.007$) compared with the control group.

Conclusions: The SoSu-life Web- and app-based tool had a modest yet beneficial effect on body weight and body fat percentage in the health care sector staff.

Trial Registration: Clinicaltrials.gov NCT02438059; <http://clinicaltrials.gov/ct2/show/NCT02438059> (Archived by WebCite at <http://www.webcitation.org/6i6y4p2AS>)

(*J Med Internet Res* 2017;19(4):e108) doi:[10.2196/jmir.6196](https://doi.org/10.2196/jmir.6196)

KEYWORDS

health promotion; workplace; smartphone; weight reduction programs; Internet; eHealth; randomized controlled trial

Introduction

Background

Public health promoting campaigns have trouble reaching the socioeconomically disadvantaged groups [1] who really need them, and most face-to-face or telephone based interventions are generally considered too costly. Disadvantaged groups often need more targeted effort in order to be able to succeed in lifestyle changes. Therefore, Web- and app-based tools have been suggested as a new and potentially cost-effective way to reach out to, motivate, and help people to improve their health [2].

Web- and app-based tools for health promotion have been shown to have potential for weight loss programs and improving other health related factors. The number of these digital tools for health promotion being developed is increasing. The aim is generally to help users improve lifestyle and develop healthier habits, and thereby improve their health status. The need for effective tools for promoting health is evident from the recent report from the World Health Organization (WHO), which shows that overall health status in Europe is improving, but that the variation in health status is increasing both within and between countries [3].

In the Danish population, overall health, and in particular the distribution of overweight and obesity, is unequally distributed, with higher levels of overweight and obesity occurring among individuals with lower educational level [4]. Employees in the social and health sector generally belong to this group. A large number of these employees are organized in the trade union of public employees (FOA), where about 90% of the members are women. A recent health examination among 1737 female members of FOA, 59% of whom were women employed in the Danish social welfare and health care sector [5], found an increased prevalence of heavy smokers, overweight and obesity, and long-term sick leave compared with other employees at the same income level. However, the majority of the employees considered it very important that they maintain or improve their health [5].

Weight management is important for alleviating related health problems such as cardiovascular disease and type 2 diabetes in overweight individuals [3]. Overweight and obesity are not only associated with somatic diseases, but they are also related to several psychosocial complications, such as depression and anxiety [6,7], and to increased sick leave [8]. Furthermore, overweight, heavy smoking, and increased prevalence of long-term sick leave are significant precursors of chronic illnesses and early retirement [5,8,9]. The increasing prevalence of overweight and obesity, the associated health risks, and the increasing expenses related to these conditions make new and better targeted, well documented, and cost-effective health promotion programs for lifestyle change even more necessary.

The workplace is increasingly used as an arena for health promotion because it is the place where many people spend a great part of their time, and WHO has recommended that the workplace be prioritized as such an arena [10].

Web-based health promoting programs have occasionally been shown to assist weight loss in overweight and obese individuals. However, results are generally inconsistent and depend greatly on the type of Web-based tool used, as well as on the study design [11]. It is therefore intrinsically difficult to compare the results of different intervention studies using Web-based weight loss tools. However, a recent meta-analysis including 23 studies comprising 8697 participants found that Web-based tools for weight loss had a modest but significant effect on weight loss: -0.68 kg compared with a nontreated control group [12]. A Cochrane analysis found an effect of -1.5 kg after 6 months [13].

The effect of workplace health promotion interventions has generally been found to be rather small and greatly dependent on the quality of the study [14]. Only a few Web- and app-based intervention studies have been conducted in workplace settings [15-17], and only two of these were randomized controlled trials [16,17]. A study by Van Vier et al [16] found that a Web-based intervention could result in a weight loss of -1.1 kg and a reduction in waist circumference of -1.2 cm, and a study by Cook et al [17] found that a Web-based intervention was more effective than printed material in inducing improvements in diet and nutrition, though this study did not examine changes in body weight.

Mobile phone apps may be useful as health promoting tools. A systematic review with meta-analysis including 17 studies with the use of mobile devices for weight loss found that the use of mobile devices induced weight loss [18]. A recent study investigating the potential of a Web-based app for promoting a healthy lifestyle found promising results [19]. However, the field has not been sufficiently explored and the potential of mobile phone apps for health promotion is not fully understood [20]. There is evidence that adding a social feature to the intervention, such as a team-based element where users can compare and compete with each other, can have a positive effect on users' willingness to use a mobile app for health promotion [21]. We have found no previous studies that have combined the use of a workplace Web-based health promoting program with a social feature.

Aim of This Study

The primary aim of this study was therefore to investigate the effect of the SoSu-life Web- and mobile phone-based app, in combination with a social feature, on changes in body weight. Secondary outcomes were changes in body fat percentage, waist circumference, blood pressure and total cholesterol (all reported in this paper), behavior change, and self-perceived well-being (both reported in a paper under preparation). User-data and qualitative data from the study are reported in another paper (Balk-Møller, under review in JMIR 2017). The control group underwent the same examinations as the intervention group, which provided information about the stand-alone effect of such health examinations.

Methods

Sampling and Study Design

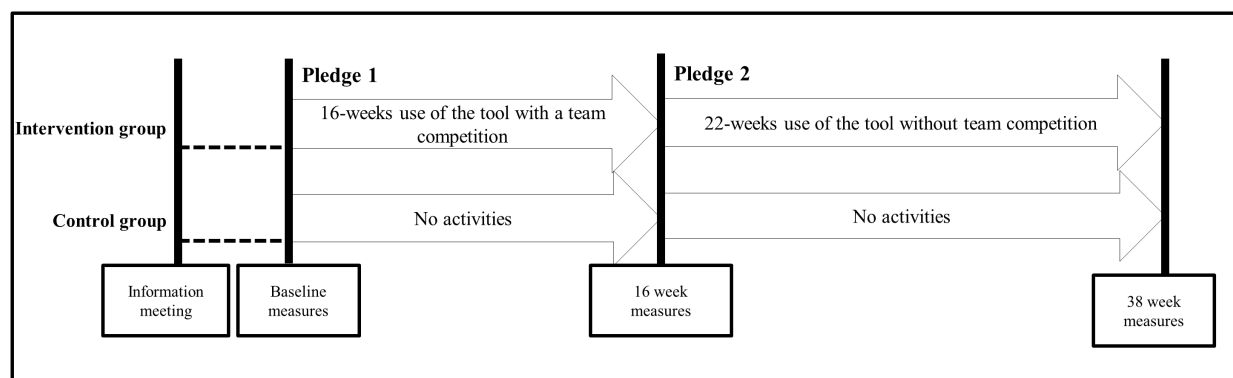
The study was conducted in 6 municipalities in Denmark from August 2012 to July 2013. In each municipality, 2-4 nursing homes were randomized to either the SoSu-life tool (called SoSu-life group) or to a control group. The randomization of the participating nursing homes was conducted continually over time within each municipality. Those municipalities with an even number of nursing homes were randomized in a 1:1 ratio and municipalities with an odd number of nursing homes in a 2:1 ratio. Of the total 20 nursing homes, 12 were randomized to SoSu-life and 8 were randomized to control. The randomization was performed in a simple blinded way (simple paper draw) by 2 of the study investigators in a collaboration between investigators and staff at an initial meeting with local staff in each municipality. Each draw was observed by independent witnesses to observe that it was performed in a fair and unbiased way. The study was divided into two distinct periods for the SoSu-life group: an initial 16-week period including a team competition, and a subsequent 22-week period without a competition. Participants chose a pledge (described later) for each period ([Figure 1](#)).

Participants were recruited from August 2012 to September 2012 when an information meeting was held at each nursing home during work hours, after which each of the interested employees completed the informed consent form. Eligibility criteria for participation were that the employees had to work under conditions in accordance with a FOA-negotiated agreement. Computer or mobile app literacy was not an inclusion criterion. A few weeks later, the participants underwent a baseline clinical examination conducted by trained staff from the research group. The clinical assessments were all performed locally at each participating nursing home, which made it impossible for the research staff to be blinded to the participants' group allocation. The anthropometrical assessments were measured using a wall-mounted stadiometer for height and a digital electronic scale (Tanita WB 100MA/WB-110MA III) for body weight. Measuring tape was used for measuring hip and waist circumference taken at the widest part of the hip and at the umbilicus, respectively. Body fat percentage was recorded with a handheld body composition monitor (Omron BF306) and blood pressure was measured using a digital blood pressure manometer (Kivex, Automatic Blood Pressure Monitor,

Model UA-787 Plus). Finger-prick blood samples were drawn to measure total cholesterol (Accutrend Plus; more information in [Multimedia Appendix 1](#)). Participants in the intervention group signed up for their first pledge (Pledge 1) immediately after the clinical examination, and were then introduced to the SoSu-life website and app by a member from the project team, and given a pamphlet about the content and functionalities of the tool. This session took approximately 30-45 min. The clinical examination, taking about 15 min in total, was repeated at 16 weeks (December 2012 to January 2013). The participants in the intervention group then made a new pledge (Pledge 2). The examination was repeated at 38 weeks (May 2013 to June 2013), when participants also answered a set of questions addressing their overall evaluation of the study. The participants in both groups spent approximately 10-15 min answering a Web-based questionnaire based on a questionnaire developed by the Danish National Research Centre for the Working Environment, regarding demographic data, health behavior, and general well-being the day before each of the 3 clinical examinations. The study was approved by the ethical committee of the Capital Region of Denmark (reference number: H-2-2012-079; a more detailed description of the study design is given in the [Multimedia Appendix 1](#)). The study is registered at Clinicaltrials.gov with ID number NCT02438059.

Developing the Tool

The SoSu-life tool is a Web- and app-based tool developed specifically for this particular group of employees. It was created in a collaboration between the company Mobile Fitness, employees from the sponsoring company PenSam, who work with health promotion for the social and health care workers on a daily basis, and scientists from the University of Copenhagen. The tool was developed through an iterative process where the users (ie, the social and health care workers) and the developers (Mobile Fitness, PenSam, and Copenhagen University) were involved. Initially, focus groups from the target group were held, where ideas and mock-ups were presented. The first edition of the SoSu-life tool was tested in a pilot study at two nonparticipating nursing homes. The ideas for mechanisms that could lead to behavior change were drawn from experience in earlier health promotion projects (nondigital) among social and health care workers, and from Mobile Fitness' expertise with a digital weight loss tool used in a similar population, both with individual users and applied in workplace settings (unpublished data).

Figure 1. Overview of the SoSu-life study.

Description of the SoSu-Life Tool

The tools' basic features were self-reporting of diet and exercise, personalized feedback, suggestions for activities and programs, and practical tips and tricks. A series of social features, including weekly assignments and colleague challenges designed to create social support and build social interactions at the workplace, played a vital role in promoting behavior change. The SoSu-life tool aimed to mobilize whole groups of colleagues to not only encourage each other in achieving personal goals, but also work on identical small weekly assignments. Points were assigned to all individual and group activities and were collected for both individuals and groups. Individual activities gave points not only to the individual but also to the group as part of the group competition. In this way, each individual's use of the digital tool benefited the whole group.

On entering the study, each participant chose one pledge out of seven to focus upon: lose weight, eat healthier, improve physical fitness, improve physical strength, quit smoking, decrease the number of cigarettes smoked, or maintain a healthy lifestyle. The program itself recommended a pledge based on information obtained from the health examination. The choice of pledge was designed to give the participants a key focus area. Subsequent communication—feedback, frequency, and content of emails and text messages provided by the program, and topics for user-to-user communication—was targeted according to the participants' individual pledge. Messages contained information about specific health issues related to the pledge, general tips, and tricks on health and well-being.

The program had various tools to help the user succeed with the pledges. The self-reporting of diet and exercise was a weight

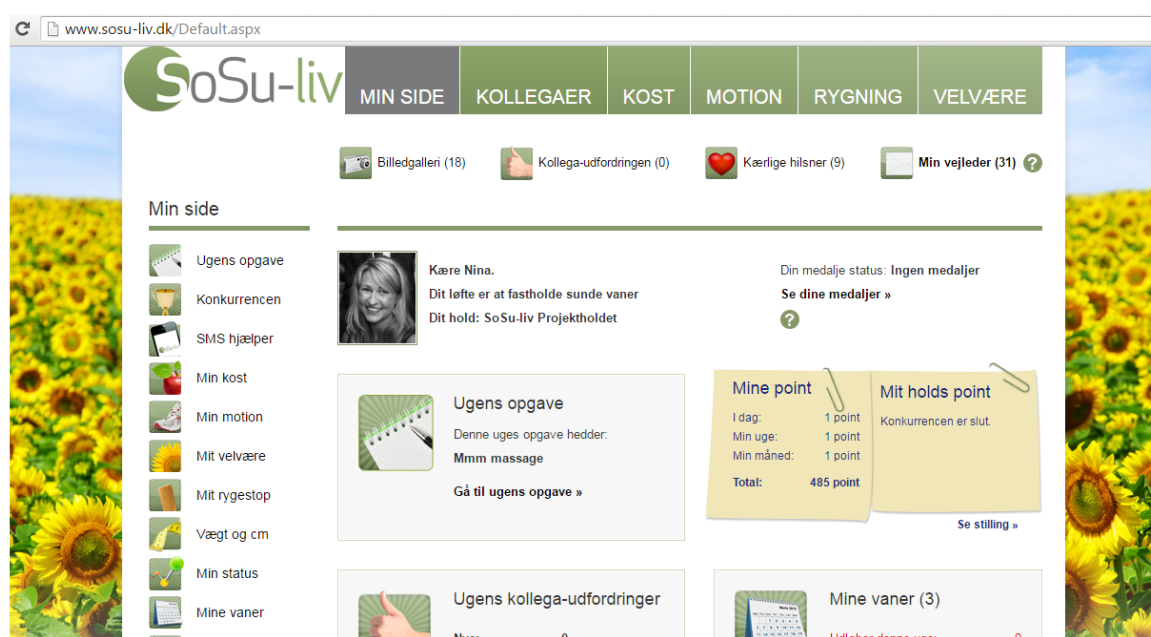
loss tool based on a unit system designed to assist either weight loss or maintenance of a weight loss [22]. All foods were assigned a number of units based on portion size, calories, and macronutrient composition [22]. Daily energy level was calculated based on the users' height and weight, and on this basis the number of units per day advisable for weight loss was suggested. The user registered his or her food intake and daily exercise, and the program gave feedback on the energy balance of the day, a green code indicating a proper energy balance, and a red code for excessive energy intake. Exercise was registered as bonus-units according to standardized energy expenditure equivalents related to specific types of exercise [23] (more information is given in the [Multimedia Appendix 1](#)). The same system was used for those participants wishing to focus on exercise alone, and feedback was given in the form of number of earned bonus-units. The website provided access to a number of video-supported exercise programs designed to increase fitness level or improve strength. Smokers wishing to either change their smoking habits or quit smoking were advised to begin by registering their habitual use of cigarettes, the time the cigarettes were smoked, and the mood they were in when they were smoking the cigarettes.

The social features included weekly assignments for all group members. Such weekly assignments could be "drink at least one liter of water every day all week" or "remember to say Good Morning to your colleagues every morning all week." The tool also included "colleague challenges" that were to be sent from colleague to colleague and were determined by the participants' individual pledge. Challenges might be: "Do not eat sugar for three days" or "Bring some fruit for us two to eat together tomorrow during the afternoon break." All features could be accessed from both the app and website ([Figures 2 and 3](#)).

Figure 2. Screenshot of the SoSu-life app menu.

The SoSu-life program entails a team competition and used a point system where all activities performed using the tool gave points to the individual user. The point system provided the highest reward for taking part in social activities. Performing the weekly challenges, and sending and carrying out “colleague challenges,” were rewarded with more points than registering diet or exercise, or taking tests or quizzes. During the first 16 weeks, each of the participating nursing homes or home care units constituted a team, and each of the individual participants’ points were added to the teams’ total points. A lottery ticket was generated for each point earned by the team. Each month the teams had a chance to win a prize from a simple lottery drawn by the research staff. The more points the team had, the bigger the chances of winning. The prizes included things such as a shopping bag for each team member, a Zumba class for the

team, or a visit from a bartender who served fruit smoothies during lunch hours. The team that collected the most points after 16 weeks also won a prize. Points were still collected in the second (22-week) intervention period, but no prizes were provided. The cost of the prizes was covered by the main sponsor of the project, who also presented the prizes at local celebration events at the nursing homes. The social features and the team competition were designed to create a supporting and encouraging work environment to help generate behavior change in the individual participant. A more complete description of the tool based on the TiDieR-checklist is given in the [Multimedia Appendix 1](#). The content in the intervention remained the same during the 38 weeks, and there were no significant bug fixes or downtimes during the project.

Figure 3. Screenshot of the SoSu-life website.

Statistical Analyses

All statistical analyses were performed in Stata release 11 (StataCorp LP). Differences in anthropometric measures from baseline to week 16, and from baseline to week 38, were analyzed using a multilevel mixed effect linear regression (mixed model) with municipality as a random effect. Significance was set at P value $\leq .05$ for all analyses.

Descriptive statistics of the sample were presented as mean and standard deviation (SD). Categorical data were presented as percentage of the sample (n). Differences between groups were tested with unpaired t test, and chi-squared for categorical data.

Use of the SoSu-life tool differed according to the pledge taken. Participants pledging change of smoking habits were therefore left out of the anthropometric outcomes. A separate analysis of a subgroup of participants with weight loss pledge, both during the first 16 weeks and the last 22 weeks (subgroup with weight loss pledge 1 and 2), was carried out.

The study was designed and powered as a “completers only” study, considering the completer analysis as the primary outcome, and that an intention-to-treat (ITT) analysis should serve as a sensitivity analysis. This was decided on the basis of two major considerations. First, the study was designed as a real-life intervention applied in a workplace setting, and participation was completely voluntary. The information provided to the study participants in the recruitment phase was relatively brief, approximating what we consider realistic if the intervention was to be disseminated on a much larger scale. Second, the pragmatic study design with only 4 visits at each workplace (information meeting, baseline meeting, week 16, and week 38) essentially gave no room for any rescheduling of clinical measures and completion of Web-based questionnaires at the workplaces. This meant that absence from work on the

clinical visit days, whether due to time off or illness, automatically led to missing data.

We believe this would have resulted in quite a significant amount of missing data. However, when performing an ITT baseline carried forward analysis, the results showed no major changes (see [Multimedia Appendix 1](#)).

Power Calculations

Change in body weight among study completers was used for the power calculations. Recruiting a total of 450 participants divided equally between the 2 study groups, and assuming (1) dropout rate of 25% and 20% over the 16 weeks and 22 weeks intervention periods, respectively, (2) a between-group difference of 2.0 kg, (3) SD of 4.0 kg in both groups, and (4) a significance level of .05, this would result in a power of 0.98. Fewer people volunteered to participate in and fewer completed the study than anticipated (152 in the intervention group vs 117 in the control group). A post hoc power analysis based on the actual results found a power of 0.27.

Results

Description of the Study Sample

Initially, 1203 employees from 20 different nursing homes in 6 municipalities were invited to participate in the study. After the information meeting, a total of 566 participants signed the consent form and underwent baseline examination ($n=355$ in the intervention group, $n=211$ in the control group). The second examination took place after 16 weeks, when 65.2% (369/566) participants were measured ($n=227$ in the intervention group, $n=147$ in the control group). The same procedure was applied after the total of 38 weeks, when 47.5% (269/566) still remained in the study ($n=152$ in the intervention group, $n=117$ in the control group; [Figure 4](#)).

There were no significant differences at baseline between the (Table 1). groups regarding social and anthropometric characteristics

Table 1. Social and anthropometric characteristic at baseline for completers in SoSu-life group and control group.

Variable	SoSu-life ^a	Control ^b	<i>P</i> value
Age (in years), mean (SD) ^c	47.0 (10.0)	47.0 (9.9)	.99 ^d
Females, n (%)	140 (92.1)	108 (92.3)	.95 ^e
Work schedule^f , n (%)			
Day duty	118 (77.6)	90 (79)	.88 ^e
Evening duty	18 (11)	15 (13.2)	
Night duty	8 (5)	4 (4)	
Varying shifts	8 (5)	5 (4)	
Marital status^g , n (%)			
Single	47 (31)	31 (27)	.44 ^e
Cohabitate	103 (68.7)	84 (73)	
Children in the household^h , n (%)			
Yes, 1 child	36 (24)	39 (34)	.08 ^e
Yes, 2 children	31 (20)	14 (12)	
Yes, 3 or more	15 (10)	16 (14)	
No	70 (46)	45 (40)	
Height (cm), mean (SD)	166.2 (7)	165.5 (7.2)	.41 ^d
Body weight (kg), mean (SD)	74.5 (15.9)	73.1 (14.8)	.44 ^d
Waist circumference (cm) ⁱ , mean (SD)	92.8 (13.9)	91.2 (11.9)	.27 ^d
Hip circumference (cm), mean (SD)	103.6 (10.8)	101.6 (10.8)	.12 ^d
Total cholesterol (mmol/L) ^j , mean (SD)	5.3 (0.9)	5.3 (0.9)	.78 ^d
Body fat percentage (%) ^k , mean (SD)	35.3 (7.4)	35.1 (7.4)	.88 ^d
Systolic blood pressure (mmHg), mean (SD)	129.0 (16.5)	130.8 (16.4)	.37 ^d
Diastolic blood pressure (mmHg), mean (SD)	80.9 (9.6)	82.1 (10.0)	.32 ^d

^aSoSu-life: n=152, unless other is specified.

^bControl: n=117, unless other is specified.

^cSD: standard deviation.

^dData are compared between groups using students *t* test.

^eData are compared between groups using chi-square test.

^fControl: n=114.

^gSoSu-life: n=150, control: n=115.

^hControl: n=114.

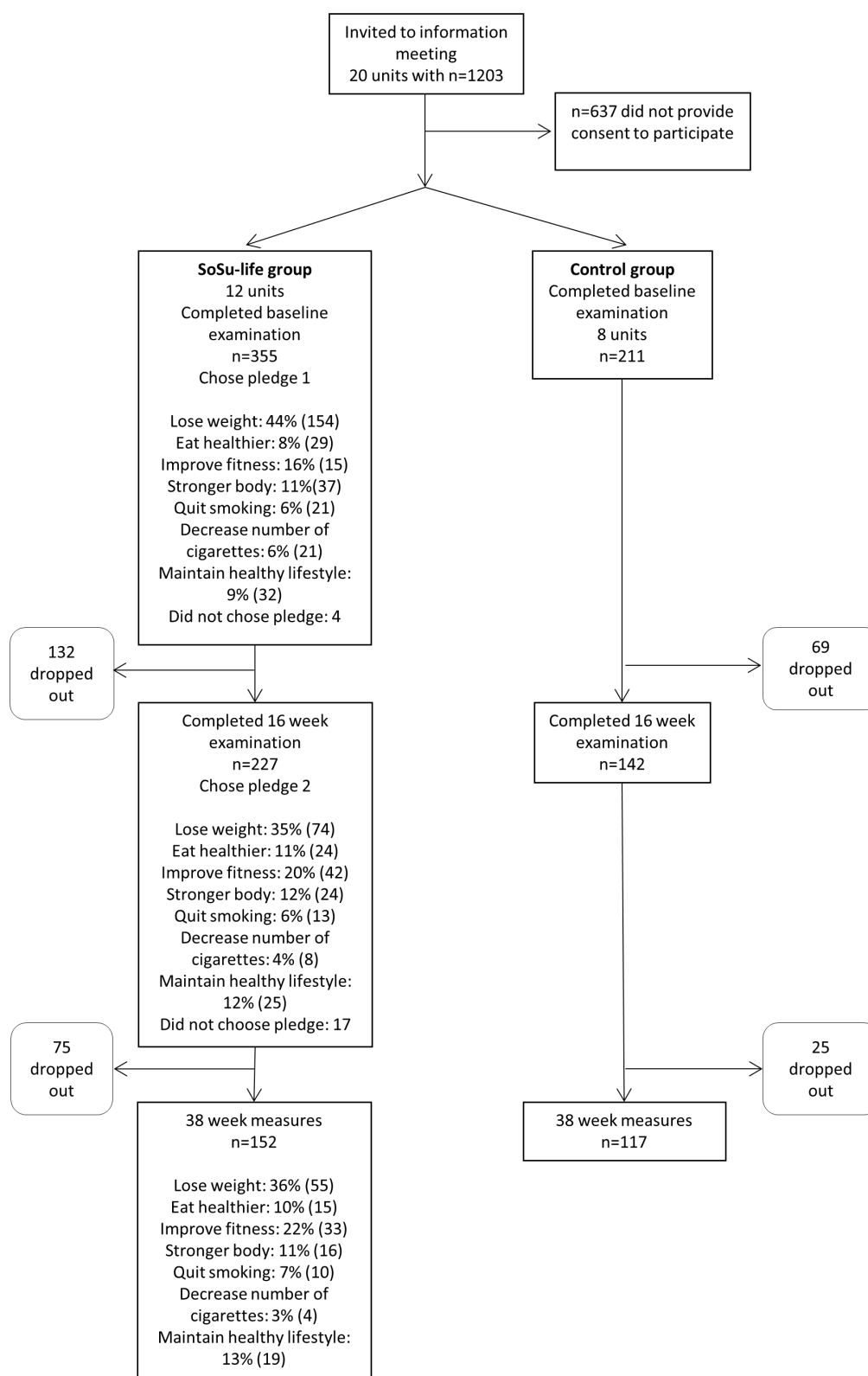
ⁱControl: n=116.

^jSoSu-life: n=151, control: n=112.

^kBody fat percentage device had a maximum of 49.9 %, SoSu-life: n=148, control: n=115.

Our study population is representative of the general characteristics of social and health care workers in Denmark [24]. The study included slightly more participants working day shifts (77%) than is normal in this sector (60%). The study population also included fewer shift workers (around 5%

compared with 15% in general) [24]. Dropout rate was higher for shift workers and for evening workers. Dropouts were otherwise generally comparable with those that completed the study, except that they had a higher waist circumference (data shown in Multimedia Appendix 1).

Figure 4. Flow diagram of the participants.

Changes in Body Weight and Anthropometric Markers

During the intervention period the SoSu-life group (excluding participants with pledges regarding smoking) had a larger decrease in body weight of -1.01 kg ($P=.03$), body fat

percentage -0.78% ($P=.03$), and waist circumference of -1.79 cm ($P=.007$) after a total of 38 weeks compared with the control group. The weight loss subgroup had an even larger decrease in body weight of -1.64 kg ($P=.02$) and waist circumference

of -2.47 cm ($P=.008$) after 38 weeks compared with the control group.

The SoSu-life group (excluding participants with pledges regarding smoking) had a larger decrease in body weight of -1.54 kg ($P<.001$) and a decrease in body fat percentage of -0.81% ($P=.003$) compared with the control group during the first 16 weeks. Here the weight loss subgroup had a decrease in body weight of -2.36 kg ($P<.001$), a decrease in body fat percentage of -0.99% ($P=.003$), and a decrease in waist circumference on -2.45 cm ($P=.003$) compared with the control group (Table 2).

No significant difference was observed between any of the SoSu-life groups or the control group in any of the other outcomes, such as blood pressure or total cholesterol.

When performing an ITT-analysis (results presented in Multimedia Appendix 1), as expected, the overall effect size between SoSu-Life and control group decreased, but remained statistically significantly different for the primary outcomes, such as body weight change. Hence, the ITT-analysis only differed from the “completers only” results in terms of the changes in body fat percentage from week 0 to week 38 between SoSu-life group and control that went from a significant ($P=.03$) to nonsignificant ($P=.20$), changes in waist circumference from week 0 to week 16 between SoSu-life group and control that went from nonsignificant ($P=.09$) to significant ($P=.03$), and from week 0 to week 38 between subgroup and control that went from significant ($P=.008$) to nonsignificant ($P=.05$).

Table 2. Changes in body weight, body fat percentage, waist circumference, systolic blood pressure, diastolic blood pressure, and total cholesterol from week 0 to week 16 and from week 0 to week 38 for completers in Web- and app-based intervention for health promotion.

Changes from week 0 to week 16 and from week 0 to week 38	n	SoSu-life Mean (SE ^a)	n	Subgroup with weight loss pledge ^b Mean (SE)	n	Control Mean (SE)	Adjusted difference ^c between SoSu-life group and control (95% CI)	<i>P</i> value ^c between SoSu-life group and control	Adjusted difference between subgroup ^b and control (95% CI)	<i>P</i> value ^c between sub- group ^b and control	
Body weight ^d (kg)	week 0 to 16	130	-1.44 (0.26)	41	-2.39 (0.52)	117	0.10 (0.19)	-1.54 (-2.18 to -0.90)	<.001	-2.36 (-3.23 to -1.49)	<.001
	week 0-38	130	-1.04 (0.34)	41	-1.68 (0.77)	117	-0.03 (0.33)	-1.01 (-1.94 to -0.08)	.03	-1.64 (-3.04 to -0.24)	.02
Body fat Percentage ^d	week 0-16	127	-0.70 (0.21)	38	-0.97 (0.29)	114	0.04 (0.16)	-0.81 (-1.35 to -0.27)	.003	-0.99 (-1.63 to -0.34)	.003
	week 0-38	126	-1.50 (0.23)	37	-1.13 (0.40)	114	-0.71 (0.27)	-0.78 (-1.48 to -0.09)	.03	-0.39 (-1.43 to 0.64)	.46
Waist Circumference ^d (cm)	week 0-16	130	-1.40 (0.47)	41	-3.12 (0.93)	116	-0.51 (0.37)	-1.05 (-2.26 to 0.16)	.09	-2.45 (-4.09 to -0.81)	.003
	week 0-38	129	-0.67 (0.50)	40	-1.90 (0.96)	116	-0.04 (0.44)	-1.79 (-3.09 to -0.49)	.007	-2.47 (-4.30 to -0.63)	.008
Systolic blood pressure (mmHg)	week 0-16	152	-2.27 (0.93)	41	-1.90 (1.43)	117	-2.57 (1.03)	0.30 (-2.41 to 3.02)	.83	-0.52 (-3.42 to 2.37)	.66
	week 0-38	152	-2.82 (1.06)	41	-2.93 (1.56)	116	-5.83 (1.22)	2.27 (-0.96 to 5.50)	.17	-0.63 (-3.42 to 2.16)	.72
Diastolic blood pressure (mmHg)	week 0-16	152	-1.85 (0.64)	41	-1.71 (2.02)	117	-1.48 (0.72)	-0.57 (-2.48 to 1.35)	.56	0.87 (-3.23 to 4.96)	.68
	week 0-38	152	-2.02 (0.62)	41	-4.02 (2.24)	117	-2.30 (0.65)	0.10 (-1.69 to 1.90)	.91	1.37 (-3.43 to 6.16)	.58
Total cholesterol (mmol/L)	week 0-16	149	0.10 (0.07)	40	0.20 (0.12)	112	0.20 (0.09)	-0.09 (-0.30 to 0.13)	.43	-0.01 (-0.33 to 0.31)	.97
	week 0-38	148	-0.12 (0.07)	40	-0.13 (0.17)	111	0.01 (0.08)	-0.10 (-0.31 to 0.12)	.39	-0.09 (-0.42 to 0.24)	.61

^aSE: standard error.

^bSubgroup defined by having weight loss pledge 1 and 2.

^cMultilevel mixed effect linier regression with municipality as a random effect.

^dParticipants with pledge regarding smoking habits was left out.

Discussion

Principal Findings

The primary aim of the study was to evaluate whether using a Web- and app-based tool for 38 weeks could stimulate beneficial changes on body weight and other anthropometric markers. We found that the SoSu-life group had lost -1.01 kg more body weight and -0.8% more body fat percentage than the control group at 38 weeks. The weight loss subgroup had an even larger decrease in body weight of -1.64 kg. These results correspond well with results from other studies [2,11,13,25]. These results suggest that a Web- and app-based weight loss tools can provide modest but statistically significant effects (-0.68 kg) compared with other weight loss methods (face-to-face-contact not included) [12]. The lower than expected completion rate may, however, have led to a lower statistical power than originally estimated.

Interpretation

Overall, these results seem to correspond well with results from other workplace interventions studies [14]. The results indicate that the first 16 weeks of using the tool were the most effective. This is very much in line with other studies that have shown more favorable results in shorter periods with a Web-based weight loss tool [26,27]. In this study, some of the effects may be due to the competitions, with prizes, and the emphasis on developing healthy habits together with colleagues in the first intervention period. The lack of social support associated with Web-based tools has previously been reported as a problem [28]. The SoSu-life tool distinguishes itself from most other Web-based tools by having incorporating social support from known colleagues, both at the workplace and in a Web-based forum, and our findings indicate the importance of social features when addressing this target group.

We found that some of the weight loss and decrease in body fat percentage was maintained throughout the 38 weeks. However, it is debatable whether a period of 38 weeks is long enough to determine the long-term effects of a Web- and app-based weight control tool. A meta-analysis by Kodama et al [12] suggests that the study period must exceed 12 months if it is to determine long-term effects. Only a few studies in this meta-analysis exceeded 12 months and these studies did not find any positive effects in using the Web-based tools [12]. The SoSu-life study shows that it is possible to maintain a weight loss for a relatively long period with the use of a Web- and app-based tool, but it cannot predict longer term effects.

The degree of motivation of participants is likely to influence results. Most previous comparable studies have recruited voluntary participants who are overweight or obese, and who have the specific intention to lose weight [25,26,29]. In this regard, our study differed significantly from previous studies. We included as many participants as possible at each workplace, regardless of their initial health status and motivation. We have therefore recruited participants with varying degrees of need for behavioral change, and different levels of motivation to change behavior. We therefore speculate that some participants took part in the study simply because their colleagues signed up and out of curiosity alone, and not necessarily with a wish

to change lifestyle, and this may explain the small effects we observed.

There were fewer shift workers in the study population compared with the total population, and more shift and evening workers dropped out of the study than those working steady daytime or weekend schedules. This could be explained by the design of the intervention study, where the health examinations primarily took place during the day time, but the social elements in the tool and the group support from colleagues probably also worked better for those with a natural interaction with several colleagues during working hours.

We did not find any effects on blood pressure and cholesterol. We ascribe this to the rather high variation in these measures, as well as the modest effects that were shown in the anthropometric markers. The intervention needs to be further improved or applied for longer time periods if it is to induce more significant clinical benefits. Furthermore, other clinical outcomes, such as insulin sensitivity, should be investigated, and subgroup assessments might identify risk profiles that would benefit more than the average user.

The extent to which the study participants were capable of using the Web- and app-based digital tools was not considered in the study, so essentially all employees were considered eligible. The participants were only given a brief introduction as to how to use the tool. This particular group of people, with lower socioeconomic status and working day-to-day work with elderly residents, may well be less capable of using digital tools than the average population. Several participants had never used a mobile phone before and were not very comfortable with using a computer. This invariably created a nonuser group of participants in the intervention group. The results may thus not be fully ascribable to the actual use of the SoSu-life tool but also in part to increased focus on health in general. Furthermore, randomization was not blinded, and as both intervention and control group underwent three health examinations, participants in the control group may have received an incentive to make lifestyle changes by virtue of the information from their health examination.

The intervention was designed to be applicable to the social and health care workers' general workplace. However, if implemented in a practical context, the health examinations might very possibly be curtailed, or dropped, due to the extra cost, and the introduction to the tool would probably take place on the Web, so the effects of the tool would probably be smaller. The SoSu-life tool was designed to run for 38 weeks. For the intervention to be sustainable in the long term, it would probably require additional resources from the developers, sponsors, and the workplaces. New content would be needed to be added to the tool, repeated team competitions with accompanying prizes (from the sponsor) would probably have to take place, and management would have to make an effort to motivate employees to keep participating and using the tool. Regular health checks for the employees could possibly be conducted by managers or other staffs, as most of them are trained health professionals.

Conclusions

In conclusion, the SoSu-life Web- and app-based intervention in employees in the social welfare and health care sector resulted in a lowering in body weight, body fat percentage, and waist circumference, particularly among those that were specifically

motivated for weight loss, and during the first 16 weeks of intervention, when encouraged with group-based prizes in a competition. The longer term effects at week 38 were modest and associated with a relatively high dropout. To maintain the effects beyond an initial period will probably require further reinforcement of the intervention.

Acknowledgments

We acknowledge the participation and support from the participating municipalities, nursing homes, and their employees in the social welfare and health care sector. We acknowledge Professor Arne Astrup for creating the initial contact between the participating stakeholders and for his key role in study conception. We acknowledge Mobile Fitness A/S who undertook the development of the Web- and mobile app-based intervention tool. Mobile Fitness A/S has a stated purpose of developing a scientifically proven health program. Mobile Fitness is founded and owned by the University of Copenhagen, CAT Innovation A/S, and private initiators. Mobile Fitness A/S had no further role in the study design, the data collection, the analysis, the interpretation of data, and the decision to submit the article for publication. We acknowledge Tina Cuthbertson for helping with proof reading the paper. The project was funded by PenSam Livsforsikringsselskab and by a grant from the Danish Agency for Science Technology and Innovation, Ministry of Science Technology and Innovation.

Conflicts of Interest

PenSam Livsforsikringsselskab manages occupational pension schemes for approximately 340,000 wage-earners employed in Danish municipalities and regions, and in private organizations. Arne Astrup is cofounder and coowner of the University of Copenhagen spin-out company Mobile Fitness A/S, Denmark. The University of Copenhagen takes full responsibility for the study design, the data collection, the analysis, the interpretation of data, and the decision to submit the article for publication.

Multimedia Appendix 1

Supplementary material.

[[PDF File \(Adobe PDF File\), 278KB - jmir_v19i4e108_app1.pdf](#)]

Multimedia Appendix 2

Presentation of the SoSu-life project and tool given to the participants in the intervention group at the information meeting. NB In Danish.

[[PDF File \(Adobe PDF File\), 5MB - jmir_v19i4e108_app2.pdf](#)]

Multimedia Appendix 3

The CONSORT eHealth checklist (v. 1.6.1) for the study.

[[PDF File \(Adobe PDF File\), 1MB - jmir_v19i4e108_app3.pdf](#)]

References

1. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Med Res Methodol* 2014;14(1):42 [[FREE Full text](#)] [doi: [10.1186/1471-2288-14-42](#)]
2. Bennett G, Glasgow R. The delivery of public health interventions via the Internet: actualizing their potential. *Annu Rev Public Health* 2009 Apr;30(1):273-292. [doi: [10.1146/annurev.publhealth.031308.100235](#)]
3. World Health Organization. Obesity: Preventing and Managing the Global Epidemic. Geneva, Switzerland: World Health Organization; 2000.
4. Christensen A, Davidsen M, Ekholm O, Hansen S, Holst M, Juel K. In: Størup M, Hajlsted B, Raahauge Hvass L, editors. Den nationale sundhedsprofil 2010 – Hvordan har du det? 1.0 edition. Copenhagen: Sundhedsstyrelsen; 2011.
5. Holst M, Ekholm O, Juel K. Sundhedstilstanden blandt FOA medlemmer 2010 og udviklingen siden 2000. Copenhagen: Statens Institut for Folkesundhed; 2010.
6. Wyatt SB, Winters KP, Dubbert PM. Overweight and obesity: prevalence, consequences, and causes of a growing public health problem. *Am J Med Sci* 2006 Apr;331(4):166-174. [Medline: [16617231](#)]
7. Handbook of Obesity Treatment. New York: The Guilford Press; 2002:144-169.

8. van Duijvenbode DC, Hoozemans MJ, van Poppel MN, Proper K. The relationship between overweight and obesity, and sick leave: a systematic review. *Int J Obes (Lond)* 2009 Aug;33(8):807-816. [doi: [10.1038/ijo.2009.121](https://doi.org/10.1038/ijo.2009.121)] [Medline: [19528969](https://pubmed.ncbi.nlm.nih.gov/19528969/)]
9. Gjesdal S, Ringdal PR, Haug K, Maeland JG. Predictors of disability pension in long-term sickness absence: results from a population-based and prospective study in Norway 1994-1999. *Eur J Public Health* 2004 Dec;14(4):398-405. [doi: [10.1093/eurpub/14.4.398](https://doi.org/10.1093/eurpub/14.4.398)] [Medline: [15542877](https://pubmed.ncbi.nlm.nih.gov/15542877/)]
10. Burton J, World Health Organization. WHO healthy workplace framework and model: background and supporting literature and practices. Geneva: World Health Organization; 2010.
11. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obes Rev* 2010 Apr;11(4):306-321. [doi: [10.1111/j.1467-789X.2009.00646.x](https://doi.org/10.1111/j.1467-789X.2009.00646.x)] [Medline: [19754633](https://pubmed.ncbi.nlm.nih.gov/19754633/)]
12. Kodama S, Saito K, Tanaka S, Horikawa C, Fujiwara K, Hirasawa R, et al. Effect of Web-based lifestyle modification on weight control: a meta-analysis. *Int J Obes (Lond)* 2012 May;36(5):675-685. [doi: [10.1038/ijo.2011.121](https://doi.org/10.1038/ijo.2011.121)] [Medline: [21694698](https://pubmed.ncbi.nlm.nih.gov/21694698/)]
13. Wieland LS, Falzon L, Sciamanna CN, Trudeau KJ, Brodney S, Schwartz JE, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. *Cochrane Database Syst Rev* 2012 Aug 15(8):CD007675 [FREE Full text] [doi: [10.1002/14651858.CD007675.pub2](https://doi.org/10.1002/14651858.CD007675.pub2)] [Medline: [22895964](https://pubmed.ncbi.nlm.nih.gov/22895964/)]
14. Rongen A, Robroek SJ, van Lenthe FJ, Burdorf A. Workplace health promotion: a meta-analysis of effectiveness. *Am J Prev Med* 2013 Apr;44(4):406-415. [doi: [10.1016/j.amepre.2012.12.007](https://doi.org/10.1016/j.amepre.2012.12.007)] [Medline: [23498108](https://pubmed.ncbi.nlm.nih.gov/23498108/)]
15. Petersen R, Sill S, Lu C, Young J, Edington DW. Effectiveness of employee internet-based weight management program. *J Occup Environ Med* 2008 Feb;50(2):163-171. [doi: [10.1097/JOM.0b013e31815c6cf6](https://doi.org/10.1097/JOM.0b013e31815c6cf6)] [Medline: [18301173](https://pubmed.ncbi.nlm.nih.gov/18301173/)]
16. van Wier MF, Ariëns GA, Dekkers J, Hendriksen IJ, Smid T, van Mechelen W. Phone and e-mail counselling are effective for weight management in an overweight working population: a randomized controlled trial. *BMC Public Health* 2009 Jan 09;9:6 [FREE Full text] [doi: [10.1186/1471-2458-9-6](https://doi.org/10.1186/1471-2458-9-6)] [Medline: [19134171](https://pubmed.ncbi.nlm.nih.gov/19134171/)]
17. Cook RF, Billings DW, Hersch RK, Back AS, Hendrickson A. A field test of a web-based workplace health promotion program to improve dietary practices, reduce stress, and increase physical activity: randomized controlled trial. *J Med Internet Res* 2007 Jun 19;9(2):e17 [FREE Full text] [doi: [10.2196/jmir.9.2.e17](https://doi.org/10.2196/jmir.9.2.e17)] [Medline: [17581811](https://pubmed.ncbi.nlm.nih.gov/17581811/)]
18. Lyzinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. *J Pers Med* 2014 Jun 30;4(3):311-385 [FREE Full text] [doi: [10.3390/jpm4030311](https://doi.org/10.3390/jpm4030311)] [Medline: [25563356](https://pubmed.ncbi.nlm.nih.gov/25563356/)]
19. Safran N, Madar Z, Shahar D. The impact of a Web-based app (eBalance) in promoting healthy lifestyles: randomized controlled trial. *J Med Internet Res* 2015 Mar 02;17(3):e56 [FREE Full text] [doi: [10.2196/jmir.3682](https://doi.org/10.2196/jmir.3682)] [Medline: [25732936](https://pubmed.ncbi.nlm.nih.gov/25732936/)]
20. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012 Jul 10;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
21. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Harnessing different motivational frames via mobile phones to promote daily physical activity and reduce sedentary behavior in aging adults. *PLoS One* 2013;8(4):e62613 [FREE Full text] [doi: [10.1371/journal.pone.0062613](https://doi.org/10.1371/journal.pone.0062613)] [Medline: [23638127](https://pubmed.ncbi.nlm.nih.gov/23638127/)]
22. Toubro S, Astrup A. Randomised comparison of diets for maintaining obese subjects' weight after major weight loss: ad lib, low fat, high carbohydrate diet v fixed energy intake. *Br Med J* 1997 Jan 04;314(7073):29-34 [FREE Full text] [Medline: [9001476](https://pubmed.ncbi.nlm.nih.gov/9001476/)]
23. Skive L. Motionsmanualen. København: Komiteen for Sundhedsoplysning; 2003.
24. Nabe NK, Garde A, Borg V, Jensen J, Høgh A. Arbejdstider i ældreplejen. København: Det Nationale Forskningscenter for Arbejdsmiljø; 2016.
25. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. *Obesity (Silver Spring)* 2010 Feb;18(2):308-313 [FREE Full text] [doi: [10.1038/oby.2009.242](https://doi.org/10.1038/oby.2009.242)] [Medline: [19696764](https://pubmed.ncbi.nlm.nih.gov/19696764/)]
26. Collins CE, Morgan PJ, Jones P, Fletcher K, Martin J, Aguiar EJ, et al. A 12-week commercial web-based weight-loss program for overweight and obese adults: randomized controlled trial comparing basic versus enhanced features. *J Med Internet Res* 2012 Apr 25;14(2):e57 [FREE Full text] [doi: [10.2196/jmir.1980](https://doi.org/10.2196/jmir.1980)] [Medline: [22555246](https://pubmed.ncbi.nlm.nih.gov/22555246/)]
27. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. *J Med Internet Res* 2009 Jan 13;11(1):e1 [FREE Full text] [doi: [10.2196/jmir.1100](https://doi.org/10.2196/jmir.1100)] [Medline: [19141433](https://pubmed.ncbi.nlm.nih.gov/19141433/)]
28. Mossavar-Rahmani Y, Henry H, Rodabough R, Bragg C, Brewer A, Freed T, et al. Additional self-monitoring tools in the dietary modification component of the Women's Health Initiative. *J Am Diet Assoc* 2004 Jan;104(1):76-85. [doi: [10.1016/j.jada.2003.10.017](https://doi.org/10.1016/j.jada.2003.10.017)] [Medline: [14702588](https://pubmed.ncbi.nlm.nih.gov/14702588/)]
29. Tate D, Jackvony E, Wing R. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. *J Am Med Assoc* 2003 Apr 09;289(14):1833-1836. [doi: [10.1001/jama.289.14.1833](https://doi.org/10.1001/jama.289.14.1833)] [Medline: [12684363](https://pubmed.ncbi.nlm.nih.gov/12684363/)]

Abbreviations

ITT: intention-to-treat

SD: standard deviation

WHO: World Health Organization

Edited by G Eysenbach; submitted 14.06.16; peer-reviewed by C Collins, F Hillier-Brown; comments to author 28.10.16; revised version received 22.12.16; accepted 14.03.17; published 10.04.17.

Please cite as:

Balk-Møller NC, Poulsen SK, Larsen TM

Effect of a Nine-Month Web- and App-Based Workplace Intervention to Promote Healthy Lifestyle and Weight Loss for Employees in the Social Welfare and Health Care Sector: A Randomized Controlled Trial

J Med Internet Res 2017;19(4):e108

URL: <http://www.jmir.org/2017/4/e108/>

doi: [10.2196/jmir.6196](https://doi.org/10.2196/jmir.6196)

PMID: [28396303](https://pubmed.ncbi.nlm.nih.gov/28396303/)

©Nina Charlotte Balk-Møller, Sanne Kellebjerg Poulsen, Thomas Meinert Larsen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 10.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Web-Based Intervention for Physical Activity and Fruit and Vegetable Intake Among Chinese University Students: A Randomized Controlled Trial

Yan Ping Duan¹, PhD; Julian Wienert², PhD; Chun Hu¹, MEd; Gang Yan Si³, PhD; Sonia Lippke^{4,5}, PhD

¹Department of Physical Education, Faculty of Social Sciences, Hong Kong Baptist University, Hong Kong, China (Hong Kong)

²Institute of Social Medicine and Epidemiology, University of Lübeck, Lübeck, Germany

³Department of Health & Physical Education, The Education University of Hong Kong, Hong Kong, China (Hong Kong)

⁴Bremen International Graduate School of Social Sciences, Jacobs University Bremen, Bremen, Germany

⁵Department of Psychology & Methods, Jacobs University Bremen, Bremen, Germany

Corresponding Author:

Yan Ping Duan, PhD

Department of Physical Education

Faculty of Social Sciences

Hong Kong Baptist University

12/F, Hong Kong Baptist University Shek Mun Campus

8 On Muk Street, Shek Mun, Shatin

Hong Kong,

China (Hong Kong)

Phone: 852 34113080

Fax: 852 34113357

Email: duanyp@hkbu.edu.hk

Abstract

Background: Ample evidence demonstrates that university students are at high risk for sedentary behaviors and inadequate fruit and vegetable intake (FVI). Internet-based interventions for multiple health behavior appear to be promising in changing such unhealthy habits. Limited randomized controlled trials have tested this assumption among Chinese university students.

Objective: Our objective was to test the efficacy of an 8-week Web-based intervention compared with a control group condition to improve physical activity (PA) and FVI in Chinese university students. The intervention content was based on the health action process approach, and developed on the basis of previous evidence from the Western hemisphere. We evaluated self-reported data including PA and FVI, stages of change for PA and FVI, and motivational (risk perception, outcome expectancies, self-efficacy), volitional (action planning, coping planning, social support), and distal (intention, habit) indicators for PA and FVI, as well as perceived mental health outcomes (quality of life, depression).

Methods: In a randomized controlled trial, we recruited 566 university students from one university in the central region of China during their general physical education class. After random allocation and exclusion of unsuitable participants, we assigned 493 students to 1 of 2 groups: (1) intervention group: first 4 weeks on PA and subsequent 4 weeks on FVI, (2) control group. We conducted 3 Web-based assessments: at the beginning of the intervention (T1, n=493), at the end of the 8-week intervention (T2, n=337), and at a 1-month follow-up after the intervention (T3, n=142). The entire study was conducted throughout the fall semester of 2015.

Results: Significant time group interactions revealed superior intervention effects on FVI; motivational, volitional, and distal indicators of FVI; and PA behavior changes, with an effect size (η^2) ranging from .08 to .20. In addition, the overall intervention effects were significant for stage progression to the action group from T1 to T2 in PA ($\chi^2_1=11.75$, $P=.001$) and FVI ($\chi^2_1=15.64$, $P=.03$). Furthermore, the intervention effect was seen in the improvement of quality of life ($F_{3,492}=1.23$, $\eta^2=.03$, $P=.02$).

Conclusions: This study provides evidence for the efficacy of a Web-based multiple health behavior intervention among Chinese university students tested with different outcome variables. Future research should address the high dropout rate and optimize the most effective components of this intervention.

Trial Registration: Clinicaltrials.gov NCT01909349; <https://clinicaltrials.gov/ct2/show/NCT01909349> (Archived by WebCite at <http://www.webcitation.org/6pHV1A0G1>)

(*J Med Internet Res* 2017;19(4):e106) doi:[10.2196/jmir.7152](https://doi.org/10.2196/jmir.7152)

KEYWORDS

Web-based intervention; physical activity; fruit and vegetable intake; university students; motivational indicators; volitional indicators

Introduction

Being physically active and eating healthily are known to reduce the risk of developing noncommunicable diseases such as cancer, cardiovascular disease, obesity, and type 2 diabetes [1]. Few young people, however, reach the recommendations of health behavior for the prevention of noncommunicable diseases. Considerable evidence indicates that more than half of university students do not achieve at least 150 accumulated minutes of moderate physical activity (PA) per week [2,3]. Additionally, some studies show that university students and young adults aged 18-24 years consume fewer than the recommended 5 daily servings of fruits and vegetables [4,5].

University students are in the transition stage from late adolescence to adulthood. Unhealthy lifestyle habits, which are reinforced during this phase, often persist into later life and lead to long-term negative health outcomes [6]. Strategies to foster a healthy lifestyle in university students are therefore essential. Some evidence reveals that multiple behavior health interventions can promote both exercise and healthy diets in university students [1,7,8]. Thus, behavioral interventions for university students are required to support long-lasting behavior changes beyond late adulthood. Most of these studies, however, have been conducted only in Western societies.

Using Internet technology to promote health behavior change has several advantages [9]. In particular, an Internet intervention can (1) be delivered to large numbers of people at a low cost, (2) ensure that the intervention is accessible at any time and any location, and (3) provide follow-ups and feedback with personalized and tailored methods. In addition, the use of Internet technology is particularly relevant to young adults, who are the major users of such technology [10]. Moreover, tailored Web-based health behavior interventions that try to address specific characteristics of participants by providing information and feedback based on previously provided information (eg, assessed via questionnaire) harness the potential to provide a more user-specific experience with higher relevance for each participant [11,12]. A meta-analysis by Lustria et al [12] concluded that tailored Web-based interventions could lead to significant improvement in health outcomes at posttesting ($d=.14$) and follow-up ($d=.16$). Such tailored Web-based behavioral interventions, however, have very rarely been systematically evaluated in China. To the best of our knowledge, this study is the first test of a Web-based multiple behavior health intervention delivered to Chinese university students.

The Health Action Process Approach

The theoretical backdrop of this study is the health action process approach (HAPA; [13]), which divides the health

behavior change process into 2 phases. First is the motivational phase, in which people who do not intend to change their behavior (nonintenders) are motivated to develop their intentions. Afterward, they enter the second, or volitional, phase, in which people initiate and perform the behavior. Within the volitional phase, a distinction can be made between people who have the intention to perform a specific behavior but do not act (intenders) and people who already perform the behavior (actors). These distinctions allow for the interventions to address those variables that are relevant to specific processes in the sequential order. For example, before people can change unhealthy habits, they must become motivated to do so. Thus, such individuals may benefit most from interventions that increase risk perception, self-efficacy, and the promotion of positive outcome expectancies [13,14]. The idea is to lead the individual toward an explicit intention, by increasing their awareness of potential risks, highlighting the positive effects of a healthy lifestyle and addressing incorrect beliefs about negative effects. Upon forming the intention, people enter the volitional process. Individuals here benefit most from action and coping planning interventions [15], as such interventions help them translate their plans into behavior. Once people start to perform a healthy behavior, self-regulatory skills are most relevant for their maintenance progress. The behavior is mainly directed by self-efficacy, which regulates how effort is invested and how persistence is managed if barriers and setbacks occur. In addition, promoting perceived social support from the individuals' social environment is equally important in preventing relapse [13].

Study Objective

On the basis of this theory, previous interventions have been developed and conducted with people interested in reducing their cardiovascular risk in the Western hemisphere, where the intervention effects have been well supported [16,17]. This study aimed at testing the effectiveness of such an intervention for Chinese students. Our objective was to test the efficacy of an 8-week Web-based intervention (first 4 weeks on PA, and subsequent 4 weeks on fruit and vegetable intake [FVI]) compared with a control group condition to improve PA and FVI in Chinese university students. We examined the effects on behaviors, stages of change movement for PA and FVI, social-cognitive indicators for PA and FVI (motivational, volitional, and distal), and perceived mental health outcomes (quality of life and depression level).

We hypothesized that the main intervention effects would be (1) more behavioral change in PA and FVI (hypothesis 1a) and more stage progression to the action stage for PA and FVI (hypothesis 1b), (2) more improvements in motivational (risk perception, outcome expectancies, self-efficacy), volitional

(action planning, coping planning, social support), and distal (intention, habit) indicators of behavior change (hypothesis 2), and (3) an increase in positive mental health outcomes, including higher quality of life and lower depression levels (hypothesis 3).

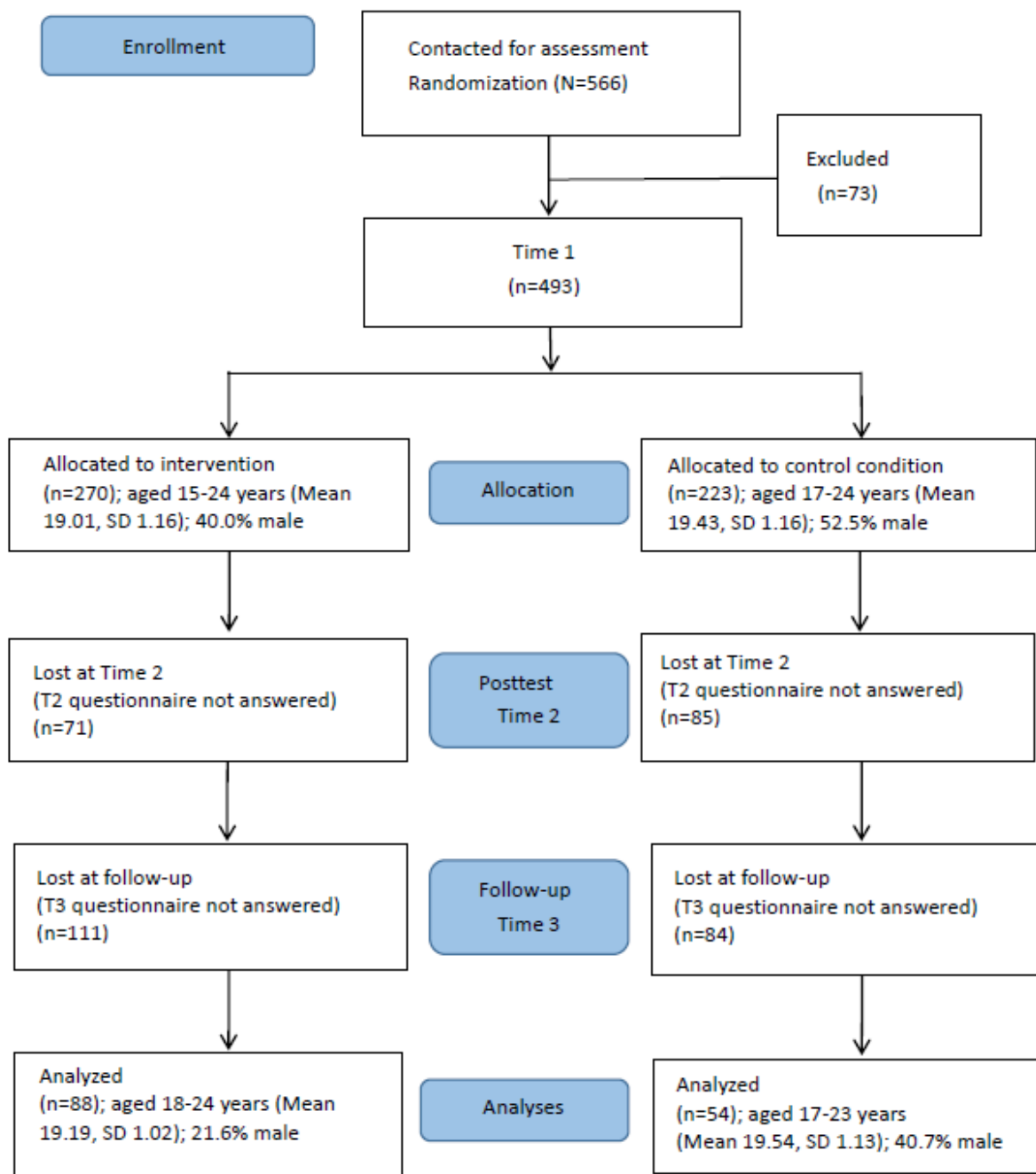
Methods

Participants and Procedure

Study participants were undergraduate students from one university in the central region of China. We contacted a total of 566 students in their first general physical education classes with the assistance of physical education lecturers during the fall semester of 2015. Of these, we excluded 73 (12.9%) after random allocation, including those who were collegiate athletes, had restrictions in terms of PA or FVI, or because they declined to participate. Subsequently, 493 (87.1%) students completed the online registration and provided personal information within 1 week, including 270 students in the intervention group (54.8%) and 223 students in the control group (45.2%). During the week after registration, we invited students in both groups to answer the first onscreen questionnaires (T1). Students in the intervention group were encouraged to participate in a Web-based program once a week during the following 8 weeks, while students in the control group did not receive any support.

Upon completion of the 8-week intervention, students in both groups were invited to fill in the second onscreen questionnaire (T2), followed by the third onscreen questionnaire administration 1 month (T3) after the intervention. By the end of the 8-week intervention, 337 (68.4%) students were still participating, with 199 (59.1%) in the intervention group and 138 (41.0%) in the control group. The final longitudinal sample consisted of 142 students (28.8%), including 88 (62.0%) in the intervention group and 54 (38.0%) in the control condition (Figure 1). All website links for the questionnaire surveys at T1, T2, and T3, as well as for the weekly intervention program, were delivered via email. To boost the engagement of students, we sent text message reminders and physical education lecturers verbally reminded participants during class during the weekly intervention process and 3 measurement points. Moreover, students were offered an additional 5 marks to their final physical education examination score as an incentive in exchange for their participation at all 3 data collection waves.

All students were informed about the purpose of the study with an informed consent form. The study procedure was approved by the Committee for the Use of Human & Animal Subjects in Teaching & Research of Hong Kong Baptist University, as well as the Deutsche Gesellschaft für Psychologie in Germany (EK-A-SL022013), and was registered with ClinicalTrials.gov (NCT01909349; Multimedia Appendix 1 [18]).

Figure 1. Flowchart of participant progress throughout the study phases.

Intervention

The intervention comprised an 8-week Web-based intervention targeting social-cognitive indicators for health behavior change for PA and FVI, based on the HAPA model [13]. The intervention applied behavior change techniques such as providing information about behavioral risks and the benefits of behavior change, motivating the formation of intentions, prompting identification of barriers, prompting setting of specific goals, reviewing behavioral goals, providing feedback on performance, prompting practice and presenting follow-up prompts, motivating the planning of social support, and prompting relapse prevention [19]. Weeks 1-4 aimed at PA and

weeks 5-8 aimed at FVI, as PA might act as a gateway behavior [20,21] and PA modules are the most favored ones in tailored eHealth lifestyle promotion [22]. Participants received 1 session per week, which lasted about 20 minutes.

The first session for each behavior targeted risk perception, outcome expectancies, and goal setting, and the second session targeted the development of action plans. The third session targeted the revision and adjustment of previously developed action plans, as well as the development of coping plans, while the fourth session targeted the revision and adjustment of previously developed coping plans and social support (Table 1).

Table 1. Intervention content and techniques for each week.

Session content	Physical activity				Fruit and vegetable intake			
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Session 1								
Risk perception	Yes				Yes			
Outcome expectancies	Yes				Yes			
Goal setting	Yes				Yes			
Session 2								
Development of action plans		Yes				Yes		
Session 3								
Revision of action plans			Yes				Yes	
Development of coping plans			Yes				Yes	
Session 4								
Revision of coping plans				Yes				Yes
Social support				Yes				Yes
Behavior change techniques [19]	Information on behavioral risks and benefits of behavior change, motivating intention formation, prompting specific goal setting	Feedback on performance, reviewing behavioral goals, prompting practice	Feedback on performance, presenting follow-up prompts, prompting relapse prevention, prompting barrier identification	Feedback on performance, prompting relapse prevention, motivating the planning of social support	Information on behavioral risks and benefits of behavior change, motivating intention formation, prompting specific goal setting	Feedback on performance, reviewing behavioral goals, prompting practice	Feedback on performance, presenting follow-up prompts, prompting relapse prevention, prompting barrier identification	Feedback on performance, prompting relapse prevention, motivating the planning of social support

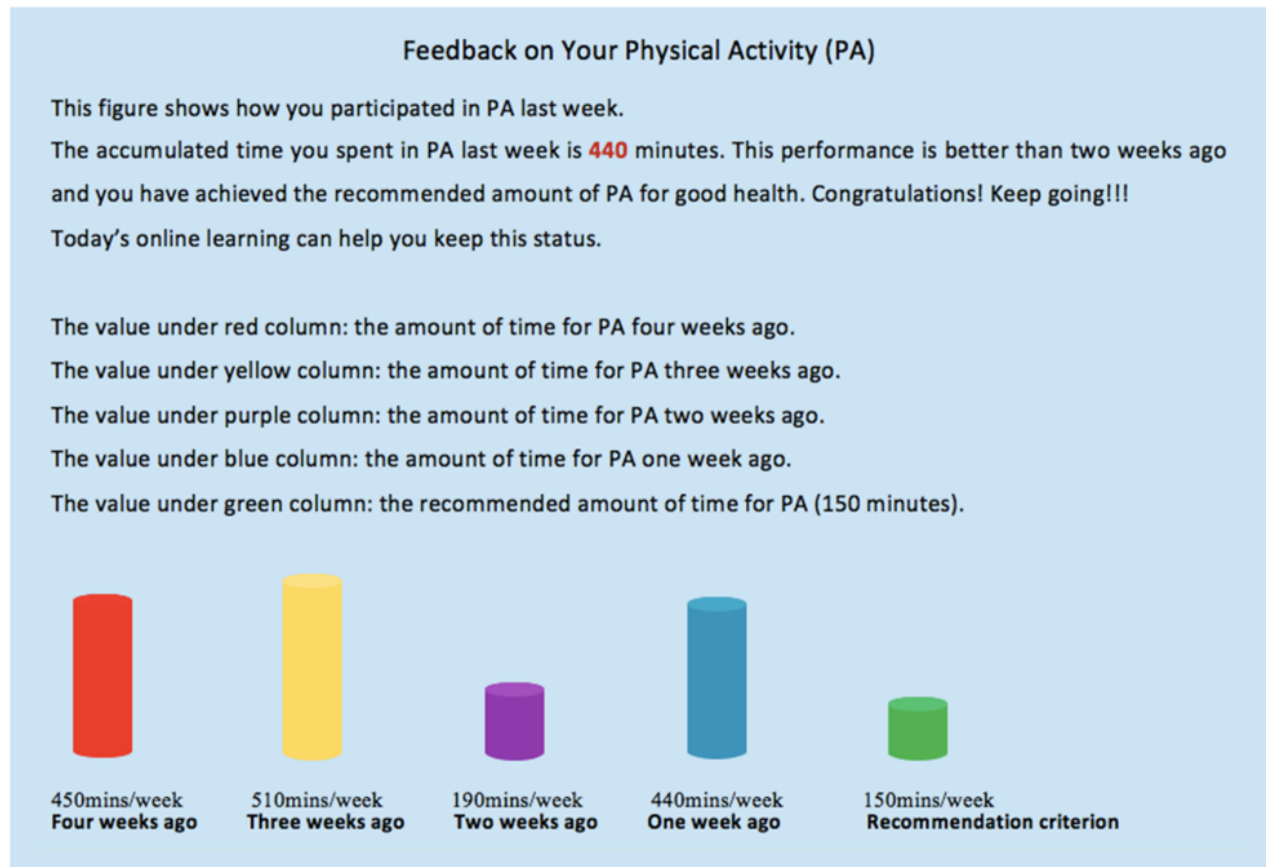
The intervention was not tailored based on the motivational and volitional stage of the HAPA. Instead, participants received tailored individualized feedback on their behavior improvement or decline at the beginning of each session, based on their prior self-report questionnaire. Moreover, during each session, participants received tailored normative feedback in the form of bar charts to compare their current behavioral performance and progress with population recommendations. Figure 2 illustrates a screenshot example (translated from Chinese) of the feedback information the participants received for PA over an accumulated amount of time (in the previous 4, 3, and 2 weeks, and the first week), as well as the PA recommendation criteria. A tailored comment was also presented, such as

You have spent an accumulated 440 minutes in PA last week. Great! This performance is better than two

weeks ago and you have achieved the recommended amount of PA for good health. Congratulations! Keep going!

In addition, we provided optional examples featuring role models throughout the intervention to support participants (eg, for setting goals or developing plans). We adopted a positive tone throughout the intervention.

The intervention in this study was based on previous research conducted in Germany and the Netherlands [23]. Prior to conducting the main study among Chinese university students, we completed preparation work on the intervention program in a prestudy, including the development and validation of adapted Chinese intervention materials, setting up intervention website modules, and optimizing website functioning by implementing Web-based tests with a small sample size.

Figure 2. Example of individual and normative feedback relating to time spent engaging in physical activity.

Measures

Demographic Information

Items addressed demographic characteristics, such as sex, age, and relationship status (single or in a relationship). We also collected self-reported body height (in centimeters) and body weight (in kilograms).

Health Behaviors

We assessed PA using the Chinese short version of the International Physical Activity Questionnaire (IPAQ-C) [24,25]. The IPAQ-C asked participants to estimate the number of days and amount of time spent on vigorous, moderate, and walking activities during the past 7 days. An individual total PA score (in minutes/week) was obtained when all questions were summed [25].

FVI during the past 7 days was assessed with 4 items: raw vegetables, fruits, fruit or vegetable juice, and cooked or steamed vegetables [26]. For each item, participants were asked to count the number of portions or glasses of liquid fruit and vegetables they consumed on average during a typical day. Each item had 11 options about the number of portions, such as 0, 0.5, 1, 1.5, 2, 2.5...5 or above. The total consumed portion was the sum of each item.

Stages of behavioral change were assessed for PA and FVI, each with 1 item on a 5-point scale asking "Did you engage in physical activity at least 5 days a week for 30 minutes or more (or 2.5 hours during the week)?" and "Please think about what

you have typically consumed during the last weeks: did you eat 5 portions of fruit and vegetables per day?" (1=no, and I do not intend to start; 2=no, but I am considering it; 3=no, but I seriously intend to start; 4=yes, but only for a brief period of time; 5=yes, and for a long period of time) [27]. People who selected point 1 were nonintenders in stage 1, those who selected point 2 or point 3 were intenders in stage 2, while those who selected point 4 or point 5 were actors in stage 3.

Motivational Indicators of Behavior Change

We adapted the risk perception scale from Perloff and Fetzner [28]. Items started with the stem "How likely is it that you will have at some time in your life...", followed by 5 items: "a high cholesterol level?," "a heart attack?," "high blood pressure?," "a stroke?," and "a cardiovascular disease?" Response used a 7-point scale ranging from 1=very unlikely to 7=very likely (Cronbach alpha=.84).

We assessed positive and negative outcome expectancies for PA with 2 items, each on a 5-point scale (1=don't agree at all; 5=totally agree), such as "If I am physically active 5 days a week for 30 minutes or more, then..." "I feel better afterward," or "it will cost me a lot of time" (positive: $\rho=.78$; negative: $\rho=.57$) [29]. We assessed positive and negative outcome expectancies for FVI, each with 2 items, such as "If I eat at least 5 portions of fruit and vegetables, then..." "this is good for my health," or "this will be a financial burden" (positive: $\rho=.88$; negative: $\rho=.74$) [29].

We assessed self-efficacy for PA with 5 items on a 5-point scale (1=don't agree at all; 5=totally agree), such as "I am certain

that I can permanently be physically active for at least 5 days a week for 30 minutes each day” (Cronbach alpha=.88) [30]. We assessed self-efficacy for FVI by 5 items on the basis of the PA scale, such as “I am certain that I can eat 5 portions of fruit and vegetables a day even if it is sometimes difficult” (Cronbach alpha=.92) [30].

Volitional Indicators of Behavior Change

We distinguished planning as action planning and coping planning. We assessed action planning by the stem “For the next month I have carefully planned...” followed by 3 items for PA, such as “which PA I will pursue,” or followed by 3 items for FVI, such as “what I will eat” (Cronbach alpha for PA=.86; Cronbach alpha for FVI=.91). We assessed coping planning by the stem “For the next month I have carefully planned...” followed by 3 items for PA, such as “what I can do in difficult situations to stick to my intentions,” or followed by 3 items for FVI, such as “how I can eat healthy, even if something happened” (Cronbach alpha for PA=.87; Cronbach alpha for FVI=.93). Answers were given on a 5-point scale ranging from 1=totally disagree to 5=totally agree [15].

We assessed social support with 3 items for PA (Cronbach alpha=.72) [31] and with 3 items for FVI (Cronbach alpha=.69) [31], such as “My partner helps me/my family helps me/my classmates and friends help me to stay physically active,” or “My partner helps me/my family helps me/my classmates and friends help me to eat healthy.” Answers were measured with a 5-point scale (1=not at all true; 5=exactly true).

Distal Indicators of Behavior Change

Regarding intention, for PA, we assessed 3 independent items with a 4-point scale (1=not true; 4=exactly true), representing different intensity levels of PA. “On 5 days a week for 30 minutes (or a minimum of 2.5 hours per week), I have the intention to perform...” “strenuous physical activity,” “moderate physical activity,” or “mild physical activity” (Cronbach alpha=.34) [27]. We assessed intention regarding FVI using 3 items: “I seriously intend to...” “eat at least five portions of fruit and vegetables daily,” “eat fruit and vegetables at every meal,” and “drink at least one fruit or vegetable juice every day” (Cronbach alpha=.73) [27].

We assessed the habit scale with the stem “Being physically active for at least 30 minutes on 5 days a week is something that...” and “Eating 5 portions of fruit and vegetables per day is something that...” followed by 2 items, such as “has become a confirmed habit” and “I do without thinking about it” (Cronbach alpha for PA=.95; Cronbach alpha for FVI=.87). Answers were indicated on a 5-point scale ranging from 1=strongly disagree to 5=strongly agree [32].

Mental Health Outcomes

We assessed quality of life using the Hong Kong version of the World Health Organization’s Quality of Life-BREF questionnaire [33]. We measured general quality of life with the question “How would you rate your quality of life?” with a 5-point scale (1=very poor; 5=very good). We also used the physical health subdomain with 7 items (Cronbach alpha=.71), such as “How satisfied are you with your ability to perform

your daily living activities?” (1=very dissatisfied; 5=very satisfied).

We assessed depression using the Chinese version of the Center for Epidemiologic Studies Depression (CES-D) scale, a self-report depression scale for research in the general population [34]. Participants were asked to indicate the frequency of symptoms on a 4-point scale (0=less than a day; 1=1 to 2 days; 2=3 to 4 days; 3=5 to 7 days) within the last week. Positively formulated items were reversed. The total score consisted of the sum of all 20 items and ranges from 0 to 60 (Cronbach alpha=.78). A CES-D score of 16 or greater indicated the likelihood of clinically significant depression.

Among all of the questionnaires above, IPAQ-C, quality of life, and depression had been well developed and validated in Chinese versions in previous studies [25,33,34]. The English questionnaires addressing motivational, volitional, and distal indicators of behavior change were translated into Chinese by a bilingual researcher. The translation was then validated using the standard back-translation technique [35]. Analysis of the data from the pilot test showed that the scales’ reliability was acceptable. In addition, we conducted pilot tests to ensure the usability and technical functionality of the electronic versions of the questionnaires prior to the main study.

Data Analysis

We conducted all data analyses using IBM SPSS version 23 (IBM Corporation). We used independent samples *t* tests and chi-square tests to analyze dropout and to compare baseline characteristics at T1. Statistical significance was set at the 5% level (2-tailed).

We tested intervention effects on PA and FVI behavior (hypothesis 1a) with a repeated-measures multivariate analysis of variance after screening the variables successfully. To evaluate hypothesis 1b, we first presented descriptive information (count and percentage) on stage distribution across T1, T2, and T3 between the intervention group and the control group for PA and for FVI. We then performed chi-square tests. To increase the cell sizes of stage groups for statistical significance, we collapsed the nonintender and intender groups into an inactive group and contrasted it to the active group. We evaluated stage movements in terms of cross-tabulating stage at T1 with T2 or T3 separately for the intervention versus the control group to compare the 2 groups with each other.

In addition, we tested the effects on combined motivational, volitional, and distal indicators of behavior change (hypothesis 2), as well as on mental health outcomes (hypothesis 3), by conducting a series of repeated-measures analyses of covariance with baseline behavior as the covariate. For both repeated-measures multivariate analysis of variance and repeated-measures analyses of covariance, we tested trends over time (T1, T2, and T3) as the within-participants factor, with treatment (intervention group vs control group) as the between-participants factor, adjusting for both baseline PA and FVI (as covariates).

We report results based on those individuals who participated in all 3 measurement points. Missing data were imputed within

each measurement point in time using the expectation-maximization method.

Results

Dropout Analysis

Results indicated that 493 people participated fully in data collection at T1, 337 at T1 and T2, and 142 at T1, T2, and T3. The dropout rates of participants were 31.6% (156/493) from T1 to T2 and 57.9% (195/337) from T2 to T3. The 142 final student sample included more women ($n=101$, 71.1%) than men, with a mean age of 19.3 years (range 17-24, SD 1.07). Most (128/142, 90.1%) of the participants were single. The average body mass index (BMI) of participants was 20.13 (SD 2.29) kg/m^2 . Participants at T1 and T2 ($n=337$) did not significantly differ from dropouts at T2 ($n=156$) with regard to relationship status ($\chi^2_1=0.06$, $P=.81$), age ($t_{491}=-0.05$, $P=.96$), BMI ($t_{491}=0.22$, $P=.96$), amount of PA at baseline ($t_{491}=1.20$, $P=.23$), and FVI at baseline ($t_{491}=-.87$, $P=.39$). More women than men participated fully in data collection at T2 ($\chi^2_1=10.67$, $P=.001$). In addition, those at T1, T2, and T3 ($n=142$) did not differ from dropouts at T2 and T3 ($n=351$) with regard to relationship status ($\chi^2_1=0.46$, $P=.49$), age ($t_{491}=-1.46$, $P=.15$), BMI ($t_{491}=1.12$, $P=.26$), and FVI at baseline ($t_{491}=-0.74$, $P=.46$). Those who participated in data collection at all 3 measurement points included more women than men ($\chi^2_1=22.60$, $P<.001$) and engaged in less PA at baseline ($t_{491}=2.26$, $P=.02$).

Randomization Check

There were no differences across the 2 groups at T1 regarding PA level, FVI, social-cognitive indicators (risk perception, outcome expectancies, self-efficacy, action plans, coping plans, social support, intention, and habit), relationship status, and BMI (all P values .23-.94). Sex and age did, however, differ between the 2 groups, with more women ($\chi^2_1=7.650$, $P=.01$) and younger participants ($t_{140}=-3.96$, $P<.001$) in the intervention group than in the control group.

Evaluation of Time and Treatment on PA, FVI, and Motivational, Volitional, and Distal indicators of Behavior Change

Table 2 presents the results, most of which were significant. Because the question of whether the intervention group and the control group developed differently over time is crucial, the interaction was of the highest interest. Out of 8 effects, 7 were statistically significant, with an effect size of η^2 ranging from .08 to .20. Only PA behavior was not statistically significant (see Table 2). However, the mean values of PA at different intensities (see Figure 3) indicated that there were descriptive differences between the 2 groups, which were clearly in favor of the intervention group. This was matched by the effects on FVI (see Figure 4), however, with significant effects ($\eta^2=.13$, $P<.001$).

Table 2. Effect sizes (η^2) and P values (significant at the 5% level, 2-tailed) for time, treatment group, and baseline behavior, as well as interaction.

Effects	Behavior ^a		Motivational ^b		Volitional ^c		Distal ^d	
	η^2	P value	η^2	P value	η^2	P value	η^2	P value
Physical activity (minutes/week)								
Time	.03	.03	.04	.49	.07	.12	.04	.22
Treatment	.01	.63	.05	.06	.11	<.001	.08	.01
Baseline behavior	N/A ^e	N/A	.14	<.001	.10	.04	.19	<.001
Interaction time treatment	.01	.95	.11	.01	.14	.01	.08	.02
Fruit and vegetable intake (portions/day)								
Time	.03	.15	.04	.52	.02	.81	.04	.22
Treatment	.06	.01	.06	.05	.05	.06	.08	.01
Baseline behavior	N/A	N/A	.18	<.001	.13	<.001	.21	<.001
Interaction time treatment	.13	<.001	.14	.01	.20	<.001	.11	.01

^aBehavior indicators: vigorous, moderate, and walking activity, or fruit and vegetable intake.

^bMotivational indicators: risk perception, outcome expectancies, self-efficacy.

^cVolitional indicators: action plans, coping plans, social support.

^dDistal indicators: intention, habit.

^eN/A: not applicable.

Figure 3. Performed physical activity (vigorous: left panel; moderate: middle panel; and walking: right panel) of the intervention group and the control group, in minutes per week, at 3 measurement points (T1: baseline; T2: end of intervention; and T3: 1-month follow-up).

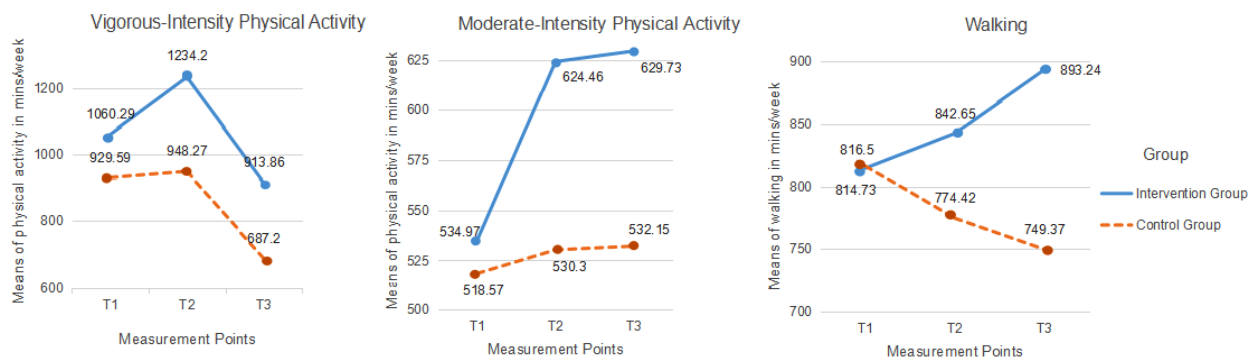
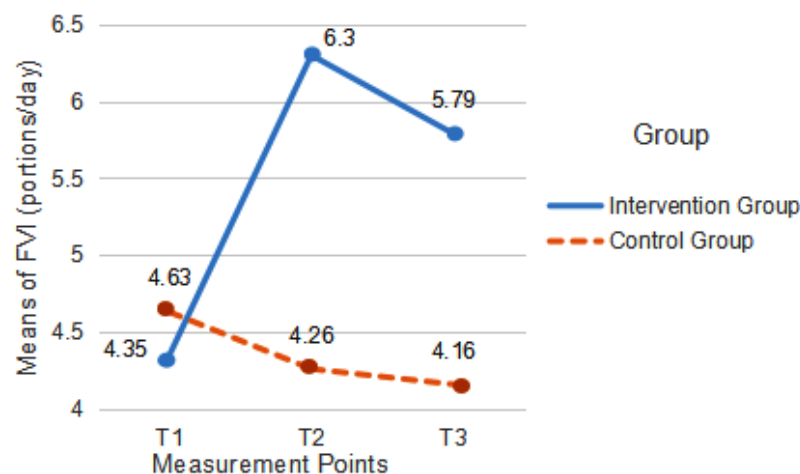


Figure 4. Fruit and vegetable intake (FVI) at T1 (baseline), T2 (end of intervention), and T3 (1-month follow-up) in the intervention group and the control group (portions per day).



Evaluation of Intervention Effects on Stage Progression to the Action Stage

Table 3 and Table 4 present the stage distributions at T1, T2, and T3 for PA and FVI. Descriptively, the stage developments underline that the intervention group was more likely than the control group to move to the action stage. After we collapsed the nonintender and intender groups as a single inactive group, there were at least 5 individuals in each cell for statistical significance testing. The findings validated previous results:

for PA, individuals who were inactive at T1 in the intervention group were more likely to move to the action stage at T2 ($\chi^2_{1}=18.57$, $P<.001$). Although descriptively this was also the case at T3, this intervention effect was not significant ($\chi^2_{1}=0.91$, $P=.34$). While the overall intervention effect for individuals being inactive and active at T1 was also significant from T1 to T2 ($\chi^2_{1}=11.75$, $P=.001$), no changes from T1 to T3 were statistically significant (all χ^2 were <2 and P values ranged from .12 to .27).

Table 3. Physical activity stage distributions at T1, T2, and T3 (n=142).

Stage	Intervention group		Control group		Total	
	n	% within group	n	% within group	n	% within group
At T1						
Nonintender	5	5.7	6	11.1	11	7.7
Intender	49	55.7	25	46.3	74	52.1
Actor	34	38.6	23	42.6	57	40.1
At T2						
Nonintender	2	2.3	8	14.8	10	7.0
Intender	30	34.1	23	42.6	53	37.3
Actor	56	63.6	23	42.6	79	55.6
At T3						
Nonintender	1	1.1	10	18.5	11	7.7
Intender	42	47.7	20	37.0	62	43.7
Actor	45	51.1	24	44.4	69	48.6

Table 4. Fruit and vegetable intake stage distributions at T1, T2, and T3 (n=142).

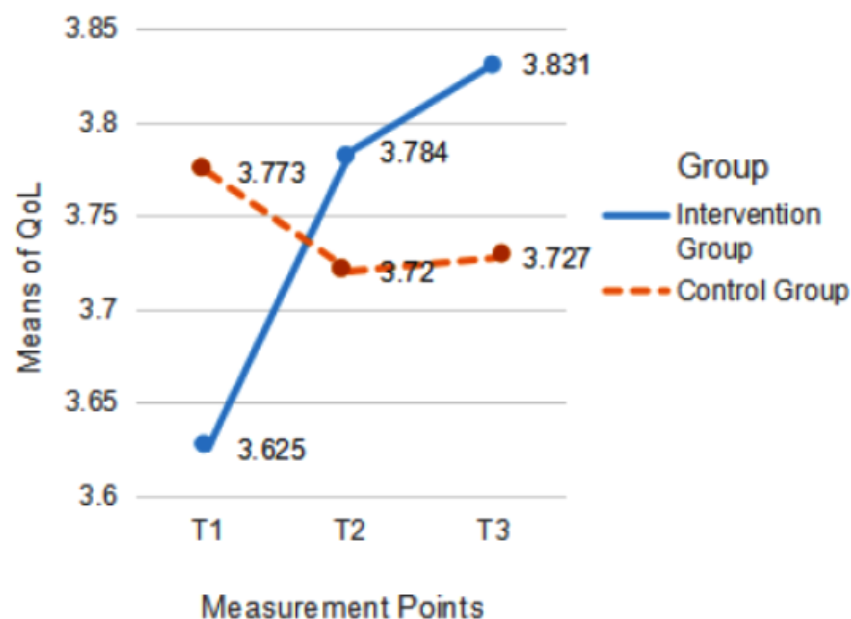
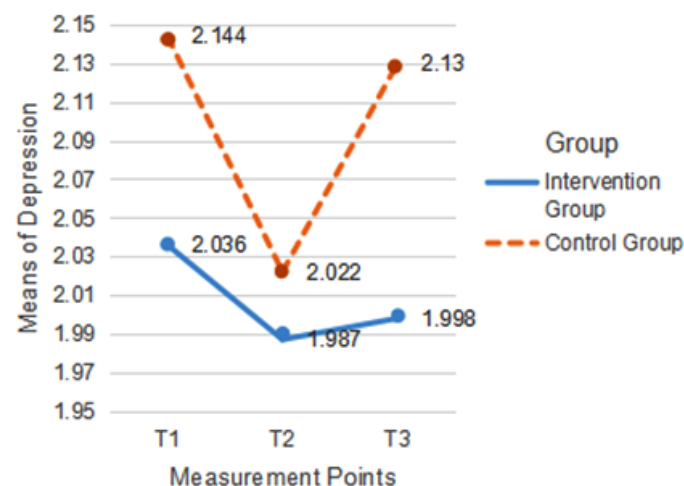
Stage	Intervention group		Control group		Total	
	n	% within group	n	% within group	n	% within group
At T1						
Nonintender	7	8.0	6	11.1	13	9.2
Intender	57	64.8	29	53.7	86	60.6
Actor	24	27.3	19	35.2	43	30.3
At T2						
Nonintender	1	1.1	8	14.8	9	6.3
Intender	34	38.6	23	42.6	57	40.1
Actor	53	60.2	23	42.6	76	53.5
At T3						
Nonintender	1	1.1	6	11.1	7	4.9
Intender	33	37.5	30	55.6	63	44.4
Actor	54	61.4	18	33.3	72	50.7

Conducting the same tests for FVI, the findings revealed more positive results, with significant changes between the intervention and the control group, both at T2 for previously inactive ($\chi^2_1=15.07$, $P<.001$) and active individuals ($\chi^2_1=4.28$, $P=.04$), and for the overall effect ($\chi^2_1=15.64$, $P=.03$). At T3, the effect was only significant in previously inactive individuals ($\chi^2_1=13.15$, $P<.001$), but not active ones ($\chi^2_1=1.35$, $P=.21$). The overall effect, however, was also significant ($\chi^2_1=10.52$, $P<.001$), suggesting the intervention's effectiveness.

Evaluation of Time and Treatment Effects on Mental Health Outcomes

Finally, we analyzed the intervention's effect on quality of life and depression. As no significant differences between the stages

emerged (tested for both behaviors), we left out the factor stage during the subsequent analysis. We found no significant differences for the group factor ($F_{3,139}=1.16$, $\eta^2=.02$, $P=.31$). The interaction of time and group was significant ($F_{3,139}=3.03$, $\eta^2=.08$, $P=.02$). The effects were, however, only due to quality of life ($F_{3,492}=1.23$, $\eta^2=.03$, $P=.02$). As for depression, the interaction was not significant ($F_{3,492}=0.17$, $\eta^2<.01$, $P=.48$). [Figure 5](#) presents the mean values for quality of life and [Figure 6](#) presents those for depression levels at the 3 measurement points for both groups.

Figure 5. Mean scores for quality of life (QoL) at T1, T2, and T3 in the intervention group and the control group.**Figure 6.** Mean scores for depression at T1, T2, and T3 in the intervention group and the control group.

Discussion

This study aimed to test the efficacy of an 8-week Web-based intervention compared with a control group condition to improve PA and FVI in Chinese university students. The majority of the study assumptions were supported.

Intervention Effects on Primary Outcomes

When testing the effects of time treatment on the 2 behavior test variables, 1 was found to be significant. In comparison with the control group, students in the intervention group reported

more consumption of fruit and vegetables over time. In addition, descriptively, the amounts of average FVI for the intervention group were all greater than the recommended amounts (5 portions per day) at the end of the 8-week intervention (6.3 portions) and at the 1-month follow-up (5.8 portions). This positive result among Chinese university students is consistent with a previous study, which was conducted with German and Dutch adults who wanted to reduce their cardiovascular risk [16,17]. Since we used intervention materials and a study design similar to the previous one, the generalization of intervention

effects on dietary behavior in this study can be warranted across Eastern and Western countries.

In terms of PA behavior, however, there was no significant interaction effect, which is not in line with other studies [16,17,36,37]. Reasons for this discrepancy may include, first, that PA levels of university students were overall relatively high at the start of the intervention (T1). According to the IPAQ-C scoring protocol, more than half of the students in this study (268/493, 54.3%) were classified as “sufficiently active” at T1, which means that individuals already participated in 3 or more days of vigorous activity for at least 20 minutes per day [24]. Therefore, ceiling effects might have influenced the intervention’s efficacy on PA change. Second, as young adults need long-term processes to establish PA behaviors as habits, the 4-week Web-based intervention dose might not have been sufficient to change their PA habits. The same issue can also be found in another Web-based PA intervention among university students [38]. Third, measuring PA after 8 weeks might have been a too-distal measurement point, as PA was addressed only throughout the first 4 weeks. Hence, PA levels could have dropped back by the time of the assessment. Taking these findings together, half of hypothesis 1a was supported.

With respect to the intervention effects on stage progression for the two behaviors, the results were positive. In comparison with students in the control group, students who were inactive in the intervention group reported more stage movements to the action stage for PA from T1 to T2. In addition, the intervention effects on stage progression for FVI were found both from T1 to T2 and from T1 to T3. It seems that the stage progression of multiple health behaviors was positively interrelated with each other in this study (T1-T2). A previous study revealed that PA and FVI appeared to facilitate rather than hinder each other [39]. There were cross-behavior associations between these two behaviors. This opinion was also supported by a study that found consistently significant correlations across stages between nutrition and PA (r range .16-.26, $P < 0.01$) [40]. Future studies should address the evaluation of mechanisms that transfer intervention effects on stage movement from PA to FVI, or vice versa. Referring back to hypothesis 1b, the data supported most of the assumption of stage progression.

Intervention Effects on Secondary Outcomes

When evaluating the outcomes of social-cognitive indicators of behavior change, all 6 tests revealed significant treatment effects on motivational, volitional, and distal indicators of PA and FVI over time, with the effect size ranging from .08 to .20. The findings are in line with previous studies, in which motivational and volitional interventions were both used to change multiple health behaviors [8,16,17]. In sum, our results support hypothesis 2.

The increase in positive mental health outcomes of quality of life and depression levels was evident in this study when combining the 2 measures together. This main effect, however, was due to the changes in quality of life only and not significantly to changes in depression. One possible reason for the lack of effect on depression might be that PA did not improve in this study, which cannot bring about mental health consequences. Another possible cause might be floor effects,

as university students reported low depression levels at the start of the study (mean 2.08, SD 0.61), which reflects the healthy mental status of participants in this study. To prevent depression in the long term, however, more components explicitly addressing mental health would need to be developed and tested in future interventions. Overall, half of hypothesis 3 was supported.

Limitations

Some limitations of this study need to be addressed. First, the dropout rate was high. Compared with male students, female students who performed less PA at baseline were more likely to comply with the instructions of the program and spent time completing data collection across the 3 time points. Thus, personal characteristics could be one explanation for the high dropout rate. Another reason might be related to the physical education lecturers’ verbal reminders in physical education class. In this study, data collection at T3 was scheduled during the week in which the last physical education class was offered. Students who attended the final physical education examination 1 week before the last class were allowed to be absent for the last physical education class, which means the lecturers could not verbally encourage those absent students to complete the final data collection. This might explain why the dropout rate from T2 to T3 (57.8%) was higher than that from T1 to T2 (31.6%). In addition, dropout and low engagement with questions could be caused by intervention features, such as the Web-based format and layout, length of the questionnaire, or browser difficulties on the intervention website [41-43]. Future studies should further address the characteristics of dropout and nonresponse to Web-based interventions, and find a solution to enlarge the sample size. This would also help provide more comprehensive subgroup analyses on the effectiveness of specific tailored components, as all participants would receive a unique intervention, but are treated equally when analyzing the data [44].

Second, the use of self-report questionnaires for behavioral outcome measures could have led to recall bias, overreporting or underreporting, and measurement errors. Thus, the inclusion of objective measures such as biomarkers, accelerometers, or pedometers is desirable in future studies. Third, the follow-up period was relatively short in this study. The longer-term impact of the intervention should be assessed in the future. Fourth, this study did not explore the mechanism of how the treatment facilitated multiple health behavior change (ie, the synthesis of PA and FVI). In other words, it is unclear which social-cognitive factors mediated the effect of the intervention on the improvements in health behavior. We advocate conducting an in-depth test in future studies.

Conclusions

This study provides evidence for the efficacy of a Web-based multiple health behavior intervention among Chinese university students. The majority of study hypotheses were supported. The initial findings suggest that the intervention was effective at increasing FVI and in enhancing perceived quality of life. In addition, all social-cognitive indicators of PA and FVI were improved in this study. The intervention did not, however, show the hypothesized effect on PA change. Future research is

warranted to address some of the limitations noted above, especially reducing high dropout rates and exploring the most effective components of the intervention, which is imperative to facilitate health promotion among university students, not only in the Western but also in the Eastern hemisphere.

Acknowledgments

JW previously worked at the Jacobs Center on Lifelong Learning and Institutional Development at Jacobs University Bremen, Germany, during the first phase of the work reported in this paper. This research was supported by a junior research group grant from the Wilhelm-Stiftung für Rehabilitationsforschung in Germany, and a Faculty Research Grant from Hong Kong Baptist University in Hong Kong (FRG2/15-16/032).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH form V1.6.

[[PDF File \(Adobe PDF File\), 527KB - jmir_v19i4e106_app1.pdf](#)]

References

1. Abu-Moghli FA, Khalaf IA, Barghoti FF. The influence of a health education programme on healthy lifestyles and practices among university students. *Int J Nurs Pract* 2010 Feb;16(1):35-42. [doi: [10.1111/j.1440-172X.2009.01801.x](#)] [Medline: [20158546](#)]
2. Hoyos I, Irazusta A, Gravina L, Gil SM, Gil J, Irazusta J. Reduced cardiovascular risk is associated with aerobic fitness in university students. *Eur J Sport Sci* 2011 Mar;11(2):87-94. [doi: [10.1080/17461391.2010.487116](#)]
3. Small M, Bailey-Davis L, Morgan N, Maggs J. Changes in eating and physical activity behaviors across seven semesters of college: living on or off campus matters. *Health Educ Behav* 2013 Aug;40(4):435-441 [[FREE Full text](#)] [doi: [10.1177/1090198112467801](#)] [Medline: [23232092](#)]
4. Chung S, Hoerr SL. Predictors of fruit and vegetable intakes in young adults by gender. *Nutr Res* 2005 May;25(5):453-463. [doi: [10.1016/j.nutres.2005.03.002](#)]
5. El Ansari W, Stock C, John J, Deeny P, Phillips C, Snelgrove S, et al. Health promoting behaviours and lifestyle characteristics of students at seven universities in the UK. *Cent Eur J Public Health* 2011 Dec;19(4):197-204 [[FREE Full text](#)] [Medline: [22432394](#)]
6. Friedman HS, Martin LR, Tucker JS, Criqui MH, Kern ML, Reynolds CA. Stability of physical activity across the lifespan. *J Health Psychol* 2008 Nov;13(8):1092-1104. [doi: [10.1177/1359105308095963](#)] [Medline: [18987082](#)]
7. Plotnikoff RC, Costigan SA, Williams RL, Hutchesson MJ, Kennedy SG, Robards SL, et al. Effectiveness of interventions targeting physical activity, nutrition and healthy weight for university and college students: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2015 Apr 01;12:45 [[FREE Full text](#)] [doi: [10.1186/s12966-015-0203-7](#)] [Medline: [25890337](#)]
8. Zhang Y, Cooke R. Using a combined motivational and volitional intervention to promote exercise and healthy dietary behaviour among undergraduates. *Diabetes Res Clin Pract* 2012 Feb;95(2):215-223 [[FREE Full text](#)] [doi: [10.1016/j.diabres.2011.10.006](#)] [Medline: [22055834](#)]
9. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. *Ann Behav Med* 2009 Aug;38(1):18-27 [[FREE Full text](#)] [doi: [10.1007/s12160-009-9133-4](#)] [Medline: [19802647](#)]
10. China Internet Network Information Center. Statistical report on internet development in China (January 2016). 2016. URL: <https://cnnic.com.cn/IDR/ReportDownloads/201604/P020160419390562421055.pdf> [accessed 2016-12-09] [[WebCite Cache ID 6mcNbSCy9](#)]
11. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. *Health Educ Res* 2008 Jun;23(3):454-466 [[FREE Full text](#)] [doi: [10.1093/her/cyn004](#)] [Medline: [18349033](#)]
12. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun* 2013;18(9):1039-1069. [doi: [10.1080/10810730.2013.768727](#)] [Medline: [23750972](#)]
13. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. *Appl Psychol* 2008 Jan;57(1):1-29. [doi: [10.1111/j.1464-0597.2007.00325.x](#)]
14. Schwarzer R, Cao DS, Lippke S. Stage-matched minimal interventions to enhance physical activity in Chinese adolescents. *J Adolesc Health* 2010 Dec;47(6):533-539. [doi: [10.1016/j.jadohealth.2010.03.015](#)] [Medline: [21094429](#)]

15. Lippke S, Schwarzer R, Ziegelmann JP, Scholz U, Schüz B. Testing stage-specific effects of a stage-matched intervention: a randomized controlled trial targeting physical exercise and its predictors. *Health Educ Behav* 2010 Aug;37(4):533-546. [doi: [10.1177/1090198109359386](https://doi.org/10.1177/1090198109359386)] [Medline: [20547760](https://pubmed.ncbi.nlm.nih.gov/20547760/)]
16. Reinwand DA, Crutzen R, Storm V, Wienert J, Kuhlmann T, de Vries H, et al. Generating and predicting high quality action plans to facilitate physical activity and fruit and vegetable consumption: results from an experimental arm of a randomised controlled trial. *BMC Public Health* 2016 Apr 12;16:317 [FREE Full text] [doi: [10.1186/s12889-016-2975-3](https://doi.org/10.1186/s12889-016-2975-3)] [Medline: [27066779](https://pubmed.ncbi.nlm.nih.gov/27066779/)]
17. Storm V, Dörenkämper J, Reinwand DA, Wienert J, De Vries H, Lippke S. Effectiveness of a Web-based computer-tailored multiple-lifestyle intervention for people interested in reducing their cardiovascular risk: a randomized controlled trial. *J Med Internet Res* 2016 Apr 11;18(4):e78 [FREE Full text] [doi: [10.2196/jmir.5147](https://doi.org/10.2196/jmir.5147)] [Medline: [27068880](https://pubmed.ncbi.nlm.nih.gov/27068880/)]
18. Eysenbach G, CONSORT- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
19. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol* 2008 May;27(3):379-387. [doi: [10.1037/0278-6133.27.3.379](https://doi.org/10.1037/0278-6133.27.3.379)] [Medline: [18624603](https://pubmed.ncbi.nlm.nih.gov/18624603/)]
20. Tucker M, Reicks M. Exercise as a gateway behavior for healthful eating among older adults: an exploratory study. *J Nutr Educ Behav* 2002;34 Suppl 1:S14-S19. [Medline: [12047824](https://pubmed.ncbi.nlm.nih.gov/12047824/)]
21. Carraça EV, Silva MN, Coutinho SR, Vieira PN, Minderico CS, Sardinha LB, et al. The association between physical activity and eating self-regulation in overweight and obese women. *Obes Facts* 2013;6(6):493-506 [FREE Full text] [doi: [10.1159/000356449](https://doi.org/10.1159/000356449)] [Medline: [24217426](https://pubmed.ncbi.nlm.nih.gov/24217426/)]
22. Schulz DN, Kremers SP, De Vries H. Tailored eHealth lifestyle promotion: which behavioral modules do users prefer? *J Health Commun* 2015;20(6):663-672. [doi: [10.1080/10810730.2015.1012243](https://doi.org/10.1080/10810730.2015.1012243)] [Medline: [25928161](https://pubmed.ncbi.nlm.nih.gov/25928161/)]
23. Reinwand D, Kuhlmann T, Wienert J, de Vries H, Lippke S. Designing a theory- and evidence-based tailored eHealth rehabilitation aftercare program in Germany and the Netherlands: study protocol. *BMC Public Health* 2013;13:1081 [FREE Full text] [doi: [10.1186/1471-2458-13-1081](https://doi.org/10.1186/1471-2458-13-1081)] [Medline: [24245493](https://pubmed.ncbi.nlm.nih.gov/24245493/)]
24. The IPAQ Group. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ): short and long forms. 2005 Nov. URL: <https://docs.google.com/viewer?a=v&pid=sites&scid=ZGVmYXVsdGRvbWVpbnx0aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWVRIZTM> [accessed 2017-12-08] [WebCite Cache ID 6manYMnit]
25. Macfarlane DJ, Lee CC, Ho EY, Chan KL, Chan DT. Reliability and validity of the Chinese version of IPAQ (short, last 7 days). *J Sci Med Sport* 2007 Feb;10(1):45-51. [doi: [10.1016/j.jsams.2006.05.003](https://doi.org/10.1016/j.jsams.2006.05.003)] [Medline: [16807105](https://pubmed.ncbi.nlm.nih.gov/16807105/)]
26. Rafferty AP, Anderson JV, McGee HB, Miller CE. A healthy diet indicator: quantifying compliance with the dietary guidelines using the BRFSS. *Prev Med* 2002 Jul;35(1):9-15. [Medline: [12079436](https://pubmed.ncbi.nlm.nih.gov/12079436/)]
27. Lippke S, Ziegelmann JP, Schwarzer R, Velicer WF. Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. *Health Psychol* 2009 Mar;28(2):183-193 [FREE Full text] [doi: [10.1037/a0012983](https://doi.org/10.1037/a0012983)] [Medline: [19290710](https://pubmed.ncbi.nlm.nih.gov/19290710/)]
28. Perloff LS, Fetzer BK. Self-other judgments and perceived vulnerability to victimization. *J Pers Soc Psychol* 1986;50(3):502-510. [doi: [10.1037/0022-3514.50.3.502](https://doi.org/10.1037/0022-3514.50.3.502)]
29. Schulz DN, Schneider F, de Vries H, van Osch LA, van Nierop PW, Kremers SP. Program completion of a web-based tailored lifestyle intervention for adults: differences between a sequential and a simultaneous approach. *J Med Internet Res* 2012;14(2):e26 [FREE Full text] [doi: [10.2196/jmir.1968](https://doi.org/10.2196/jmir.1968)] [Medline: [22403770](https://pubmed.ncbi.nlm.nih.gov/22403770/)]
30. Luszczynska A, Sutton S. Physical activity after cardiac rehabilitation: evidence that different types of self-efficacy are important in maintainers and relapsers. *Rehabil Psychol* 2006;51(4):314-321. [doi: [10.1037/0090-5550.51.4.314](https://doi.org/10.1037/0090-5550.51.4.314)]
31. Jackson J, Lippke S, Gray C. Stage-specific prediction of physical activity in orthopaedic patients after rehabilitation treatment. *Int J Sport Psychol* 2011;42(6):586-609.
32. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. *J Appl Social Psychol* 2003 Jun;33(6):1313-1330. [doi: [10.1111/j.1559-1816.2003.tb01951.x](https://doi.org/10.1111/j.1559-1816.2003.tb01951.x)]
33. Yao G, Wu C. Similarities and differences among the Taiwan, China, and Hong-Kong versions of the WHOQOL questionnaire. *Soc Indic Res* 2008 Sep 30;91(1):79-98. [doi: [10.1007/s11205-008-9326-4](https://doi.org/10.1007/s11205-008-9326-4)]
34. Rankin SH, Galbraith ME, Johnson S. Reliability and validity data for a Chinese translation of the Center for Epidemiological Studies-Depression. *Psychol Rep* 1993 Dec;73(3 Pt 2):1291-1298. [doi: [10.2466/pr0.1993.73.3f.1291](https://doi.org/10.2466/pr0.1993.73.3f.1291)] [Medline: [8115582](https://pubmed.ncbi.nlm.nih.gov/8115582/)]
35. Marin G, Marin B. Research With Hispanic Populations. Oakland, CA: Sage Publications; 1991.
36. Lippke S, Fleig L, Wiedemann AU, Schwarzer R. A computerized lifestyle application to promote multiple health behaviors at the workplace: testing its behavioral and psychological effects. *J Med Internet Res* 2015;17(10):e225 [FREE Full text] [doi: [10.2196/jmir.4486](https://doi.org/10.2196/jmir.4486)] [Medline: [26429115](https://pubmed.ncbi.nlm.nih.gov/26429115/)]
37. King AC, Castro CM, Buman MP, Hekler EB, Urizar GG, Ahn DK. Behavioral impacts of sequentially versus simultaneously delivered dietary plus physical activity interventions: the CALM trial. *Ann Behav Med* 2013 Oct;46(2):157-168 [FREE Full text] [doi: [10.1007/s12160-013-9501-y](https://doi.org/10.1007/s12160-013-9501-y)] [Medline: [23609341](https://pubmed.ncbi.nlm.nih.gov/23609341/)]

38. Epton T, Norman P, Dadzie A, Harris PR, Webb TL, Sheeran P, et al. A theory-based online health behaviour intervention for new university students (U@Uni): results from a randomised controlled trial. *BMC Public Health* 2014;14:563 [FREE Full text] [doi: [10.1186/1471-2458-14-563](https://doi.org/10.1186/1471-2458-14-563)] [Medline: [24903620](https://pubmed.ncbi.nlm.nih.gov/24903620/)]
39. Fleig L, Küper C, Lippke S, Schwarzer R, Wiedemann AU. Cross-behavior associations and multiple health behavior change: a longitudinal study on physical activity and fruit and vegetable intake. *J Health Psychol* 2015 May;20(5):525-534. [doi: [10.1177/1359105315574951](https://doi.org/10.1177/1359105315574951)] [Medline: [25903240](https://pubmed.ncbi.nlm.nih.gov/25903240/)]
40. Lippke S, Nigg CR, Maddock JE. Health-promoting and health-risk behaviors: theory-driven analyses of multiple health behavior change in three international samples. *Int J Behav Med* 2012 Mar;19(1):1-13 [FREE Full text] [doi: [10.1007/s12529-010-9135-4](https://doi.org/10.1007/s12529-010-9135-4)] [Medline: [21234735](https://pubmed.ncbi.nlm.nih.gov/21234735/)]
41. Brouwer W, Kroeze W, Crutzen R, de Nooijer NJ, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. *J Med Internet Res* 2011;13(1):e2 [FREE Full text] [doi: [10.2196/jmir.1639](https://doi.org/10.2196/jmir.1639)] [Medline: [21212045](https://pubmed.ncbi.nlm.nih.gov/21212045/)]
42. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
43. Reinwand DA, Crutzen R, Elfeddali I, Schneider F, Schulz DN, Smit ES, et al. Impact of educational level on study attrition and evaluation of web-based computer-tailored interventions: results from seven randomized controlled trials. *J Med Internet Res* 2015 Oct 07;17(10):e228 [FREE Full text] [doi: [10.2196/jmir.4941](https://doi.org/10.2196/jmir.4941)] [Medline: [26446779](https://pubmed.ncbi.nlm.nih.gov/26446779/)]
44. Wienert J, Kuhlmann T. A stitch in time saves nine: things to consider when tailoring your online intervention. *Eur Health Psychol* 2015;17(1):12-17 [FREE Full text]

Abbreviations

BMI: body mass index

CES-D: Center for Epidemiologic Studies Depression

FVI: fruit and vegetable intake

HAPA: health action process approach

IPAQ-C: International Physical Activity Questionnaire Chinese short version

PA: physical activity

Edited by G Eysenbach; submitted 22.12.16; peer-reviewed by R Schwarzer, S Alley; comments to author 26.01.17; revised version received 01.03.17; accepted 18.03.17; published 10.04.17.

Please cite as:

Duan YP, Wienert J, Hu C, Si GY, Lippke S

Web-Based Intervention for Physical Activity and Fruit and Vegetable Intake Among Chinese University Students: A Randomized Controlled Trial

J Med Internet Res 2017;19(4):e106

URL: <http://www.jmir.org/2017/4/e106/>

doi: [10.2196/jmir.7152](https://doi.org/10.2196/jmir.7152)

PMID: [28396306](https://pubmed.ncbi.nlm.nih.gov/28396306/)

©Yan Ping Duan, Julian Wienert, Chun Hu, Gang Yan Si, Sonia Lippke. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 10.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Mobile Phone-Delivered Cognitive Behavioral Therapy for Insomnia: A Randomized Waitlist Controlled Trial

Corine HG Horsch¹, PhD; Jaap Lancee², PhD; Fiemke Griffioen-Both³, PhD; Sandor Spruit³, MSc; Siska Fitrianie³, PhD; Mark A Neerincx¹, PhD; Robbert Jan Beun³, PhD; Willem-Paul Brinkman¹, PhD

¹Department of Intelligent Systems, Delft University of Technology, Delft, Netherlands

²Department of Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands

³Department of Information and Computing Sciences, Utrecht University, Utrecht, Netherlands

Corresponding Author:

Corine HG Horsch, PhD

Department of Intelligent Systems

Delft University of Technology

Mekelweg 4

Delft, 2628 CD

Netherlands

Phone: 31 152784145

Fax: 31 152787141

Email: corinehorsch@gmail.com

Abstract

Background: This study is one of the first randomized controlled trials investigating cognitive behavioral therapy for insomnia (CBT-I) delivered by a fully automated mobile phone app. Such an app can potentially increase the accessibility of insomnia treatment for the 10% of people who have insomnia.

Objective: The objective of our study was to investigate the efficacy of CBT-I delivered via the Sleepcare mobile phone app, compared with a waitlist control group, in a randomized controlled trial.

Methods: We recruited participants in the Netherlands with relatively mild insomnia disorder. After answering an online pretest questionnaire, they were randomly assigned to the app (n=74) or the waitlist condition (n=77). The app packaged a sleep diary, a relaxation exercise, sleep restriction exercise, and sleep hygiene and education. The app was fully automated and adjusted itself to a participant's progress. Program duration was 6 to 7 weeks, after which participants received posttest measurements and a 3-month follow-up. The participants in the waitlist condition received the app after they completed the posttest questionnaire. The measurements consisted of questionnaires and 7-day online diaries. The questionnaires measured insomnia severity, dysfunctional beliefs about sleep, and anxiety and depression symptoms. The diary measured sleep variables such as sleep efficiency. We performed multilevel analyses to study the interaction effects between time and condition.

Results: The results showed significant interaction effects ($P<.01$) favoring the app condition on the primary outcome measures of insomnia severity ($d=-0.66$) and sleep efficiency ($d=0.71$). Overall, these improvements were also retained in a 3-month follow-up.

Conclusions: This study demonstrated the efficacy of a fully automated mobile phone app in the treatment of relatively mild insomnia. The effects were in the range of what is found for Web-based treatment in general. This supports the applicability of such technical tools in the treatment of insomnia. Future work should examine the generalizability to a more diverse population. Furthermore, the separate components of such an app should be investigated. It remains to be seen how this app can best be integrated into the current health regimens.

Trial Registration: Netherlands Trial Register: NTR5560; <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=5560> (Archived by WebCite at <http://www.webcitation.org/6noLaUdJ4>)

(*J Med Internet Res* 2017;19(4):e70) doi:[10.2196/jmir.6524](https://doi.org/10.2196/jmir.6524)

KEYWORDS

insomnia; smartphone app; cognitive behavioral therapy; eHealth

Introduction

Approximately 10% of the western adult population have chronic insomnia [1]. People with insomnia experience difficulties falling asleep, staying asleep, or both, and as a consequence they are sleep deprived during the day [2]. For example, insomnia is associated with low levels of concentration, greater fatigue, and impaired cognitive functioning [3-5]. Another consequence of insomnia is an increased risk of developing mental disorders such as depression and anxiety [6,7], or physical disorders such as diabetes and high blood pressure [8,9]. Insomnia also leads to societal costs such as reduced productivity, higher levels of sick leave, and more accidents [10].

One of the most common nonpharmacological treatments for insomnia is cognitive behavioral therapy (CBT-I). CBT-I is an effective treatment in either a face-to-face [11-13] or a self-help format [14,15]. Recently it has become more common to offer these self-help formats via the Internet. A recent meta-analysis [16] demonstrated that Internet-delivered CBT-I showed large treatment effects (Cohen $d=1.0$) on the Insomnia Severity Index (ISI). In addition to the efficacy of computerized CBT-I (CCBT-I), this format has multiple other advantages over face-to-face treatments. Potentially it can save costs, because less therapist time is needed, and the treatment can be offered to a larger number of people who can go through the treatment in their own time.

Until now, studies on computerized treatments have been mostly limited to Web-based treatments. A possible next step is delivering CCBT-I via a mobile phone app. CCBT-I delivered via a mobile phone has similar advantages to existing CCBT-I, such as wide and easy accessibility, reduced stigma, and greater cost-efficiency [17], but it could potentially exceed those advantages because mobile phones are portable. People carry their phones with them all the time and they are ubiquitous, unobtrusive, and intimate. Therefore, an effective app-based treatment for insomnia would increase the possible coverage for CBT-I. Furthermore, mobile phones are rich in sensors, computationally powerful, and remotely accessible, which provides opportunities for personalization, ecological momentary access, and real-time tracking [18,19].

In the domain of sleep, several kinds of sleep apps have been studied. For example, there is an app that unobtrusively increases awareness of sleep hygiene recommendations [20], an app that applies active sleep sampling for measuring sleep [21], a social app that shares time in bed based on alarm usage [22], and an app that supports and was used alongside of face-to-face CBT-I. To the best of our knowledge, however, no studies have evaluated the efficacy of stand-alone CCBT-I apps. To bridge this gap, we conducted a randomized controlled trial to compare

a group using a CCBT-I-based app with a waitlist control group. We expected that this app would have an ameliorating effect on insomnia severity and sleep impairment compared with the waitlist control group, assessed by a sleep diary.

The app offered a sleep diary, a relaxation exercise, sleep restriction exercise, and sleep hygiene and education. Since sleep restriction is seen as the most effective exercise [23,24], it was the main focus of the app. The goal was to demonstrate the app's efficacy in a sample of patients with relatively mild insomnia in order to test the proof of principle before investigating it in a more severely affected population.

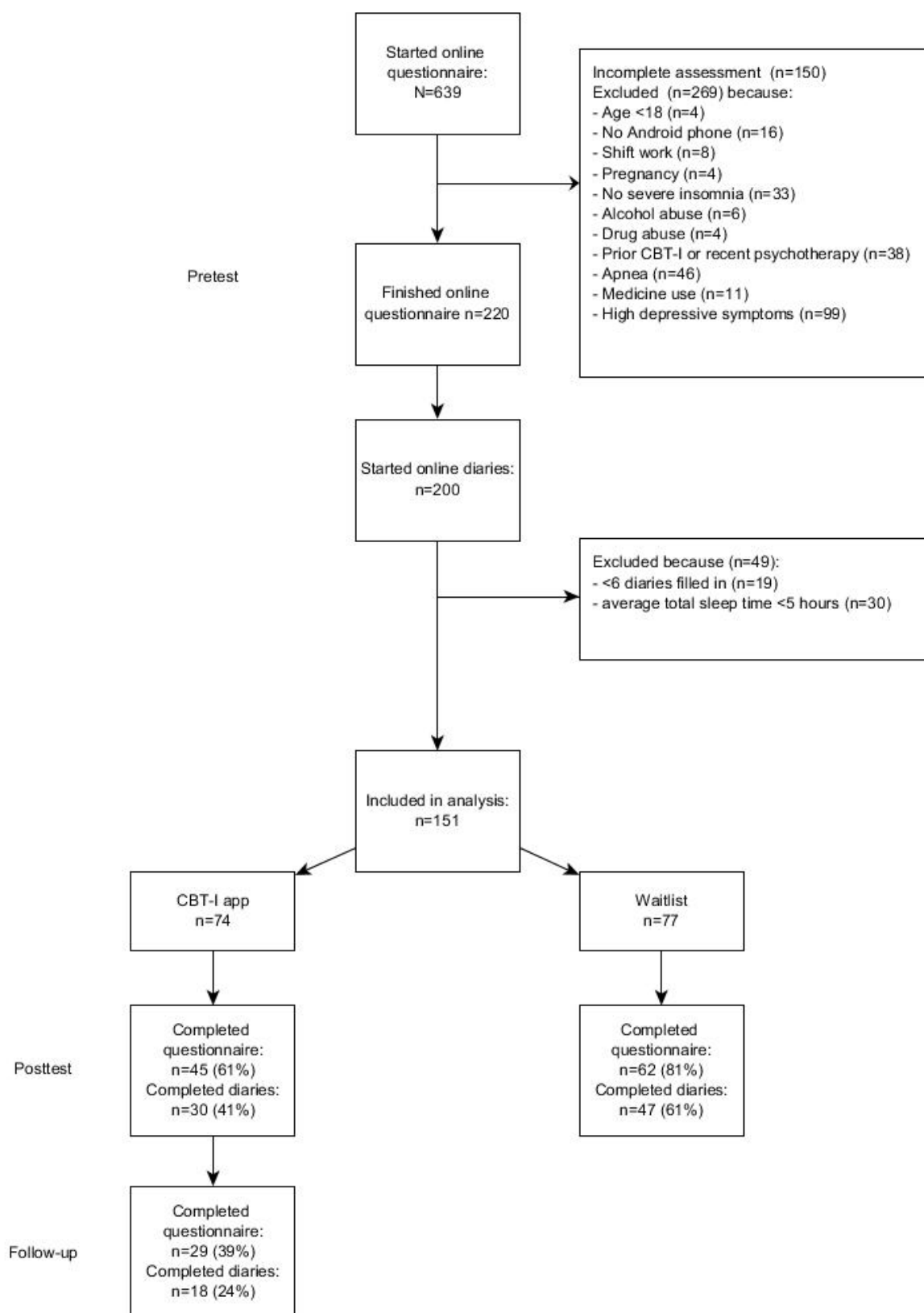
Methods

This study had a between-participants design with 2 arms: a waitlist condition and an intervention condition, with preintervention, postintervention, and 3-month follow-up measures.

Participants

We recruited participants from August 15 to October 21, 2015, via websites, social media, online advertisements, flyers, and a press release in the Netherlands. An initial group of 639 interested individuals completed an informed consent form and started the online questionnaire. Of this group, we excluded 269 people based on the inclusion and exclusion criteria (see [Figure 1](#); [Multimedia Appendix 1](#) [25]). Inclusion criteria were (1) difficulty with initiating or maintaining sleep for at least 30 minutes a night, for at least 3 nights a week, for at least 3 months, causing clinically significant distress or impairment in daily functioning, in accordance with the criteria for a *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) diagnosis of insomnia [2], (2) stable medication use, (3) aged ≥ 18 years, and (4) a valid email address, connected to the Internet, and in possession of an Android mobile phone (operating system version 4.1 or higher). Exclusion criteria were (1) ISI score [26] < 7 , (2) previous treatment with CBT-I (3) having started other psychotherapy in the last 6 months, (4) self-reported diagnosis of schizophrenia or psychosis, (5) alcohol or marijuana abuse (> 3 glasses of alcohol a day for at least 21 days a month, or use of marijuana more than once a week), (6) possible sleep apnea (determined with a subscale of the SLEEP-50 questionnaire; cut off ≥ 15 [27]), (7) shift work, (8) pregnant or breast-feeding, (9) symptoms of depression (determined with a subscale of the Centre for Epidemiological Studies Depression [CES-D] scale [28,29]; cutoff ≥ 27), or (10) total sleep time ≤ 5 hours on average as reported in a consecutive 7-day sleep diary prior to the experiment. We measured all inclusion and exclusion criteria, except the last one, using an online questionnaire asking the participants directly about the criteria (see [Multimedia Appendix 2](#)).

Figure 1. Consolidated Standards of Reporting Trial (CONSORT) flow diagram of recruitment, reasons for exclusion, and experimental compliance. CBT-I: cognitive behavioral therapy for insomnia.



Intervention

The Sleepcare app [30,31] was based on previously published protocols (eg, [32,33]) and followed a talk-and-tool design

principle, which is based on the idea that people interact symbolically and physically with their environment (see [34] for more details on the rationale and design decisions). The app packaged a sleep diary, a relaxation exercise, sleep restriction

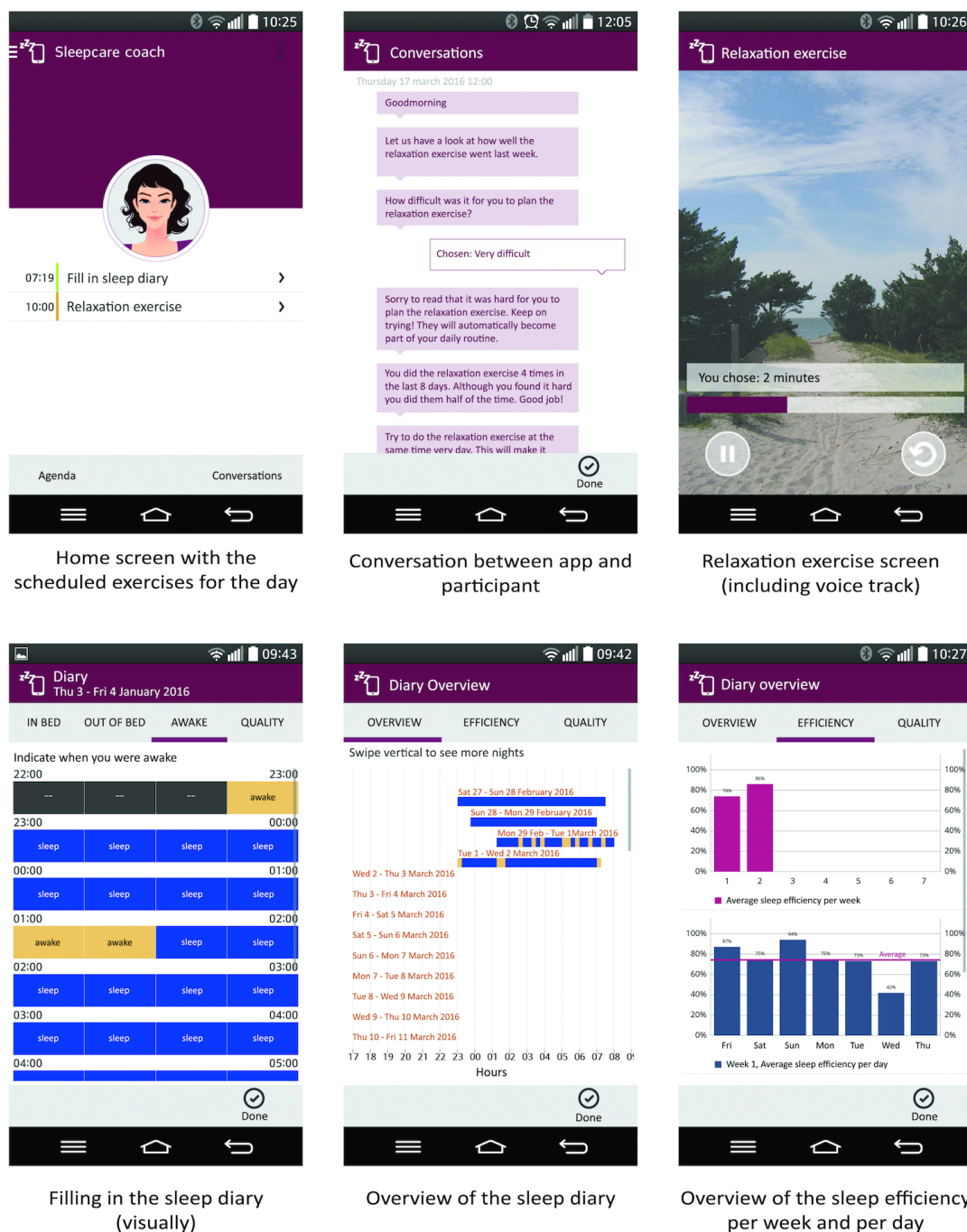
exercise, and sleep hygiene and education (Figure 2). The app offered these exercises in Dutch, adjusted them to the participant, and reminded participants to perform the exercises. The basic program duration was 6 to 7 weeks, depending on a participant's adherence. For example, if a participant had filled out <6 sleep diaries since starting the app, the app explained to the participant that the sleep restriction exercise could start only after they had completed 6 diaries. If they had completed <6 diaries, the introduction of the sleep restriction exercise was postponed until participants met this prerequisite. The app was fully automated and did not require any input from therapists or a human administrator. Automatic warnings were built in when participants slept for <5 hours on average. The first

warning appeared after 5 days and warned against activities such as driving a car while feeling sleepy. Follow-up warnings also included a referral to the general practitioner, and the app automatically stopped the sleep restriction exercise.

The app consisted of a home screen that displayed the scheduled exercises for that day. Furthermore, there was a menu, a calendar, and a conversation screen (Figure 3). The menu provided access to all components of the app and the CBT-I exercises. The calendar displayed all the scheduled activities for the whole 7 weeks, which the participants could browse through at any time. The app interacted with the participants via dialogues on the conversation screen.

Figure 2. Sleepcare app treatment protocol.



Figure 3. Mock-up screenshots of the Sleepcare app translated from Dutch.

Conversations

The conversation screen (Figure 3) displayed the dialogues between the app and the participants, which were inspired by face-to-face consultations. New exercises were introduced, and the progress of the participants was evaluated. Typically, the app gave information and asked multiple-choice questions. Participants could open new conversations only after the previous conversation was finished. Conversations were initiated by the app based on a participant's adherence and progress. For example, if a relaxation exercise was done <3 times within 4 days of its introduction, a conversation would start to address the participant's adherence. Additionally, the app started conversations based on a participant's progress; for example,

after the participant had filled out the sleep diary for a week, the app calculated and showed weekly sleep efficiency averages in an evaluation conversation. A detailed description of the underlying design principles can be found in Beun et al [34].

CBT-I Exercises

Sleep Diary

The sleep diary was a visual translation of the core Consensus Sleep Diary [35] consisting of 4 screens asking participants to fill in their bedtimes and wake times, and their subjective sleep quality. The sleep diary could only be filled out for the previous night. Connected to the sleep diary was the sleep overview, which gave a visual summary of the participant's sleep.

Relaxation Exercise

Relaxation exercises have a long history for treating insomnia. Based on the idea that hyperarousal is a main determinant of insomnia, relaxation exercises are the designated treatment [36]. The relaxation exercise was a progressive muscle relaxation exercise of varying durations, ranging from 1 to 16 minutes. The participants could choose the length of the exercise themselves before starting the exercise. By being offered short exercises, participants were able to gradually develop the habit of relaxing. The participants were guided by a voice track, which told them which muscles to contract and when to relax. The app advised participants to do the relaxation exercise once a day, although participants could do the exercises as often as they wanted [37].

Modified Sleep Restriction

Sleep restriction is seen as the most effective exercise [23,24], and therefore was the main focus of the app. After a week, the app introduced the modified sleep restriction exercise, on condition that participants had filled ≥ 6 sleep diaries and had an average sleep efficiency of $< 85\%$. We developed an algorithm to calculate the ideal and maximum time in bed for that specific participant based on the average sleep times of the previous week (FGB, unpublished data, 2016). The algorithm used rules such as (1) the ideal time in bed is equal to the average time in bed, (2) the maximum time in bed is always at least 1 hour less than the average time in bed, (3) the advised time in bed is never < 5 hours [38], and (4) the advised time in bed lies between average total sleep time and average time in bed. Participants had the opportunity to negotiate their sleep time. The app first suggested that the participant abide by the ideal time in bed. The participant was then given the option to accept that time in bed, to negotiate longer time in bed up to the calculated maximum time in bed, or to refrain from sleep restriction completely. Participants were allowed to negotiate about and refrain from the sleep restriction exercise in order to enhance self-empowerment, set realistic goals, and thereby increase adherence. Every week the app evaluated the adherence and effect of the sleep restriction exercise. When participants' sleep efficiency was $> 85\%$, they were allowed 15 minutes extra in bed. When a participant's average sleep efficiency was $< 85\%$, the app suggested the same or a further restriction of 15 minutes, depending on a participant's adherence. We used a modified sleep restriction protocol because we assumed this would increase the possibility of completing the restriction exercise, on the basis that more lenient sleep restriction is better than no sleep restriction at all.

Sleep Hygiene and Education

Sleep hygiene and education consist of lifestyle recommendations and knowledge about sleep. As a single-component intervention, it lacks efficacy [13]. However, it lays the foundation for CBT-I, since it increases participants' knowledge about sleep and the factors influencing sleep. By increasing understanding, sleep hygiene and education increase the efficacy of the other CBT-I exercises, and are therefore included in many multicomponent interventions [36]. Sleep hygiene and education information was presented on different screens as tips and facts in text format. The tips were divided

into 3 categories: (1) food and drink, (2) bedroom, and (3) behavior. "Use your bedroom only for sleeping, not for working" was an example of a bedroom tip. The sleep facts were categorized into 8 groups: (1) sleep cycles, (2) amount of sleep, (3) age, (4) animals, (5) disorders, (6) causes, (7) sleep medicine, and (8) fun facts. An example of a fact about age is "The amount of sleep a person needs is age dependent."

Persuasive Strategies

We implemented various kinds of persuasive strategies to support the participants' adherence. First, the app sent notifications for both the exercises and the conversations. So, for a scheduled exercise such as filling in the sleep diary, the app sent the participants a notification. For unfinished conversations, additional reminders were sent every day at noon. Second, the app provided room for negotiation about the sleep restriction assignment as described earlier. Furthermore, the app was designed to be easy to use and attractive, to improve adherence [31,39].

Measurements

Questionnaire Measures

Primary Measure

We measured the severity of insomnia with a Dutch translation of the widely used ISI. This is a 7-item questionnaire with scores ranging from 0 (no insomnia) to 28 (severe insomnia). A cutoff score of 7 determined relatively mild insomnia [26]. We set the cutoff score at this level instead of the clinically more relevant cutoff score of 10 [40] because the goal of this study was to demonstrate the feasibility of the app in a sample of patients with at least subclinical levels of insomnia.

Secondary Measures

The Pittsburgh Sleep Quality Index (PSQI) is a self-rating scale that measures sleep disturbances over a 1-month period [41]. The PSQI consists of 19 items with scores from 0 to 3 evaluating 7 subdomains. The scores of these subdomains are summed to calculate a global score ranging from 0 to 21. A global score > 5 indicates severe impairment in at least two domains, or moderate impairment in at least three domains.

We measured dysfunctional beliefs with the Dutch translation of the brief Dysfunctional Beliefs and Attitudes about Sleep (DBAS-16) scale [42]. The DBAS-16 consists of 16 statements with scores from 0 to 10 to indicate how much people agree with the statement. The average is calculated so that the total score ranges from 0 (no dysfunctional beliefs) to 10 (severe dysfunctional beliefs).

Anxiety symptoms were assessed with the 7 anxiety items of the Dutch version of the Hospital Anxiety and Depression Scale (HADS) [43,44]. The summed score ranges from 0 (no symptoms of anxiety) to 21 (severe symptoms of anxiety).

Depressive symptoms were measured using a Dutch translation of the CES-D scale. The CES-D consists of 20 items with scores ranging from 0 to 3, which are summed, with higher scores indicating more depressive symptoms [28,29].

Diary Measures

We used an online Dutch translated version of the consensus sleep diary [35]. Participants filled out the sleep diary for 7 days. In the diary they recorded the time they went to bed, the time they tried to go to sleep, their time of final awakening, their time out of bed, sleep onset latency, wake after sleep onset, terminal wakefulness, number of awakenings, sleep quality (1 = “very bad” to 10 = “very good”), and use of sleep medication. From these variables, we calculated the time in bed (time in bed=final arising time–time of going to bed), sleep time (total sleep time=time in bed–sleep onset latency–wake after sleep onset–terminal wakefulness), and sleep efficiency (sleep efficiency=[total sleep time/time in bed]×100). Sleep efficiency was the second primary measure in this trial.

Process Measures

We measured motivation to use the app with the Situational Motivation Scale [45] and acceptance of the app with the unified theory of acceptance and use of technology [46]. The focus of this paper is on the outcome measures, so we do not include the results of the process measures.

Procedure

Participants gave online informed consent and filled out the questionnaire addressing the inclusion and exclusion criteria, demographic information, and the outcome measures. The participants who met the study criteria received a link to an online sleep diary by email for 7 successive days. Emails for the diary were sent at 6.00 AM, and a reminder email was sent at 10.00 AM. We excluded participants who reported an average total sleep time of <5 hours and then randomly assigned the others to either the app or the waitlist condition. Randomization was carried out by a third party who was not part of this study. They used an online tool [47] to generate blocks of 20 participants. The list of the randomization sequence was kept in a locked office cupboard by the third party. After participants were assigned to a condition, participants and the principal investigator (CH) were no longer blinded to the condition allocation.

Then, 3 weeks after starting with the app or the waitlist condition, all participants received an interim measurement consisting of the ISI and DBAS-16, supplemented with questions regarding motivation (Situational Motivation Scale) and app acceptance for the app group. These interim measures are not reported in this paper. Both groups received a postintervention questionnaire, 7 weeks after random assignment, consisting of all the outcome measures (ISI, PSQI, DBAS-16, CES-D, and HADS) and a 7-day diary. In addition, participants in the app group received questions regarding the effect and utility of the app, which are not reported in this paper. After completing the diary, participants in the waitlist condition received the app. Participants in the app condition additionally received a 3-month follow-up questionnaire and diary.

The study was approved by the internal Ethical Review Board of the University of Amsterdam, and was registered with the Netherlands Trial Register (NTR5560).

Statistical Analysis

Required Statistical Power

To our knowledge, this study is the first large-scale randomized controlled trial to study an app to treat chronic insomnia, so expected effects were unknown. Earlier research about Web-based CBT-I found a Cohen $d > 1.0$ [48,49]. We were uncertain whether these large effects could also be obtained by an app, so we anticipated an average effect. A power calculation for a mixed analysis of variance design (effect: $f^2 = 0.15$, power 80%, $\alpha = .05$) indicated that a total of 90 participants were needed to detect a potential difference between the 2 conditions. As a meta-analysis showed that on average 50% of people adhere to technology-mediated insomnia treatment [50], the goal was set to include 180 participants.

Analyses

We tested the effects of the intervention using multilevel analyses, which allows for the inclusion of participants with 1 measurement and therefore is appropriate for intention-to-treat analyses [51]. Models were built in R version 3.1.3 (The R Foundation) to explore within-group (time), between-group (condition: app vs waitlist), and interaction (time × condition) effects. Model 0 is the basic model and includes only the participants as a random intercept. Model 1 adds the fixed factor time to model 0. Model 2 was built on model 1 and adds the condition as a fixed effect. Finally, model 3 adds the interaction effect between time and condition. Models 4 and 5 concern the premeasurement and follow-up data. Model 4 is the null model that includes only the participants as a random intercept. In model 5 time is added. Since there were no follow-up measurements for the waitlist, condition is not included. Dropout analyses for the postmeasurements showed that age, sleep quality, and terminal wakefulness were associated with nonresponse in the app condition. In the waitlist condition, number of awakenings was related to nonresponse. Dropout analyses for the follow-up measurements showed that terminal wakefulness was associated with nonresponse. Therefore, they were added as covariates in all models in the multilevel regression analyses [51]. We calculated chi-squares for the various models to compare the ability of the models to fit the data. Furthermore, we calculated R^2 values to indicate the level of variance explained by the level-1 variables [52]. R^2 values of .10 indicate a small effect, $R^2 = .30$ indicates a medium effect, and $R^2 = .50$ indicates a large effect [53].

To enhance comparability with other studies, we calculated between-group Cohen d values. Table 1 shows the means and effect sizes based on an imputed dataset. First, we used multiple imputation in IBM SPSS version 22 (IBM Corporation) to insert missing cases [54]. Data from 41–44 participants (27.2%–29.1%) were missing for the outcome questionnaires. Diaries were missing from 76 of the 151 participants (50.3%). The follow-up measurements were not imputed due to a large amount of missing data. For imputation, we used the pre- and post measures of the ISI, PSQI, DBAS-16, CES-D, HADS, sleep quality, sleep onset latency, wake after sleep onset, number of awakenings, time in bed, terminal wakefulness, total sleep time, and sleep efficiency, next to sex and age. With a predictive

mean matching procedure, 10 separate datasets were generated. The values in Table 1 are based on these imputed datasets. Second, we calculated Cohen d values by dividing the difference in change scores by the pooled standard deviation of that change score ($d = [\text{mean}_{\text{change score waitlist}} - \text{mean}_{\text{change score app}}] / \text{SD}_{\text{pooled}}$), whereby $\text{mean}_{\text{change score}} = \text{mean}_{\text{pre}} - \text{mean}_{\text{post}}$. We calculated within-group Cohen d values using the pre- and post scores per

condition and the pooled standard deviation ($d = [\text{mean}_{\text{pre}} - \text{mean}_{\text{post}}] / \text{SD}_{\text{pooled}}$). Additionally, we calculated within-group Cohen d values with the pre- and follow-up scores per condition and the pooled standard deviation ($d = [\text{mean}_{\text{pre}} - \text{mean}_{\text{follow-up}}] / \text{SD}_{\text{pooled}}$) (Table 2). A Cohen d of 0.20 indicates a small effect, 0.50 a moderate effect, and 0.80 a large effect [53].

Table 1. Observed baseline and imputed posttest means and effect sizes.

Measure	Group	Score, mean (SD)		Cohen <i>d</i>		95% CI (change scores)
		Baseline	Posttest	Within group	Between groups	
Questionnaire						
ISI ^a	WL ^b	16.4 (3.3)	13.2 (4.5)	0.80	−0.66	−0.99 to −0.33 ^c
	App	16.4 (3.1)	9.9 (4.9)	1.33		
DBAS-16 ^d	WL	5.2 (1.3)	4.8 (1.6)	0.24	−0.15	−0.47 to 0.17
	App	5.3 (1.3)	4.7 (1.4)	0.41		
CES-D ^e	WL	15.0 (5.8)	15.5 (9.5)	−0.06	−0.94	−1.28 to −0.61 ^c
	App	16.5 (6.0)	11.0 (5.6)	0.98		
HADS ^f	WL	5.6 (3.1)	6.2 (3.8)	−0.15	−0.75	−1.08 to −0.42 ^c
	App	6.1 (3.0)	4.1 (2.5)	0.81		
PSQI ^g	WL	10.6 (2.8)	9.7 (2.9)	0.32	−0.77	−1.10 to −0.44 ^c
	App	11.0 (2.8)	7.4 (3.3)	1.09		
Diary						
Sleep efficiency	WL	77.0 (8.2)	78.3 (7.6)	−0.16	0.71	0.37 to 1.04 ^c
	App	77.6 (7.3)	84.8 (5.3)	−1.37		
Time in bed	WL	500 (46)	513 (34)	−0.32	−0.55	−.88 to −0.22 ^c
	App	506 (44)	495 (31)	0.36		
Total sleep time	WL	386 (49)	401 (47)	−0.32	0.24	−0.08 to 0.56
	App	393 (52)	421 (37)	−0.74		
Sleep onset	WL	31 (21)	30 (19)	0.06	−0.45	−0.77 to −0.12 ^c
	App	33 (20)	20 (12)	1.01		
Wake after sleep onset	WL	44 (30)	44 (25)	−0.02	−0.70	−1.03 to −0.36 ^c
	App	45 (32)	24 (11)	1.79		
Terminal wakefulness	WL	37 (20)	37 (15)	0.02	−0.26	−0.58 to 0.07
	App	35 (22)	29 (13)	0.52		
Number of awakenings	WL	2.22 (1.14)	2.14 (1.08)	0.07	−0.38	−0.70 to −0.05 ^c
	App	1.94 (0.99)	1.46 (1.14)	0.42		
Sleep quality	WL	2.93 (0.52)	3.10 (0.55)	−0.33	0.29	−0.04 to 0.61
	App	2.97 (0.41)	3.33 (0.54)	−0.67		

^aISI: Insomnia Severity Index.^bWL: waitlist condition.^cThe confidence interval does not contain zero, meaning the effect apparently exists.^dDBAS-16: brief Dysfunctional Beliefs and Attitudes about Sleep.^eCES-D: Centre for Epidemiological Studies Depression.^fHADS: Hospital Anxiety and Depression Scale.^gPSQI: Pittsburgh Sleep Quality Index.

Table 2. Completers sample: baseline, posttest, and follow-up mean scores for the app condition and within-group effect sizes for the baseline (pre) to follow-up measurements.

Measure	Baseline score mean (SD)	Posttest score mean (SD)	Follow-up score mean (SD)	Cohen <i>d</i>
Questionnaire				
ISI ^a	16.4 (3.1)	9.8 (4.8)	10.0 (5.3)	1.20
DBAS-16 ^b	5.3 (1.3)	4.7 (1.4)	4.3 (1.8)	0.58
CES-D ^c	16.5 (6.0)	10.3 (5.3)	11.0 (7.2)	0.75
HADS ^d	6.1 (3.0)	4.0 (2.4)	4.3 (2.8)	0.67
PSQI ^e	11.0 (2.8)	7.6 (3.1)	9.1 (3.6)	0.53
Diary				
Sleep efficiency	77.6 (7.3)	83.8 (8.3)	83.8 (10.9)	−0.57
Time in bed	506 (44)	496 (50)	483 (39)	0.57
Total sleep time	393 (52)	417 (62)	403 (57)	−0.17
Sleep onset	33 (20)	22 (14)	21 (15)	0.80
Wake after sleep onset	45 (32)	27 (21)	25 (24)	0.84
Terminal wakefulness	35 (22)	31 (22)	35 (36)	0.13
Number of awakenings	1.94 (0.99)	1.58 (1.10)	1.75 (1.44)	0.01
Sleep quality	2.97 (0.41)	3.38 (0.51)	3.41 (0.60)	−0.74

^aISI: Insomnia Severity Index.^bDBAS-16: brief Dysfunctional Beliefs and Attitudes about Sleep.^cCES-D: Centre for Epidemiological Studies Depression.^dHADS: Hospital Anxiety and Depression Scale.^ePSQI: Pittsburgh Sleep Quality Index.

Results

Baseline Characteristics of the Sample

We randomly assigned 151 participants to the app ($n=74$) or a waitlist condition ($n=79$). Participants had a mean age of 39.66 years ($SD\ 13.44$; range 18–80). Of the total 151 participants, 94 were female (62.3%). Table 3 shows demographic information about the participants, as well as the randomization check. At baseline, the groups did not differ significantly on any demographic characteristics (all P values $>.05$).

Efficacy: Intention-to-Treat Analyses

Table 1 displays the mean scores for all the outcome measures and corresponding Cohen d values for the baseline and postmeasurements. Table 2 displays the mean scores for the follow-up measures. Figure 4 depicts the scores for the main outcome measures ISI and sleep efficiency. Table 4 and Table

5 present the results of the multilevel analyses. Table 4 shows the coefficients for model 3 and their standard errors, and whether the coefficients were significant. Table 5 shows whether there was a significant difference between the models (χ^2) and the level of variance explained (R^2). The multilevel analyses showed significant interaction effects between time and condition on the primary outcome measures ISI ($d=-0.66$) and sleep efficiency ($d=0.71$) at posttest. These effects indicate that the app was more effective than the waitlist condition. Furthermore, wake after sleep onset, number of awakenings, PSQI, CES-D, and HADS improved and showed significant interaction effects (Table 4 and Table 5), but sleep onset latency, time in bed, terminal wakefulness, total sleep time, and DBAS-16 showed no significant effects at posttest. At follow-up, improvements on all outcome measures remained significant, except for number of awakenings.

Table 3. Demographic characteristics of participants in the Sleepcare mobile phone app and waitlist control conditions (n=151).

Characteristics	Group		Statistic	P value
	Waitlist (n=74)	App (n=79)		
Age in years, mean (SD)	41 (13.9)	39 (13.0)	$t_{149}=1.02$.31
Sex, n (%)			$\chi^2_1=0.13$.72
Female	49 (64)	45 (61)		
Male	28 (36)	29 (39)		
Living together, n (%)			$\chi^2_1=0.26$.61
Yes	49 (64)	50 (68)		
No	28 (36)	24 (32)		
Employed, n (%)			$\chi^2_1=0.65$.42
Yes	56 (73)	58 (78)		
No	21 (27)	16 (22)		
Educational level, n (%)			$\chi^2_3=1.10$.78
Lower general secondary education	7 (9)	4 (5)		
Higher general secondary education	10 (13)	9 (12)		
Community college	11 (14)	9 (12)		
University	49 (64)	52 (70)		
Duration of insomnia in years, n (%)			$\chi^2_4=5.40$.25
<1	8 (10)	9 (12)		
1-5	38 (49)	27 (36)		
>5-10	10 (13)	13 (18)		
>10	12 (16)	20 (27)		
Unclear answer	9 (12)	5 (7)		
Insomnia due to a physical condition, n (%)			$\chi^2_1=0.20$.66
Yes	9 (12)	7 (9)		
No	68 (88)	67 (91)		
Used sleep medication, n (%)			$\chi^2_1=2.67$.10
Yes	3 (4)	8 (11)		
No	74 (96)	66 (89)		
Prescribed sleep medication, n (%)			$\chi^2_1=0.92$.35
Yes	3 (100)	6 (75)		
No	0 (0)	2 (25)		

Table 4. Multilevel analysis results of the diary and questionnaire variables for model 3: coefficients, standard errors, and *P* values^a.

Variables	Intercept	Time	Condition	Interaction
Diary variables				
Sleep efficiency				
B	77.38	1.97	−0.24	5.34
SE	0.78	0.99	1.11	1.55
<i>P</i> value	<.001	.05	.83	.001
Sleep onset latency				
B	32.75	−3.59	0.12	−7.86
SE	2.47	2.63	3.52	4.16
<i>P</i> value	<.001	.18	.97	.06
Wake after sleep onset				
B	42.87	−3.74	3.17	−22.42
SE	2.99	3.79	4.27	5.97
<i>P</i> value	<.001	.33	.46	<.001
Number of awakenings				
B	2.20	−0.17	−0.11	−0.48
SE	0.13	0.14	0.19	0.22
<i>P</i> value	<.001	.21	.55	.03
Time in bed				
B	499.74	9.75	5.46	−16.07
SE	5.02	5.22	7.16	8.28
<i>P</i> value	<.001	.07	.45	.06
Terminal wakefulness				
B	36.93	−0.44	−1.50	−3.23
SE	2.35	3.24	3.35	5.08
<i>P</i> value	<.001	.89	.65	.53
Total sleep time				
B	387.6	18.6	2.76	14.9
SE	5.53	6.18	7.89	9.77
<i>P</i> value	<.001	.004	.73	.13
Questionnaire variables				
Insomnia Severity Index				
B	16.26	−3.04	0.20	−3.53
SE	0.44	0.53	0.63	0.79
<i>P</i> value	<.001	<.001	.75	<.001
Pittsburgh Sleep Quality Index				
B	10.50	−0.97	0.56	−2.52
SE	0.33	0.36	0.47	0.54
<i>P</i> value	<.001	.008	.23	<.001
Dysfunctional Beliefs and Attitudes about Sleep				
B	5.17	−0.34	0.18	−0.31
SE	0.16	0.15	0.22	0.23
<i>P</i> value	<.001	.03	.41	.18

Variables	Intercept	Time	Condition	Interaction
Centre of Epidemiological Studies Depression scale				
B	14.95	0.66	1.54	−6.58
SE	0.79	0.93	1.12	1.40
<i>P</i> value	<.001	.48	.17	<.001
Hospital Anxiety and Depression Scale				
B	5.69	0.41	0.38	−2.36
SE	0.35	0.43	0.50	0.65
<i>P</i> value	<.001	.35	.45	<.001

^aThe covariates are not reported in this table.

Table 5. Multilevel analysis results of the diary and questionnaire variables: model comparisons^a.

Variables	Models				
	0 vs 1	1 vs 2	2 vs 3	0 vs 3	4 vs 5
Diary variables					
Sleep efficiency					
χ^2_1	23.23	1.53	11.54	36.30	18.66
R^2	.07	.01	.03	.11	.04
<i>P</i> value	<.001	.22	<.001	<.001	<.001
Sleep onset latency					
χ^2_1	9.83	0.39	3.57	13.79	12.21
R^2	.02	.00	.01	.03	.01
<i>P</i> value	.002	.53	.06	.003	<.001
Wake after sleep onset					
χ^2_1	14.82	0.62	13.29	28.73	11.99
R^2	.05	.00	.03	.08	.02
<i>P</i> value	<.001	.43	<.001	<.001	<.001
Number of awakenings					
χ^2_1	9.28	1.89	4.63	15.81	3.59
R^2	.00	.01	.01	.02	.00
<i>P</i> value	.002	.17	.03	.001	.06
Time in bed					
χ^2_1	0.69	0.03	3.82	4.54	5.38
R^2	.00	.00	.01	.02	.01
<i>P</i> value	.41	.85	.05	.21	.02
Terminal wakefulness					
χ^2_1	0.41	0.65	0.41	1.46	0.85
R^2	.00	.01	.00	.01	.00
<i>P</i> value	.52	.42	.52	.69	.36
Total sleep time					
χ^2_1	23.45	0.85	2.35	26.66	4.35
R^2	.07	.00	.00	.08	.00
<i>P</i> value	<.001	.36	.13	<.001	.04
Questionnaire variables					
Insomnia Severity Index					
χ^2_1	86.00	4.65	19.07	109.73	66.02
R^2	.25	.02	.05	.31	.30
<i>P</i> value	<.001	.03	<.001	<.001	<.001
Pittsburgh Sleep Quality Index					
χ^2_1	43.05	0.91	20.65	64.61	9.28
R^2	.11	.01	.05	.15	.04

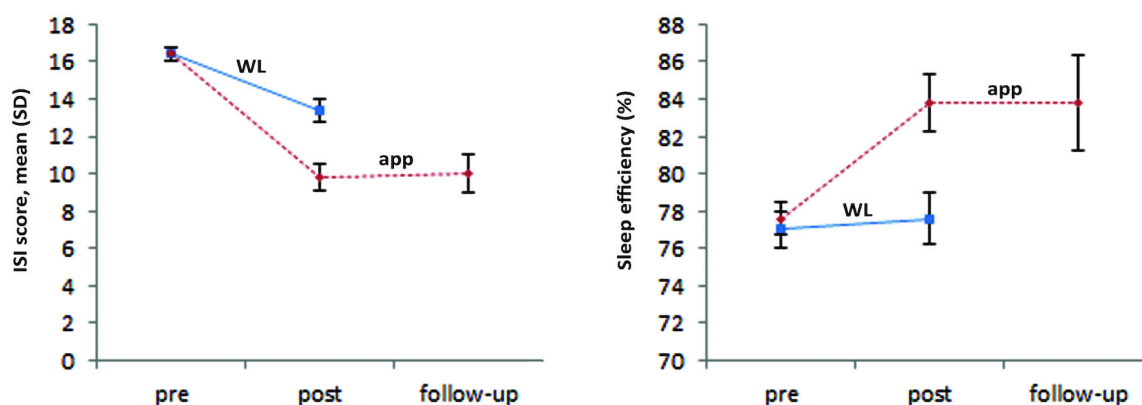
Variables	Models				
	0 vs 1	1 vs 2	2 vs 3	0 vs 3	4 vs 5
<i>P</i> value	<.001	.34	<.001	<.001	.002
Dysfunctional Beliefs and Attitudes about Sleep					
χ^2_1	16.32	0.12	1.86	18.31	11.65
R^2	.02	.00	.00	.02	.05
<i>P</i> value	<.001	.07	.17	<.001	<.001
Centre of Epidemiological Studies Depression scale					
χ^2_1	8.34	0.96	20.67	29.97	18.66
R^2	.02	.01	.05	.08	.08
<i>P</i> value	.004	.33	<.001	<.001	<.001
Hospital Anxiety and Depression Scale					
χ^2_1	3.17	1.49	12.75	17.41	10.03
R^2	.01	.01	.04	.06	.03
<i>P</i> value	.07	.22	<.001	<.001	.002

^aThe covariates are not reported in this table.

Among the participants who completed the pre- and posttest, we found a clinically meaningful change on the ISI ($\Delta\text{ISI} \geq 8$) [40] between the waitlist and the app conditions. We also observed a significant clinically meaningful change 20 times in the app condition (20/45, 44%) and 7 times in the waitlist condition (7/62, 11%) at the posttest. In the app condition, significantly more people reached a meaningful clinical change ($\chi^2_1=15.19$, $P<.001$). Before treatment, all participants had an

ISI score >7 [26]. Of the participants who completed the posttest, 17 in the app condition (17/45, 38%) and 6 in the waitlist condition (6/62, 10%) had an ISI score ≤ 7 . In the app condition, significantly more participants dropped below the insomnia threshold of ISI ≤ 7 than in the waitlist condition ($\chi^2_1=12.20$, $P<.001$). At follow-up, 7 of the 29 participants had an ISI score ≤ 7 (7/29, 24%).

Figure 4. Completers sample: Insomnia Severity Index (ISI) scores and sleep efficiency of the Sleepcare mobile phone app group compared with the waitlist (WL) control group at baseline, posttest, and 3-month follow-up. Error bars represent standard error.



Treatment Adherence

We divided treatment adherence into 4 components objectively measured by the app: (1) the number of sleep diaries filled out, (2) the number of relaxation exercises performed, (3) the number of conversations completed, and (4) the deviation between real time in bed and agreed-upon time in bed (Table 6). Only 2 of the 74 participants did not download the app. Furthermore, the adherence data showed different adherence patterns (see Multimedia Appendix 3). Most participants ($n=35$) filled out >35 diaries, 13 participants filled out <7 diaries, and the other 24 participants filled out between 7 and 35 diaries. This pattern follows a U-shaped curve. For the relaxation exercise, another pattern can be distinguished. More than half of the participants ($n=41$) performed a maximum of 7 relaxation exercises, of whom 11 participants did not do any relaxation exercises at all;

7 participants did >35 relaxation exercises. For the conversations the opposite was true: most participants ($n=47$) finished $\geq 90\%$ of the conversations. Only a few participants ($n=4$) finished <10% of the conversations. A total of 38 participants started and committed to the sleep restriction exercise, meaning that they came to an agreement with the coach about their time in bed. Participants could stay in bed for longer or for less time than the agreement, and both situations occurred. Of these 38 participants, 32 stayed in bed for longer than agreed on for most of the nights involving sleep restriction. When participants stayed in bed too long it was by 67 minutes on average (SD 45); when they shortened their time in bed it was by 42 minutes on average (SD 29). Of the 38 participants, 26 (68%) were adherent, meaning that their time in bed deviated by an average of <60 minutes from the agreed-upon time in bed.

Table 6. Treatment adherence among 72 of 74 participants who downloaded the Sleepcare app.

Adherence component	Participants who performed the component, n	Proportion of the component completed, mean (SD)	Range	Participants with adequate dose, n (%) ^a
Number of diaries filled out	72	29.1 (16.4) times/49 (59%)	0-48 times	35/72 (49%)
Number of relaxation exercises performed	72	10.8 (12.0) times/49 (22%)	0-45 times	7/72 (10%)
Number of completed conversations in the training	72	83% (27%)	0%-100%	47/72 (65%)
Deviation from sleep restriction in minutes	38	59.2 (46.4) minutes	9-285 minutes	26/38 (68%)

^aAdequate dose: diaries >35, relaxation exercises >35, deviation from sleep restriction <60 minutes, conversations >90%.

Discussion

In this large-scale randomized controlled trial, we investigated the efficacy of CBT-I delivered via a mobile phone app. The results show that the app had moderate significant effects than in a waitlist on the primary measures of insomnia severity ($d=-0.66$, 95% CI -0.99 to -0.33) and sleep efficiency ($d=0.71$, 95% CI 0.37 - 1.04). The following secondary measures also improved compared with the waitlist: wake after sleep onset, number of awakenings, PSQI, depression, and anxiety. Other measures, such as time in bed, total sleep time, and dysfunctional beliefs about sleep did not show significant improvements, which could be explained by the exercises included in the treatment. The focus on modified sleep restriction in the app could explain the lack of improvement in time in bed and total sleep time, since sleep restriction aims at increasing sleep efficiency, starting by decreasing time in bed and thereby also influencing total sleep time. In addition, the app did not contain a cognitive exercise, which could explain why we found no difference in the dysfunctional beliefs about sleep. Nevertheless, at posttest, 44% of the participants in the app condition had achieved a clinically meaningful change compared with 11% in the waitlist condition. The improvements were largely sustained at the 3-month follow-up. The observed effects on the primary measures are similar to those reported in a recent meta-analysis on sleep efficiency (Hedge $g=0.58$) and somewhat lower (but in the same range) than those reported for insomnia severity (Hedge $g=1.09$) [16]. Note that the meta-analysis was based on studies with various levels of human involvement, ranging from no human support to personal contact as part of the intervention. Earlier research indicated that human support

increases efficacy [55]. However, the effect sizes in our study were achieved without any form of human support.

Regarding automated support, this study most closely resembles the trials by Espie and colleagues [48] and Ritterband and colleagues [56], which both offered automated Web-based CBT-I. These Web-based treatments packaged the full scope of CBT-I and demonstrated large effects. Espie and colleagues found a Cohen d of 0.95 for sleep efficiency. Ritterband and colleagues found a Cohen d of 1.26 for insomnia severity and 0.68 for sleep efficiency. Again, the observed effect sizes in our study were more or less in the same range as these published results, and our effects were achieved without including the full CBT-I package (eg, we did not include cognitive therapy and stimulus control). The app concentrated on sleep restriction, and as a result the effects for sleep efficiency are more pronounced than those for insomnia severity. The focus on sleep restriction may also explain the absence of an effect on total sleep time.

Zachariae and colleagues [16] found in their meta-analysis that 58.7% to 100% of the participants in the CCBT-I conditions completed postintervention assessments, with an average of 75.3%. In our study, 61% of the participants in the app condition completed postintervention assessment questionnaires, while 81% of the participants in the waitlist condition did so. This difference can probably be explained by the fact that the participants in the waitlist received the app only after they had filled out the postintervention assessment. However, the number of participants filling out assessments may not necessarily correspond to the number of participants who complete interventions. Therefore, we also report treatment adherence

numbers and adequate doses. Previously, Espie and colleagues [48] found that 88% of their participants received an adequate dose (≥ 4 sessions). Lancee and colleagues [55] found that 83% received an adequate dose of the modules in the support condition, and 60% in the no-support condition. In our trial, adherence was measured for the different components, with adherence rates ranging from 10% to 68%. Apart from the relaxation exercise adherence (where only 10% of the participants received an adequate dose), the other adherence rates are comparable with the 60% found by Lancee and colleagues [55] in their no-support condition.

In general, there exists a positive relationship between treatment adherence and treatment outcome [50]. This relationship suggests that this trial could have been more successful if relaxation adherence rates would have been higher, for example. Relaxation exercises have been noted as an effective treatment as part of multicomponent treatment [36]. Therefore, it is likely that higher adherence rates to the relaxation exercise would have improved the overall efficacy of the app. However, relaxation exercises as stand-alone treatment were inferior to other CBT-I exercises, such as stimulus control and sleep restriction [12]. So, as in all multicomponent treatments, it remains unclear how adherence to specific components affected the overall efficacy of the app. In this trial, overall adherence rates were adequate, but there was also a considerable number of people who did not start the modified sleep restriction exercise at all. Beforehand, we decided that it was better to keep people in no or a suboptimal sleep restricting schedule rather than letting them drop out of the treatment altogether. A possible risk of suboptimal treatments could be that people may not seek further help after a less-successful treatment. However, this is a general health care risk, which is not typical for eHealth or self-help interventions. Nevertheless, the optimal tradeoff between individual autonomy and strictness in mobile phone app regimens has yet to be determined in future studies.

Limitations and Future Work

This study has a number of limitations that should be considered in relation to the findings. Since the goal of the study was to demonstrate the efficacy of the app first in a group with insomnia disorder but without too much sleep impairment, we used an ISI score of >7 , meaning that we excluded people who slept <5 hours as measured by a sleep diary. This exclusion criterion may have led to a floor effect and the inclusion of participants with relatively little room for improvement. Although it is hard to compare studies because of different inclusion criteria, it seems that Espie and colleagues [48] only included participants with more severe insomnia (baseline sleep efficiency of 55%–65%). It may be possible to achieve larger effects in samples with higher levels of symptoms. However, it remains the case that the efficacy of our mobile phone app has not yet been demonstrated in a sample with severe insomnia. Because this was one of the first times a stand-alone app has been used to deliver CBT-I, we also excluded participants with comorbidities such as depression. This and the issues mentioned above limit the generalizability of our results, especially given the high comorbidity of depression and insomnia. Now that the app has proven its efficacy in a relatively mildly affected sample,

future research could expand the inclusion criteria (eg, severe insomnia, depression) to study the efficacy of a CBT-I app in a more severely affected population.

A methodological limitation was that no other Web-based or face-to-face treatment group was included. Several other studies have already demonstrated the efficacy of CCBT-I and CBT-I programs. However, a similar Web-based condition could provide insight into the added value of a mobile app. Another related limitation is that there was no placebo control group. It may very well be that nonspecific factors played a role in the treatment effects of the app. Other methodological limitations were that this study used self-report measures, and polysomnography would be needed to confirm the objective changes in sleep. Furthermore, the participants in this study were a self-selected sample and may have been an unusual group of people who were interested in solving their sleep difficulties with self-help. A high percentage of the sample consisted of university-educated participants, who represent only a part of society.

Another limitation was that the app focused on sleep restriction and relaxation. Future work should include more of the other CBT-I components, such as cognitive exercises, and evaluate those. Mobile phone apps provide us with the unique opportunity to study the separate components of CBT-I in a controlled way. Future research could focus on studying the separate components, so that more insight can be gained into the individual efficacy of these CBT-I components. In addition, future research should compare the efficacy of the classical sleep restriction versus the modified sleep restriction exercise.

Lastly, there were some technical issues during the randomized controlled trial that made it impossible for some participants to continue to the next conversation. The occurrence of this problem was monitored and solved when needed. In these cases, a new conversation was manually planned in the database for a specific participant, and an email with instructions to update the app was sent to that participant.

Conclusion

We are confident that this study has produced insights into the domain of automated e-coaching apps for insomnia. These apps provide an opportunity to investigate separate treatment components while minimizing the influence of nonspecific therapist factors such as therapeutic alliance. Keeping the limitations in mind, this study demonstrated the efficacy of a mobile phone app in the treatment of insomnia. These effects were clinically meaningful and in the range of what is found for Web-based treatment in general. This supports the applicability of these kinds of technical tools in the treatment of insomnia. Through these apps, many more people can be offered effective insomnia treatment with probable reduced costs. We are confident that mobile phone apps will prove to be useful in the realm of prevention treatments; it remains to be determined how they should best be offered: in a stand-alone format for (prevention) treatment, or within a blended care framework where the sleep specialist uses an app to improve and accelerate insomnia treatment.

Acknowledgments

This study was funded by the Philips and Technology Foundation STW, the Nationaal Initiatief Hersenen en Cognitie (NIHC) under the Healthy Lifestyle Solutions partnership program. Lisanne Bogaard helped with the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.1 [25].

[[PDF File \(Adobe PDF File\), 978KB - jmir_v19i4e70_app1.pdf](#)]

Multimedia Appendix 2

Pretest questionnaire.

[[PDF File \(Adobe PDF File\), 56KB - jmir_v19i4e70_app2.pdf](#)]

Multimedia Appendix 3

Adherence patterns.

[[PDF File \(Adobe PDF File\), 92KB - jmir_v19i4e70_app3.pdf](#)]

References

1. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med Rev* 2002 Apr;6(2):97-111. [Medline: [12531146](#)]
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th edition. Arlington, VA: APA; 2013.
3. Espie CA, Kyle SD, Hames P, Cyhlarova E, Benzeval M. The daytime impact of DSM-5 insomnia disorder: comparative analysis of insomnia subtypes from the Great British Sleep Survey. *J Clin Psychiatry* 2012 Dec;73(12):e1478-e1484. [doi: [10.4088/JCP.12m07954](#)] [Medline: [23290331](#)]
4. Kyle SD, Morgan K, Espie CA. Insomnia and health-related quality of life. *Sleep Med Rev* 2010 Feb;14(1):69-82. [doi: [10.1016/j.smrv.2009.07.004](#)] [Medline: [19962922](#)]
5. LeBlanc M, Beaulieu-Bonneau S, Mérette C, Savard J, Ivers H, Morin CM. Psychological and health-related quality of life factors associated with insomnia in a population-based sample. *J Psychosom Res* 2007 Aug;63(2):157-166. [doi: [10.1016/j.jpsychores.2007.03.004](#)] [Medline: [17662752](#)]
6. Baglioni C, Battagliese G, Feige B, Spiegelhalder K, Nissen C, Voderholzer U, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord* 2011 Dec;135(1-3):10-19. [doi: [10.1016/j.jad.2011.01.011](#)] [Medline: [21300408](#)]
7. Taylor DJ, Lichstein KL, Durrence HH, Reidel BW, Bush AJ. Epidemiology of insomnia, depression, and anxiety. *Sleep* 2005 Nov;28(11):1457-1464. [Medline: [16335332](#)]
8. Mallon L, Broman J, Hetta J. High incidence of diabetes in men with sleep complaints or short sleep duration: a 12-year follow-up study of a middle-aged population. *Diabetes Care* 2005 Nov;28(11):2762-2767. [Medline: [16249553](#)]
9. Suka M, Yoshida K, Sugimori H. Persistent insomnia is a predictor of hypertension in Japanese male workers. *J Occup Health* 2003 Nov;45(6):344-350 [FREE Full text] [Medline: [14676413](#)]
10. Daley M, Morin CM, LeBlanc M, Grégoire J, Savard J. The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. *Sleep* 2009 Jan;32(1):55-64 [FREE Full text] [Medline: [19189779](#)]
11. Irwin MR, Cole JC, Nicassio PM. Comparative meta-analysis of behavioral interventions for insomnia and their efficacy in middle-aged adults and in older adults 55+ years of age. *Health Psychol* 2006 Jan;25(1):3-14. [doi: [10.1037/0278-6133.25.1.3](#)] [Medline: [16448292](#)]
12. Morin CM, Bootzin RR, Buysse DJ, Edinger JD, Espie CA, Lichstein KL. Psychological and behavioral treatment of insomnia: update of the recent evidence (1998-2004). *Sleep* 2006 Nov;29(11):1398-1414. [Medline: [17162986](#)]
13. Morin CM, Hauri PJ, Espie CA, Spielman AJ, Buysse DJ, Bootzin RR. Nonpharmacologic treatment of chronic insomnia: an American Academy of Sleep Medicine review. *Sleep* 1999 Dec 15;22(8):1134-1156. [Medline: [10617176](#)]

14. Ho FY, Chung K, Yeung W, Ng TH, Kwan K, Yung K, et al. Self-help cognitive-behavioral therapy for insomnia: a meta-analysis of randomized controlled trials. *Sleep Med Rev* 2015 Feb;19:17-28. [doi: [10.1016/j.smrv.2014.06.010](https://doi.org/10.1016/j.smrv.2014.06.010)] [Medline: [25104471](https://pubmed.ncbi.nlm.nih.gov/25104471/)]
15. van Straten A, Cuijpers P. Self-help therapy for insomnia: a meta-analysis. *Sleep Med Rev* 2009 Feb;13(1):61-71. [doi: [10.1016/j.smrv.2008.04.006](https://doi.org/10.1016/j.smrv.2008.04.006)] [Medline: [18952469](https://pubmed.ncbi.nlm.nih.gov/18952469/)]
16. Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia: a systematic review and meta-analysis of randomized controlled trials. *Sleep Med Rev* 2016 Dec;30:1-10. [doi: [10.1016/j.smrv.2015.10.004](https://doi.org/10.1016/j.smrv.2015.10.004)] [Medline: [26615572](https://pubmed.ncbi.nlm.nih.gov/26615572/)]
17. Voinescu B, Szentagotai A, David D. Internet-administered cognitive-behavioral therapy for insomnia. *J Cogn Behav Psychother* 2013;13(1 A):225-237.
18. Konrath S. Positive technology: using mobile phones for psychosocial interventions. In: Yan Z, editor. *Encyclopedia of Mobile Phone Behavior*. Hershey, PA: IGI Global; 2015:871-897.
19. Kaltenthaler E, Cavanagh K. Computerised cognitive behavioural therapy and its uses. *Prog Neurol Psychiatry* 2010;14(3):22-29.
20. Bauer J, Consolvo S, Greenstein B, Schooler J, Wu E, Watson N, et al. ShutEye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. 2012 Presented at: CHI 2012; May 5-10, 2012; Austin, TX, USA.
21. Lawson S, Jamison-Powell S, Garbett A, Linehan C, Kucharczyk E, Verbaan S, et al. Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY: ACM; 2013 Presented at: CHI '13; April 27-May 2, 2013; Paris, France p. 2497-2506. [doi: [10.1145/2470654.2481345](https://doi.org/10.1145/2470654.2481345)]
22. Shirazi AS, Clawson J, Hassanpour Y, Tourian MJ, Schmidt A, Chi EH, et al. Already up? using mobile phones to track & share sleep behavior. *Int J Hum Comput Stud* 2013 Sep;71(9):878-888. [doi: [10.1016/j.ijhcs.2013.03.001](https://doi.org/10.1016/j.ijhcs.2013.03.001)]
23. Miller CB, Espie CA, Epstein DR, Friedman L, Morin CM, Pigeon WR, et al. The evidence base of sleep restriction therapy for treating insomnia disorder. *Sleep Med Rev* 2014 Oct;18(5):415-424. [doi: [10.1016/j.smrv.2014.01.006](https://doi.org/10.1016/j.smrv.2014.01.006)] [Medline: [24629826](https://pubmed.ncbi.nlm.nih.gov/24629826/)]
24. Harvey L, Inglis SJ, Espie CA. Insomniacs' reported use of CBT components and relationship to long-term clinical outcome. *Behav Res Ther* 2002 Jan;40(1):75-83. [Medline: [11762429](https://pubmed.ncbi.nlm.nih.gov/11762429/)]
25. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
26. Morin C. *Insomnia: Psychological Assessment and Management*. New York, NY: Guilford Press; 1993.
27. Spoormaker VI, Verbeek I, van den Bout J, Klip EC. Initial validation of the SLEEP-50 questionnaire. *Behav Sleep Med* 2005;3(4):227-246. [doi: [10.1207/s15402010bsm0304_4](https://doi.org/10.1207/s15402010bsm0304_4)] [Medline: [16190812](https://pubmed.ncbi.nlm.nih.gov/16190812/)]
28. Radloff L. The CES-D scale a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1(3):385-401.
29. Bouma J, Ranchor A, Sanderma R, van Sonderen S. *Het Meten van Symptomen van Depressie met de CES-D: Een Handleiding [Dutch translation of the Epidemiological Studies-Depression scale]*. Groningen, Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken; 1995.
30. Beun R, Griffioen-Both F, Ahn R, Fitrianie S, Lancee J. Modeling interaction in automated e-coaching: a case from insomnia therapy. 2014 Presented at: COGNITIVE 2014, The Sixth International Conference on Advanced Cognitive Technologies and Applications; May 2014; Venice, Italy p. 25-29.
31. Beun RJ, Brinkman WP, Fitrianie S, Griffioen-Both F, Horsch C, Lancee J, et al. Improving adherence in automated e-coaching: a case from insomnia therapy. 2016 Presented at: 11th International Conference on Persuasive Technology; April 5-7, 2016; Salzburg, Austria p. 276-287.
32. Morin C, Espie C. *Insomnia: A Clinical Guide to Assessment and Treatment*. New York, NY: Springer; 2003.
33. Verbeek I, Klip E. *Slapeloosheid [Insomnia]*. Amsterdam, Netherlands: Boom; 2005.
34. Beun R, Fitrianie S, Griffioen-Both F, Spruit S, Horsch C, Lancee J, et al. Talk and Tools: The Best of Both Worlds in Mobile User Interfaces of E-coaching. Technical Report UU-CS-2016-008. Utrecht, Netherlands: Computing and Information Sciences, Utrecht University; 2016.
35. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep* 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: [10.5665/sleep.1642](https://doi.org/10.5665/sleep.1642)] [Medline: [22294820](https://pubmed.ncbi.nlm.nih.gov/22294820/)]
36. Bootzin RR, Epstein DR. Understanding and treating insomnia. *Annu Rev Clin Psychol* 2011;7:435-458. [doi: [10.1146/annurev.clinpsy.3.022806.091516](https://doi.org/10.1146/annurev.clinpsy.3.022806.091516)] [Medline: [17716026](https://pubmed.ncbi.nlm.nih.gov/17716026/)]
37. van Eijk RM. Ambient coaching of progressive relaxation. 2013 Presented at: IEEE/WIC/ACM International Conference on Web Intelligence Intelligent Agent Technology; November 17-20, 2013; Atlanta, GA, USA p. 17-20. [doi: [10.1109/WI-IAT.2013.177](https://doi.org/10.1109/WI-IAT.2013.177)]

38. Kyle SD, Aquino MRJ, Miller CB, Henry AL, Crawford MR, Espie CA, et al. Towards standardisation and improved understanding of sleep restriction therapy for insomnia disorder: a systematic examination of CBT-I trial content. *Sleep Med Rev* 2015 Oct;23:83-88. [doi: [10.1016/j.smrv.2015.02.003](https://doi.org/10.1016/j.smrv.2015.02.003)] [Medline: [25771293](#)]
39. Consolvo S, McDonald D, Landay J. Theory-driven design strategies for technologies that support behavior change in everyday life. 2009 Presented at: CHI 2009: ACM Conference on Human Factors in Computing Systems; April 4-9, 2009; Boston, MA, USA p. 4-9.
40. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep* 2011 May 01;34(5):601-608 [FREE Full text] [Medline: [21532953](#)]
41. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989 May;28(2):193-213. [Medline: [2748771](#)]
42. Morin CM, Vallières A, Ivers H. Dysfunctional beliefs and attitudes about sleep (DBAS): validation of a brief version (DBAS-16). *Sleep* 2007 Nov;30(11):1547-1554 [FREE Full text] [Medline: [18041487](#)]
43. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [Medline: [6880820](#)]
44. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert A. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997 Mar;27(2):363-370. [Medline: [9089829](#)]
45. Guay F, Vallerand R, Blanchard C. On the assessment of situational intrinsic and extrinsic motivation: the Situational Motivation Scale (SIMS). *Motivation Emotion* 2000;24(3):175-213.
46. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. *MIS Q* 2003;27(3):425-478.
47. Dallal GE. Welcome to Randomization.com!!! (where it's never the same thing twice). 2013 Mar 29. URL: <http://randomization.com/> [accessed 2017-01-26] [WebCite Cache ID 6noLOByZx]
48. Espie CA, Kyle SD, Williams C, Ong JC, Douglas NJ, Hames P, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. *Sleep* 2012 Jun;35(6):769-781 [FREE Full text] [doi: [10.5665/sleep.1872](https://doi.org/10.5665/sleep.1872)] [Medline: [22654196](#)]
49. Lancee J, van Straten A, Morina N, Kaldo V, Kamphuis JH. Guided online or face-to-face cognitive behavioral treatment for insomnia? A randomized wait-list controlled trial. *Sleep* 2016 Jan 01;39(1):183-191 [FREE Full text] [doi: [10.5665/sleep.5344](https://doi.org/10.5665/sleep.5344)] [Medline: [26414893](#)]
50. Horsch C, Lancee J, Beun RJ, Neerincx MA, Brinkman W. Adherence to technology-mediated insomnia treatment: a meta-analysis, interviews with users, and focus groups with users and experts. *J Med Internet Res* 2015;17(9):e214 [FREE Full text] [doi: [10.2196/jmir.4115](https://doi.org/10.2196/jmir.4115)] [Medline: [26341671](#)]
51. Hox J. Multilevel Analysis, Techniques and Applications. 2nd edition. New York, NY: Routledge; 2010.
52. Snijders TA, Bosker RJ. Modeled variance in two-level models. *Sociol Methods Res* 1994 Feb;22(3):342-363. [doi: [10.1177/0049124194022003004](https://doi.org/10.1177/0049124194022003004)]
53. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: L Erlbaum; 1988.
54. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393 [FREE Full text] [Medline: [19564179](#)]
55. Lancee J, van den Bout J, Sorbi MJ, van Straten A. Motivational support provided via email improves the effectiveness of internet-delivered self-help treatment for insomnia: a randomized trial. *Behav Res Ther* 2013 Dec;51(12):797-805. [doi: [10.1016/j.brat.2013.09.004](https://doi.org/10.1016/j.brat.2013.09.004)] [Medline: [24121097](#)]
56. Ritterband LM, Thorndike FP, Gonder-Frederick LA, Magee JC, Bailey ET, Saylor DK, et al. Efficacy of an Internet-based behavioral intervention for adults with insomnia. *Arch Gen Psychiatry* 2009 Jul;66(7):692-698 [FREE Full text] [doi: [10.1001/archgenpsychiatry.2009.66](https://doi.org/10.1001/archgenpsychiatry.2009.66)] [Medline: [19581560](#)]

Abbreviations

CBT-I: cognitive behavioral therapy for insomnia
CCBT-I: computerized cognitive behavioral therapy for insomnia
CES-D: Centre for Epidemiological Studies Depression
DBAS-16: brief Dysfunctional Beliefs and Attitudes about Sleep
DSM-5: Diagnostic and Statistical Manual of Mental Disorders
HADS: Hospital Anxiety and Depression Scale
ISI: Insomnia Severity Index
PSQI: Pittsburgh Sleep Quality Index

Edited by G Eysenbach; submitted 22.08.16; peer-reviewed by F Ho, N Siriwardena; comments to author 06.10.16; revised version received 11.11.16; accepted 23.11.16; published 11.04.17.

Please cite as:

Horsch CHG, Lancee J, Griffioen-Both F, Spruit S, Fitrianie S, Neerincx MA, Beun RJ, Brinkman WP

Mobile Phone-Delivered Cognitive Behavioral Therapy for Insomnia: A Randomized Waitlist Controlled Trial

J Med Internet Res 2017;19(4):e70

URL: <http://www.jmir.org/2017/4/e70/>

doi: [10.2196/jmir.6524](https://doi.org/10.2196/jmir.6524)

PMID: [28400355](https://pubmed.ncbi.nlm.nih.gov/28400355/)

©Corine HG Horsch, Jaap Lancee, Fiemke Griffioen-Both, Sandor Spruit, Siska Fitrianie, Mark A Neerincx, Robbert Jan Beun, Willem-Paul Brinkman. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 11.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Using Persuasive Technology to Increase Physical Activity in People With Chronic Obstructive Pulmonary Disease by Encouraging Regular Walking: A Mixed-Methods Study Exploring Opinions and Preferences

Yvonne Kiera Bartlett¹, BSc, MSc, PhD; Thomas L Webb², BSc, MSc, PhD; Mark S Hawley³, BSc, PhD

¹Manchester Centre for Health Psychology, Manchester Academic Health Science Centre, The School of Health Sciences, University of Manchester, Manchester, United Kingdom

²Department of Psychology, University of Sheffield, Sheffield, United Kingdom

³Centre for Assistive Technology and Connected Healthcare and School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom

Corresponding Author:

Yvonne Kiera Bartlett, BSc, MSc, PhD

Manchester Centre for Health Psychology

Manchester Academic Health Science Centre, The School of Health Sciences

University of Manchester

Coupland 1 Building

Oxford Road

Manchester, M139PL

United Kingdom

Phone: 44 161306436 ext 65436

Fax: 44 161 275 2588

Email: kiera.bartlett@manchester.ac.uk

Abstract

Background: People with chronic obstructive pulmonary disease (PwCOPD) often experience breathlessness and fatigue, making physical activity challenging. Although many persuasive technologies (such as mobile phone apps) have been designed to support physical activity among members of the general population, current technologies aimed at PwCOPD are underdeveloped and only use a limited range of persuasive technology design principles.

Objective: The aim of this study was to explore how acceptable different persuasive technology design principles were considered to be in supporting and encouraging physical activity among PwCOPD.

Methods: Three prototypes for mobile apps using different persuasive technology design principles as defined by the persuasive systems design (PSD) model—namely, dialogue support, primary task support, and social support—were developed. Opinions of these prototypes were explored through 28 interviews with PwCOPD, carers, and the health care professionals (HCPs) involved in their care and questionnaires completed by 87 PwCOPD. Participants also ranked how likely individual techniques (eg, competition) would be to convince them to use a technology designed to support physical activity. Data were analyzed using framework analysis, Friedman tests, and Wilcoxon signed rank tests and a convergent mixed methods design was used to integrate findings.

Results: The prototypes for mobile apps were received positively by participants. The prototype that used a dialogue support approach was identified as the most likely to be used or recommended by those interviewed, and was perceived as more persuasive than both of the other prototypes ($Z=-3.06$, $P=.002$; $Z=-5.50$, $P<.001$) by those who completed the questionnaire. PwCOPD identified dialogue support and primary task support techniques as more likely to convince them to use a technology than social support techniques ($Z=-5.00$, $P<.001$; $Z=-4.92$, $P<.001$, respectively). Opinions of social support techniques such as competition and collaboration were divided.

Conclusions: Dialogue support and primary task support approaches are considered to be both acceptable and likely to be persuasive by PwCOPD, carers, and HCPs. In the future, these approaches should be considered when designing apps to encourage physical activity by PwCOPD.

KEYWORDS

persuasive technology; chronic obstructive pulmonary disease; physical activity; walking; mHealth; mobile apps

Introduction

The term “persuasive technology” describes “any interactive computing system designed to change people’s attitudes and behaviors” (p.1; [1]). There are estimated to be over 40,000 mobile phone apps that aim to persuade users to change health behaviors such as physical activity, diet, and smoking [2]. However, despite a rising number of publications in the area [3], content analysis of existing apps reveal that they currently make little use of theories on behavior change or persuasive technology and include little evidence-based content [4-8]. As a consequence, there have been multiple calls for people to make better use of theory and evidence when designing apps intended to promote behavior change [9-11].

The persuasive systems design (PSD) model was developed to provide a framework for the design and evaluation of persuasive technologies [12]. Expanding on Fogg’s (2003) conceptualization of persuasive technology [1], Oinas-Kukkonen

and Harjuma (2009) identified 28 techniques that can be used to design persuasive systems. These are organized into four persuasive design principles, or approaches, to persuasion: (1) techniques that help the user to carry out the target behavior (termed *primary task support*), (2) techniques that motivate the user through feedback and interaction with the technology (termed *dialogue support*), (3) techniques that leverage social influence (termed *social support*), and finally, (4) techniques that make the system appear credible to the user (termed *credibility support*) [12]. Table 1 provides examples of techniques associated with each of these design principles. Initial evidence suggests that including dialogue support techniques in technology-based interventions may increase adherence [13], the extent to which the intervention is perceived to be persuasive and, through increasing people’s motivation to use the intervention, effect actual use of the persuasive system [14]. Despite this evidence, however, theoretical approaches to designing the interactive elements of persuasive technologies remain underused [13,15].

Table 1. Examples of persuasive technology techniques by design principle [12].

Design principle	Example persuasive technology technique	Description
Primary task support	Reduction	A system that reduces complex behavior into simple tasks helps users perform the target behavior, and may increase the cost-benefit ratio of a behavior.
Dialogue support	Social role	If a system adopts a social role, users will be more likely to use it for persuasive purposes.
Social support	Competition	A system can motivate users to adopt a target attitude or behavior by leveraging human beings’ natural drive to compete.
Credibility support	Trustworthiness	A system that is viewed as trustworthy will have increased powers of persuasion

Increasing Physical Activity in PwCOPD

Chronic obstructive pulmonary disease (COPD) is an umbrella term for a number of lung diseases such as chronic emphysema and chronic bronchitis. In 2013, COPD was the third most common cause of death in the United States [16], and the World Health Organization estimates that by 2030 it may become the third most common cause of death worldwide [17]. On average, each year, treating COPD typically costs the US health care system US \$30 billion and the UK National Health Service (NHS) around £800 million [16-18]. PwCOPD experience symptoms such as breathlessness on exertion, muscle weakness, increased sputum production, and a chronic cough. In the short-term, these symptoms can reduce people’s ability to complete daily activities and reduce quality of life [19] and, in the long-term, these symptoms can lead to hospitalization and respiratory failure [20]. PwCOPD can enter a negative cycle as their symptoms make it harder to remain active, and the less active they are, the worse their symptoms become [21,22]. Indeed, evidence suggests that physical inactivity is associated with higher numbers of hospital admissions, exacerbations, and mortality in PwCOPD [23-25]. Currently, it is recommended that stable COPD is managed with a combination of medications

and pulmonary rehabilitation (PR) [26,27]. PR has been shown to increase people’s capacity for exercise and their health-related quality of life [28]. However, studies have suggested that completing a course of PR does not necessarily increase levels of physical activity [29,30]. This finding suggests that although structured rehabilitation can increase the ability to perform physical activity, additional support may be needed to integrate physical activity into everyday life. Walking is a low-intensity, free physical activity that does not require specialist equipment or tuition. Regular walking after PR has been associated with higher quality of life and health-related quality of life [31], and increases in distance walked daily and daily step count have been found to predict fewer acute exacerbations in people with COPD [32,33]. Encouraging daily walking could therefore be an effective way to help increase physical activity in everyday life.

Using Technology to Support Physical Activity in PwCOPD

Previous research on how technology can be used to support physical activity in PwCOPD has explored how both Internet and mobile phone-based technologies can encourage and support physical activity among PwCOPD, either alone or in conjunction

with counseling [34-36]. The findings indicate that technology-based interventions are usually acceptable to PwCOPD and, although the studies thus far are mostly feasibility and pilot studies, early findings suggest that technology-based interventions have the potential to increase levels of physical activity in this population [36,37]. Indeed, to date, randomized trials have shown improvements in physical activity up to 3 months [38] and daily step count up to 4 months [39]. However, it should be noted that, in the latter study, engagement with the intervention had decreased by 12 months, and daily step count was not significantly different from baseline at this point [39].

An additional problem (which may contribute to the limited efficacy) is that, to date, interventions have tended to rely on a relatively small pool of techniques for promoting changes in behavior, namely, self-monitoring, providing feedback, motivational suggestions, and goal-setting [34-39]. As a result, it is not currently known how acceptable a wider range of persuasive technology techniques would be and what design principles would be most attractive and persuasive to PwCOPD, their carers, and the health care professionals (HCPs) involved in their care.

Aims

This research therefore sought to explore how acceptable and persuasive technologies following different persuasive design principles are in increasing physical activity through encouraging daily walking among people with chronic obstructive pulmonary disease (PwCOPD). Our aim was to inform the choice of design principles and specific persuasive techniques in the design of an app that could be used to encourage physical activity in this population. To achieve this, we designed three prototype apps and used them to investigate the following research questions:

RQ1: What are the opinions and preferences of PwCOPD, their carers, and HCPs involved in their care toward systems using different persuasive technology approaches?

RQ2: Which individual techniques, or design principles, are perceived to be most persuasive?

Methods

Design

A convergent mixed methods design was used to assess the opinions and preferences of PwCOPD, carers, and HCPs toward persuasive technology [40]. Ethical approval for the study was granted by the ethics committee at the University of Sheffield and permission to recruit was granted by both the NHS (for HCPs) and the British Lung Foundation.

Participant Recruitment

Participants for the interviews were recruited through four British Lung Foundation Breathe Easy support groups in South Yorkshire (PwCOPD and carers) and from an NHS service specializing in care for PwCOPD (HCPs). Following approval from the moderators of the groups and a manager at the NHS service, potential participants were given information about the project. Anyone who was interested then contacted the

researchers to participate. All participants provided informed consent.

A second sample of PwCOPD was invited to complete a questionnaire, either through a website or by post. A letter was sent to 140 Breathe Easy support groups (excluding those in South Yorkshire), which contained a link for Web-based completion and a number to call if participants preferred to receive a paper copy of the questionnaire. In addition, 34 moderators of online support groups for PwCOPD were contacted and 6 agreed to post a link to the questionnaire on their websites. All participants provided either written or electronic consent and no incentives were provided for participation.

Materials

Prototypes

Three “medium-fidelity” prototypes were created to show how the screens within each system might look and to describe how users might navigate through the system. Medium-fidelity prototypes present the layout and content of each screen as it would look; not a sketch, as would be expected for a low-fidelity prototype; however, also they are not interactive, as would be expected for a high-fidelity prototype [41]. Navigation through the screens was shown with an arrow indicating which button should be pressed to move to the next screen. The prototypes were presented on a screen or the screenshots were presented on paper. Each prototype focused on a different persuasive technology design principle as delineated by the PSD model: dialogue support, primary task support, and social support. It was decided not to develop a prototype describing credibility support as the research was being conducted through a University and this may have inferred some credibility. It would therefore be difficult to know whether to attribute credibility to the persuasive technology technique or techniques or the context of the research. Each prototype used the same behavior change techniques associated with control theory [42]; namely, prompt intention formation, prompt specific goal setting, prompt self-monitoring of behavior, receive feedback, and prompt review of behavior [42,43]. All prototypes were designed to monitor a daily walk, in addition, certain persuasive technology techniques were present in all three prototypes such as self-monitoring, tailoring, and reduction [12]. The prototypes were further informed by looking at the most popular physical activity apps available for Android, iPhone, and Windows phone at the time of development to see how techniques were operationalized in popular apps designed to promote physical activity. The look and feel of the prototypes was standardized as far as possible, with each using the same font, font sizes, button design, and color scheme.

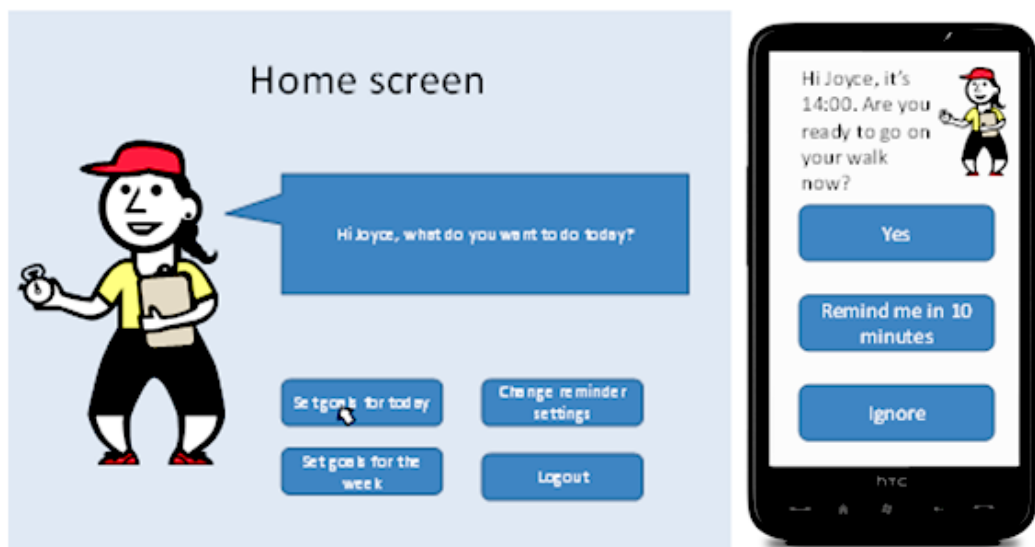
Prototype 1: Virtual Coach System

This prototype used a dialogue support approach that was designed to encourage interaction between the user and the system. In this prototype, the virtual coach represented by a static picture, used the name of the user (“Joyce”) to personalize the system and encourage interaction by taking a social role. The coach led the user through progressive goals. Although there was the option to change the goals, suggestions were made

by the coach. The user could then choose to receive reminders to complete the activity (see Figure 1). The prototype explained that, while the user is walking, they would have the option to receive audio encouragement from the coach (in the form of recorded messages telling the user how many minutes they have

been walking, or when they are halfway to their goal). Feedback would be presented as a graph, accompanied by praise and encouragement from the virtual coach. The prototype also outlined a suggested exercise plan with daily walking goals that increased to reach an overall goal (walking for 30 minutes).

Figure 1. Screenshots from the virtual coach prototype showing the homepage and a reminder.

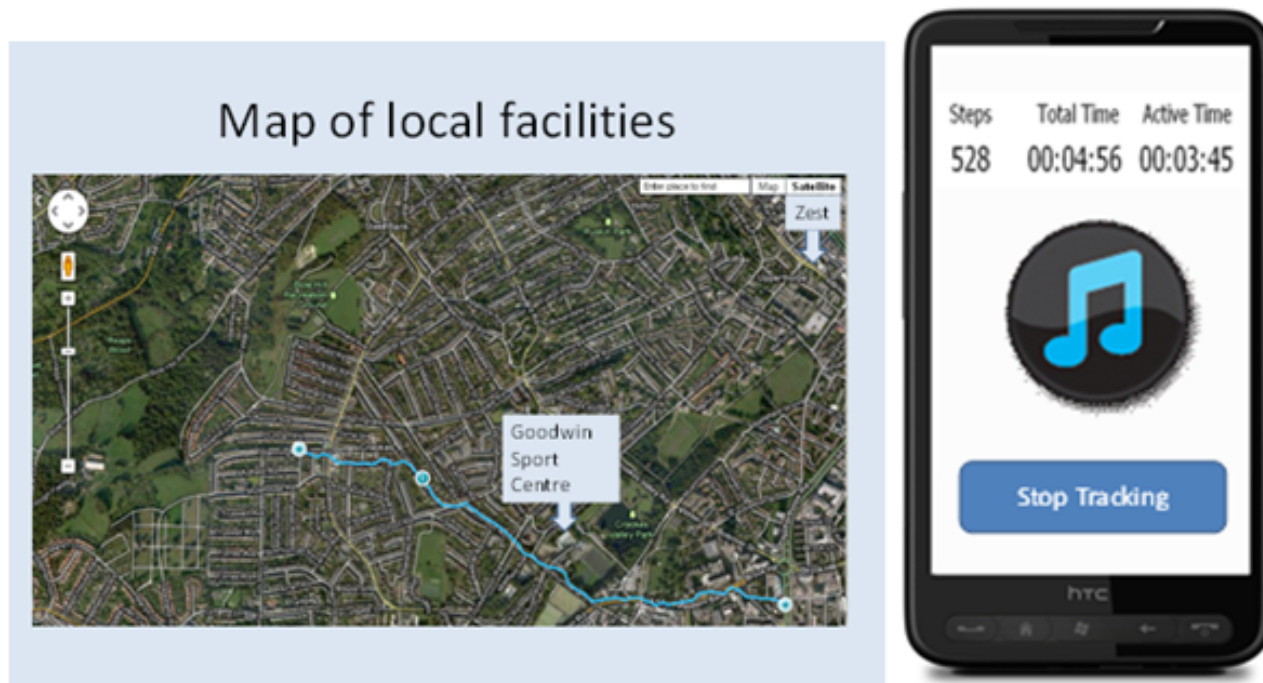


Prototype 2: Music and Maps System

This prototype used a primary task support approach that was intended to help the user to achieve their daily walking goal. The prototype was based on the format used by many of the existing apps designed to promote physical activity. In this prototype, participants could see that the user could set goals

and track their activity using their mobile phone. It was explained that, while walking, the user could choose music to listen to. Following the walk, feedback would be offered on a satellite map, as a summary table, or on a calendar (with activity levels shown for each day). It was further explained that local exercise facilities would be highlighted on the map (see Figure 2).

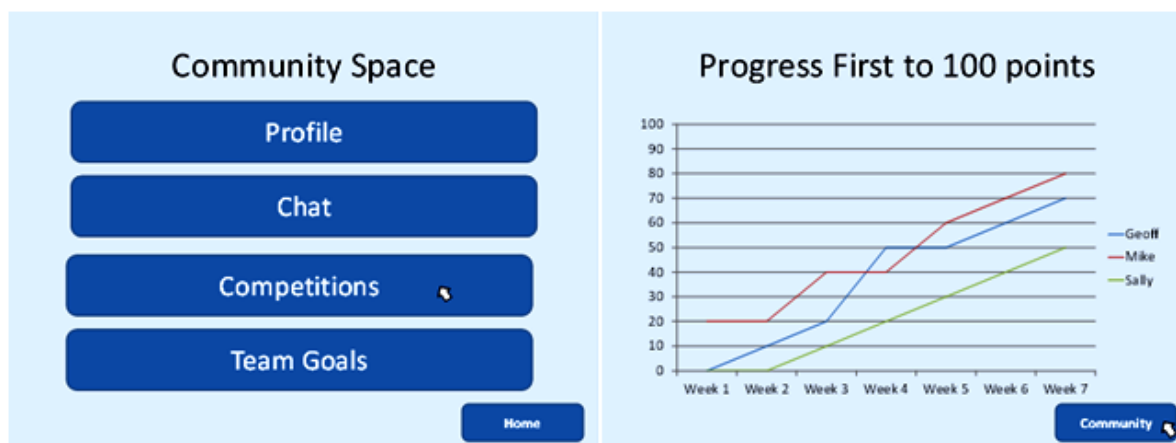
Figure 2. Screenshots from the music and maps prototype showing the map feedback and playing music.



Prototype 3: Online Community System

This prototype used a social support approach and was based on the idea of building a community of like-minded users to support increases in physical activity (see Figure 3). The prototype described how the app would provide a mechanism for computer-mediated communication between peers, while encouraging interaction through competitions and collaborations.

Figure 3. Screenshots from the online community prototype showing the community space and competition graph.



Interview Guide and Questionnaire

Both the questionnaire and the interview schedule followed the same basic structure. PwCOPD were first asked background questions related to COPD and their current level of physical activity. All participants were asked about their use of technology and whether they had any previous experience with persuasive technology. This was followed by a presentation of each of the three prototypes with an explanation to describe how the user might navigate through the system, following which participants' opinions were sought before the next prototype was presented. Once participants had expressed their opinions on all three prototypes, their overall opinion of using persuasive technology to support and encourage physical activity within PwCOPD, and preferences for specific prototypes as well as techniques and features within the prototypes, were sought.

The interview was piloted with a person with COPD. As this participant reported no problems in understanding the interview material, and the timing was appropriate, it was decided to include the data from this participant within the main analysis. The Web-based questionnaire was piloted to ensure that all of the branch questions were working effectively, and that the prototypes were displayed appropriately. A paper version of the questionnaire was piloted with 4 people aged 31-60 years in order to test whether the branch questions were clear, and to establish how long it typically took people to complete the questionnaire. Time taken ranged from 15 to 30 minutes. No problems were reported with the branch questions. However, we did reverse the items identified as negative items in the measures section (so a high score would indicate a more positive response) and correct an error in the information section. As

The prototype showed that users could track their activity using a mobile phone, and share information with other users. It was explained that points would be given when users achieved their goals (the details of the goal completed would not be shared) and that there was the potential for users to earn both virtual (stars or trophies on their profile) and "real-world" rewards (either through vouchers or donating money to charity).

those who piloted the questionnaire did not have COPD, their data were not used in the analysis.

Questionnaire Measures

The questionnaire was divided into four sections:

Section 1 "Questions about you" included questions about demographics (eg, age and gender), how long the participant had been diagnosed with COPD, and the Medical Research Council (MRC) breathlessness scale [44,45] which comprises 5 statements that ask participants to grade their current experience of breathlessness from "Not troubled by breathlessness except on strenuous exercise" (MRC grade 1) to "Too breathless to leave the house or breathless when dressing or undressing" (MRC grade 5).

Section 2, "Questions about physical activity" asked participants to estimate the number of minutes per week that they engaged in light, moderate, and vigorous intensity activity (definitions were provided of each of these, based on the information in [46]).

Section 3, "Questions about technology" was developed for this study and included questions related to computer and mobile phone ownership and use. Participants were also asked if they had "heard of or seen," "ever used," "still use," or "would consider using" "any technology (ie, on the computer, on the Internet, or mobile phone) that is designed to try and help change people's behaviors, for example, increasing exercise, encouraging healthy eating, or stopping smoking."

Finally, Section 4 assessed participants' opinion of each prototype using 8 items, of which 4 items were translated from the *perceived persuasiveness* measure [14]; that is, participants were asked how much they agreed that each prototype "was

interesting,” “would have an influence on me,” “is personally relevant to me,” and “makes me think about my physical activity.” An additional 4 items were added to assess the extent to which participants expected to enjoy using the technology (“I would not enjoy using this system” [negative item]) and how effective they thought it might be (“I think the system would be useful in increasing my physical activity,” “If I wanted to increase my activity levels, I would not use this system” [negative item], and “this system makes me want to increase my physical activity”). All the items were answered on a 7-point Likert scale from “strongly agree” to “strongly disagree,” and the negative items identified above were reverse-scored, and thus a higher score indicates a more positive response. Reliability across the 8 items was assessed using Cronbach alpha and was found to be high for each prototype (virtual coach, Cronbach alpha=.93; music and maps, Cronbach alpha=.93; and online community, Cronbach alpha=.95) [47]. Therefore, the 8 items were summed to create a single scale score representing how persuasive each prototype was deemed to be. Participants were also given a full list of the features across the prototypes and asked to rank the top five that they believed might convince them to use the technology and also to indicate with an “x” any features that would definitely not convince them to use the technology.

In the interview, preferences were elicited by asking participants to identify which prototype they would use (or recommend to others in the case of HCPs). To allow comparison with the questionnaire data, these responses were coded [48] such that if a clear choice was made, then the prototype or feedback screen was given a score of 1; whereas, if a participant reported that they would choose a combination of two prototypes or that two were equally favored, each was given a score of 0.5.

Data Analysis

Framework analysis [49] was used to analyze the interview data in Nvivo 9 (QSR International Pty Ltd). The questionnaire data were analyzed with SPSS 17.0 (SPSS Inc). Preferences for features were calculated using the ranked score given to the

individual techniques (recoded such that a ranking of 1 gave the highest score). Individual techniques were grouped according to the PSD (primary task support, dialogue support, or social support) [12] and the Friedman and Wilcoxon signed rank test were used to identify differences in ranking between the principles.

The data from the interviews and questionnaires were analyzed separately, and then an integration matrix was designed to compare the two strands of data [40,48,50,51]. The integrated findings are presented under thematic headings in Results section below.

Results

Sample Characteristics

Table 2 describes the characteristics of the sample. In total, 23 interviews were conducted; 11 with PwCOPD on their own, 5 with PwCOPD and their carers, and 7 with HCPs (providing a total N of 28). Questionnaires were returned by 121 PwCOPD; however, 34 were excluded due to missing data. The analyses reported below are therefore based on those who rated how persuasive PwCOPD found each prototype (n=87). Mild COPD was underrepresented in both samples: The modal MRC breathlessness grade reported by PwCOPD who completed the questionnaire was 4, and 69% (11/16) of PwCOPD who were interviewed reported needing help when walking outside.

Most of the participants who were interviewed had a mobile phone (although some rarely used it). However, very few of these participants had ever heard of or used any form of persuasive technology. The participants who responded to the questionnaire seemed to be more familiar with technology, with 91% (79/87) of the participants having a mobile phone, and 52% (41/79) using it at least daily. Furthermore, 46% (40/87) of the participants had heard of persuasive technology, and 63% (46/73) reported that they would consider using persuasive technology.

Table 2. Participant characteristics.

Characteristics	Interview (n=28)	Questionnaire (n=87)
Mean age (SD ^a)	70.8 (8.3) ^b	64.0 (8.5)
% of female participants	16 (57%)	59 (69%) ^c
Nationality UK	28 (100%)	58 (67%)
MRC^dbreathlessness grade		
1		5 (6%)
2		21 (24%)
3		19 (22%)
4		28 (32%)
5		14 (16%)
≥150 min moderate activity per week ^e		39 (65%)

^aSD: standard deviation.

^bOnly PwCOPD, n=16.

^cn=85.

^dMRC: Medical Research Council.

^en=60.

Participants' Opinions of the Prototypes and Preferences

Prototype 1: The Virtual Coach System

Participants who were interviewed tended to think that encouragement from the virtual coach would be motivating, which could indicate that the prototype was successful in describing a system that could fulfill a social role:

You are motivated when you're encouraged.
[PwCOPD, female, aged 65 years]

The virtual coach system was thought to be good for people who were more mobile and those living alone. Carers described similar sentiments, namely, that the virtual coach system would be good for people who are mobile, on their own, and able to use the technology.

If they can manage the technology... And if it was somebody on their own who needed (it). [Carer, female, aged 75]

Participants' reasons for deeming the virtual coach system as suitable for someone on their own were that if a carer was motivating the person with COPD, then there would be no need to duplicate this role through technology, again suggesting that participants viewed this form of technology as fulfilling a social role. HCPs were very positive about the prototype, although they felt that the novelty of a virtual coach may wear off and that not everyone would understand the technology. They

therefore felt that there should be an opportunity to simplify the virtual coach if the full system was deemed too complex. HCPs also tended to think that the technology suggesting goals would be useful.

I think that's what a lot of people need because if you're just doing it yourself, you just switch off and say oh, another day. [HCP, female]

When asked which system they would use (or recommend for use) the virtual coach system was given the highest score by HCPs, but it was not scored as highly as the other prototypes by PwCOPD or their carers (see Figure 4).

Participants who completed the questionnaire perceived the virtual coach prototype to be the most persuasive (mean 40.71, SD 11.46). A Friedman test identified a significant main effect of prototype ($\chi^2_2=28.1$, $P<.001$), and a Wilcoxon signed rank test (with Bonferroni correction applied) found significant differences between the virtual coach and the music and maps prototypes ($Z=-3.06$, $P=.002$) and between the virtual coach and the online community prototypes ($Z=-5.50$, $P<.001$; see Figure 5).

In summary, both samples agreed that the virtual coach system could be persuasive, although the PwCOPD who were interviewed were less positive about the prototype than PwCOPD who answered the questionnaire and the HCPs and carers who were interviewed.

Figure 4. Preference for prototypes expressed during the interviews. HCP: health care professional; PwCOPD: people with chronic obstructive pulmonary disease.

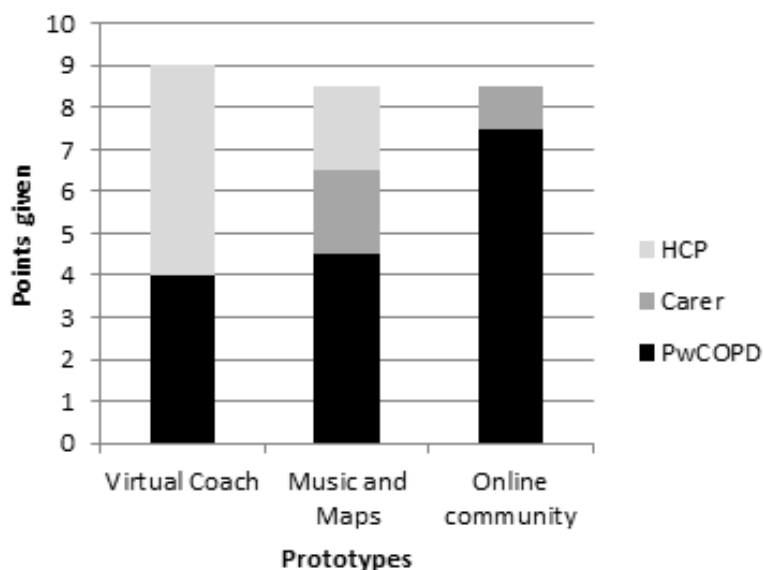
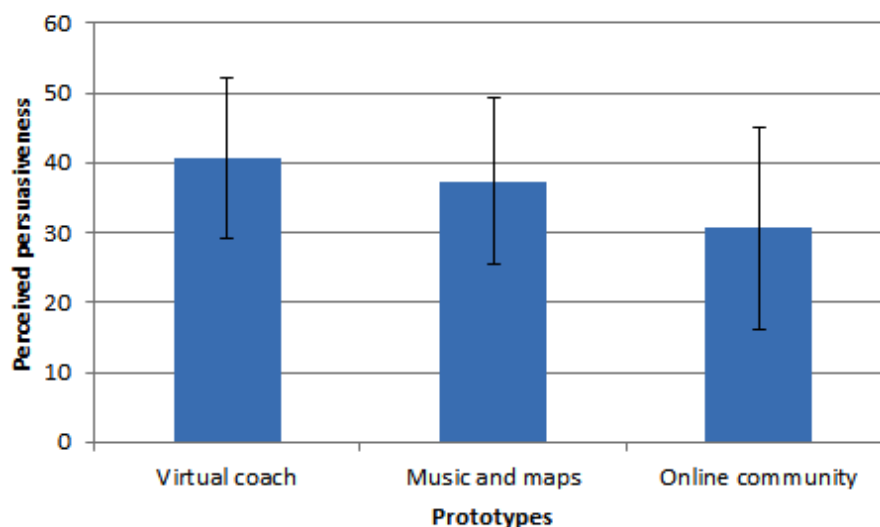


Figure 5. Perceived persuasiveness of each prototype (questionnaire data).



Prototype 2: The Music and Maps System

When interviewed, both PwCOPD and HCPs mentioned that the maps and the summary information provided in the music and maps prototype seemed like it might be interesting, but that it may not be sufficient to promote physical activity. Participants felt that goals that were suggested by the system may be a useful addition to help to persuade people to perform the activity suggested.

I do like the maps and where you've walked, and tracked and I think that's good, because look I've

done that, I might do a bit more tomorrow. [PwCOPD female, aged 70 years]

I would, ideally, like a combination of both, in that if you want the system to suggest goals then it can do but if you're quite happy with setting your own goals and you know what you want to achieve then it sort of takes a more stand back approach and lets you do it basically. [HCP, female]

Again, the technology described in the music and maps prototype was thought to be better for people who were relatively mobile as the feedback would likely be more interesting if the person was walking further. Some PwCOPD

mentioned that this system would not provide them with any useful information if they either walked the same routes or walked very little outside.

The idea of walking with music divided opinion, with some participants feeling that music was useful as a distraction while exercising, whereas others thought that it would be dangerous to walk while using headphones:

There are an awful lot of accidents caused by...walking along, their head's in the clouds their big bopping through their ears...I personally think it's not a good idea. [PwCOPD, male, aged 73 years]

Some PwCOPD, HCPs, and carers indicated that they may use (or recommend that those they care for use) the music and maps prototype (see Figure 4). The PwCOPD who completed the questionnaire rated this prototype as less persuasive (mean 37.40, SD 11.85) than the virtual coach prototype ($Z=-3.06$, $P=.002$), but more persuasive than the online community prototype ($Z=-3.82$, $P<.001$) (see Figure 5).

Prototype 3: The Online Community System

Participants tended to think that an appreciation of the social components of this prototype would depend on the user's personality; that is, it may appeal to some but not to others. One user had previously had a negative experience with an online support group and therefore said that they would not use one again; another remarked that the success of online communities depended on who else was using the website:

I'm a person person rather than a computer person. So for me, in my age group I have doubts, the younger end...they're into Facebook, they're into Twitter and whatever. I'm not. [PwCOPD, male, aged 74 years]

Participants who liked the prototype describing the online community appreciated the potential for competition and for communicating with people who were going through similar experiences. Some PwCOPD felt that competition would motivate them and likened it to other competitive activities that they enjoyed like playing cards or quizzes:

It encourages you to do it both for your own sake and for the competition. [PwCOPD, male, aged 76 years]

Some participants felt that an online community would be better for those who are more mobile; whereas, other participants thought that people who could not do much activity would be more likely to use such technology. HCPs tended to support the latter view, stating that the online community would be the best for people who cannot go out, but that the approach would only work if the user themselves chose it. One HCP said that it would be hard for her to suggest this system to PwCOPD as she did not like it herself.

Others felt that incorporating competition might promote an unhealthy desire to win and, relatedly, that losing may have a detrimental effect on the user's feelings; or that being in competition was not in keeping with the purpose of this technology (which is to promote the self-management of COPD, and ultimately to feel better):

Is making it competitive taking the idea away from what you're actually doing it for? [PwCOPD, male, aged 68 years]

Some participants felt that their opinion of the prototype describing the online community system may be influenced by their own competitive nature; some described that being competitive would make losing harder, whereas others felt that being a competitive person would encourage them to try more. HCPs made the point that they try to discourage competition between people during PR, as it can result in people over-exerting themselves or feeling disheartened. When the idea that the actual goal (ie, how many minutes) would not be revealed to other users was reiterated, some HCPs changed their minds and became more supportive of the idea, whereas others felt that having hidden goals might encourage cheating and that the points would not necessarily go to the right people. One HCP felt that using persuasive technology to make any comparison between users would be inappropriate:

We try and avoid encouraging that sort of behavior, erm, and I'm quite a competitive person and you know the whole first person to get to 500 points and I would be, and I know this sounds really bad but I would be really inclined to decrease the amount of activity that I did to get my points quicker to beat someone. [HCP, female]

When participants who were interviewed were asked which persuasive technology they would be most likely to use, the prototype describing the online community was chosen more often than the other two prototypes by PwCOPD. It was chosen by a few carers but not by any HCPs (see Figure 4). Respondents to the questionnaire rated the online community prototype as significantly less persuasive than the prototypes describing the music and maps and the virtual coach (mean 30.67, SD 14.52; $Z=-5.50$, $P<.001$ and $Z=-3.82$, $P<.001$, respectively; see Figure 5).

Opinions of Individual Features and Persuasive Technology Techniques

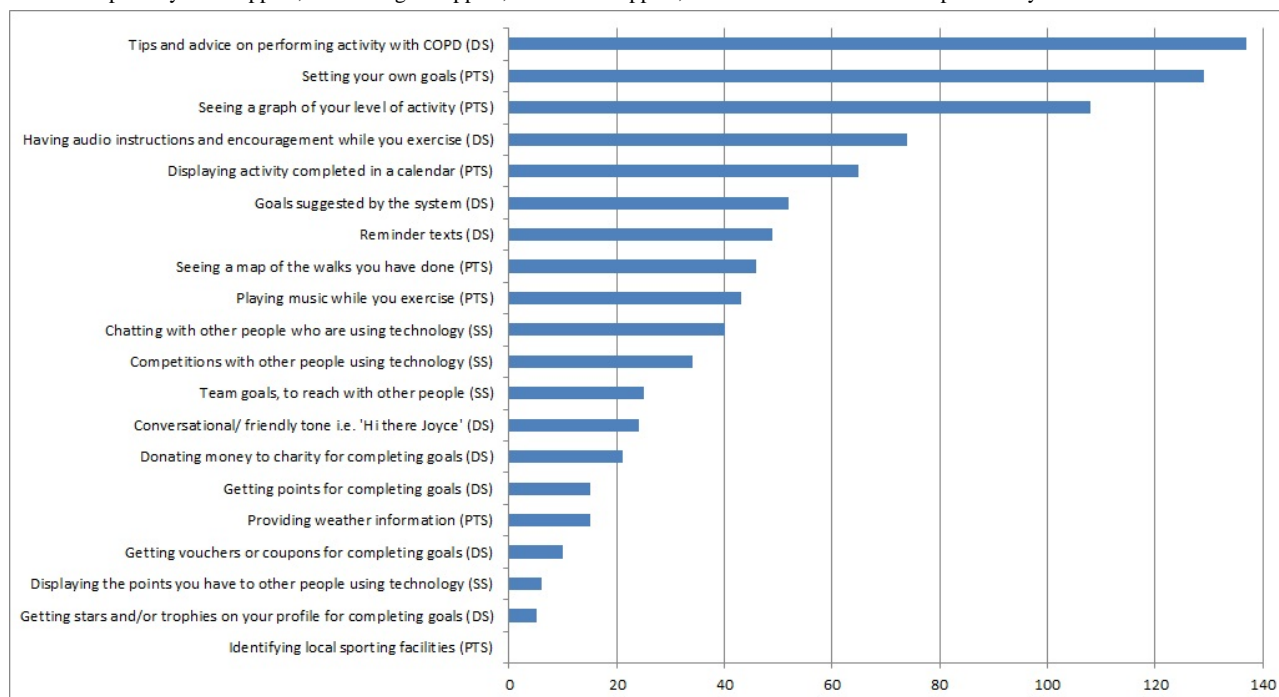
Among the respondents to the questionnaire that ranked at least 5 features of the persuasive technologies ($n=54$), scores were reversed so that a high ranking was associated with a high score. Figure 6 shows how participants ranked the different features. The feature that was ranked as the most likely to convince participants to use the technology was "Tips and advice on performing activity with COPD."

The features were then grouped according to the element of the PSD model that they addressed independently of the prototype they were presented within, namely, primary task support, dialogue support, or social support. A Friedman test identified a significant difference between the persuasiveness of features associated with different elements of the PSD model ($\chi^2_2=33.0$, $P<.001$), and a Wilcoxon signed rank test (with Bonferroni correction applied) found that features associated with primary task support were rated significantly more likely to convince PwCOPD to use the technology (mean 7.52, SD 4.63) than those associated with social support (mean 1.94, SD 2.97; $Z=-5.00$, $P<.001$). Features associated with dialogue support were also

rated significantly higher (mean 7.17, SD 5.53) than those associated with social support ($Z=-4.92$, $P<.001$). There was no significant difference between participants' ratings of features associated with primary task support and those associated with dialogue support ($Z=-.25$, $P=.80$).

It was found that 32 participants (59% 32/54) also indicated that some features would definitely not convince them to use persuasive technology; the most commonly identified features being identifying local sporting facilities ($n=23$), getting stars or trophies on your profile for completing goals ($n=22$), and displaying the points that you have to other people who are using this technology ($n=21$).

Figure 6. Sum of ranks given to individual techniques and features. Principles of the persuasive systems design model they relate to is indicated in brackets. PTS: primary task support; DS: dialogue support; SS: social support; COPD: chronic obstructive pulmonary disease.



Discussion

The findings of this research suggest that persuasive technology techniques designed to encourage and support physical activity among PwCOPD were received relatively positively by PwCOPD, their carers, and HCPs involved in their care. Below, we discuss the findings in the context of the PSD model and other theoretical literature, as well as pointing to the practical implications of the findings.

Principal Findings

The prototype based on the dialogue support principle that described a virtual coach was, overall, deemed most likely to be used or recommended, and it was also deemed to be the most persuasive. In addition, the techniques associated with the primary task support principle were ranked as most likely to convince participants to use a technology (although not significantly different to the techniques associated with dialogue support). The prototype based on the social support design principle that described the online community was the least likely to be recommended for use by HCPs and was also rated as the least persuasive by PwCOPD who completed the questionnaire.

The positive view of both dialogue support and primary task support expressed in this research suggests that these persuasive design principles and associated techniques are acceptable to PwCOPD and are likely to be used. In support of this idea, a

review of apps designed to promote behavior change identified that primary task support elements are used most frequently in apps that target smoking, drinking, and weight loss, and the authors additionally suggest that dialogue support techniques might also be used to promote other behaviors [52]. The use of techniques associated with dialogue support has been shown to have a direct effect on how persuasive interventions are deemed to be, which in turn has been found to predict intentions to use and actual use of a website aimed to encourage weight loss and increase positive mood [14].

A recent systematic review found that self-monitoring was the most commonly used persuasive technology technique in apps that are designed to increase physical activity. Suggestions and praise were the most commonly used dialogue support techniques, whereas the technology taking a social role (also a dialogue support technique) was only used in a single paper [8]. Participants in this research felt that the social role provided by technology should not duplicate other forms of support. For example, those who received encouragement from a carer did not feel that they would also need it from a virtual coach. However, other participants reported that they would value encouragement from technology, perhaps because they did not receive it from other sources.

The competition element of the online community prototype divided opinion. Some PwCOPD felt that competition would be encouraging, whereas others felt that the danger of becoming

disheartened was too great and that any system should aim to provide positive support and encouragement only. This opinion was echoed by the HCPs who were interviewed. Unlike rewards and maps, which were thought to be potential extras that might be ignored, HCPs seemed to recognize that competition could be persuasive, but felt that this approach to persuasion was inappropriate. Competition has been identified as a design element that is commonly used in games [53], and thus, could be considered a form of “gamification” (where game elements are used in nongame contexts such as health care [54]). Research that looks more broadly at the application of gamification has identified that the context that gamification is used in as well as characteristics of the user (eg, age) may influence the effectiveness of using game strategies to persuade [55,56], although little work thus far has explored the use of gamification in an older adult population [57].

One explanation for the divided opinion of competition in this study might be that different people anticipate experiencing different emotions as a result of competing. Some participants compared the competitive element provided by the online community to other competitive activities that they enjoyed such as playing cards or quizzes. These participants perhaps anticipated experiencing positive emotions when competing and viewed competition as fun. Other participants, however, may have feared failing and so reacted negatively to the prospect of competition. In the PSD model, competition is reported to leverage human beings’ “natural drive to compete” (p.495; [12]), and the inclusion of competitive elements has been found to have a positive effect on levels of physical activity interventions among healthy adults in some studies [8], as well as being frequently used in apps designed to promote weight loss [52]. It could be suggested, however, that the challenges that physical activity presents to PwCOPD means that competition is associated with stronger emotions. PwCOPD who appreciate the importance of physical activity may connect a failure to be active with a decline in their health, and therefore, the consequences of not succeeding in the competition may be viewed as more serious than they might be among healthy adults. Therefore, this hypothesis, as well as further exploration of the potential for other gamification strategies in the context of COPD care, warrants further consideration.

The findings suggest that a system that supports dialogue between the user and the technology alongside supporting the primary task (here, walking) to promote the self-regulation of physical activity is likely to be acceptable to PwCOPD and perceived as persuasive. Previous research indicates that these design principles are associated with adherence to Web-based health interventions [13] as well as intention to use and actual use of a Web-based intervention designed to promote healthy eating [14]. In contrast, the use of the social support design principle, while potentially engaging for some, is less likely to appeal to the majority of users. As discussed above, this is likely to be especially true for techniques that encourage any form of social comparison or competition. Further research should explore the use of persuasive technology techniques not only to promote both initial interest in the technology but also to support continued engagement. As this research also focused only on encouraging regular walking, it may also be helpful for future research to consider a wider range of physical activities that are suitable for PwCOPD (eg, wall push-ups) over the longer term. If PwCOPD are willing to engage with persuasive technology, then applications could also extend beyond promoting physical activity to other aspects of managing COPD such as promoting the use of breathing exercises, and providing relevant information.

Conclusions

This research investigated the opinions of PwCOPD, their carers, and HCPs involved in their care toward the use of different forms of persuasive technology to support and encourage increases in physical activity among PwCOPD. Opinions of persuasive technology were on the whole positive; however, opinions depended on personal preferences and initial levels of capability and motivation to engage with both physical activity and technology. Our findings suggested that a prototype describing a virtual coach designed to support interactions between the user and the technology was the most popular, and that techniques related to both supporting dialogue and primary task support were better supported by participants than those related to social support. We therefore recommend that future research integrate dialogue and primary support techniques into apps for PwCOPD and build on these findings to further explore how persuasive technology can be used to engage and meet the needs of this population.

Acknowledgments

The work was funded by the EPSRC. The sponsors did not have a role in the preparation or publication of this manuscript. This research was supported by the NIHR Collaboration for Leadership in Applied Health Research and Care Yorkshire and Humber (NIHR CLAHRC YH). The views and opinions expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. The authors would like to thank Paschal Sheeran for advice on the initial design of the study and Claire Bentley, The British Lung Foundation, Breathe Easy support groups and Breathing Space Rotherham for their help with recruitment.

Authors' Contributions

YKB designed the study, recruited participants, collected the data, analyzed the data, and prepared the initial draft of this manuscript. MH advised on the initial concept and design of the study. MH and TLW advised on data analyses, interpretation, and revised drafts of the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. San Fransisco: Morgan Kaufmann; 2003.
2. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J Public Health Inform* 2014;5(3):229 [FREE Full text] [doi: [10.5210/ojphi.v5i3.4814](https://doi.org/10.5210/ojphi.v5i3.4814)] [Medline: [24683442](https://pubmed.ncbi.nlm.nih.gov/24683442/)]
3. Hamari J, Koivisto J, Pakkanen T. Do persuasive technologies persuade? a review of empirical studies. In: Proceedings of the 9th International conference. 2014 Presented at: PERSUASIVE 2014; 2014; Padua, Italy. [doi: [10.1007/978-3-319-07127-5_11](https://doi.org/10.1007/978-3-319-07127-5_11)]
4. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. *Health Educ Behav* 2013 Apr;40(2):133-139. [doi: [10.1177/1090198112452126](https://doi.org/10.1177/1090198112452126)] [Medline: [22991048](https://pubmed.ncbi.nlm.nih.gov/22991048/)]
5. West JH, Hall PC, Hanson CL, Barnes MD, Giraud-Carrier C, Barrett J. There's an app for that: content analysis of paid health and fitness apps. *J Med Internet Res* 2012;14(3):e72 [FREE Full text] [doi: [10.2196/jmir.1977](https://doi.org/10.2196/jmir.1977)] [Medline: [22584372](https://pubmed.ncbi.nlm.nih.gov/22584372/)]
6. Breton ER, Fuemmeler BF, Abrams LC. Weight loss-there is an app for that! but does it adhere to evidence-informed practices? *Transl Behav Med* 2011 Dec;1(4):523-529 [FREE Full text] [doi: [10.1007/s13142-011-0076-5](https://doi.org/10.1007/s13142-011-0076-5)] [Medline: [24073074](https://pubmed.ncbi.nlm.nih.gov/24073074/)]
7. Azar KMJ, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. *Am J Prev Med* 2013 Nov;45(5):583-589. [doi: [10.1016/j.amepre.2013.07.005](https://doi.org/10.1016/j.amepre.2013.07.005)] [Medline: [24139771](https://pubmed.ncbi.nlm.nih.gov/24139771/)]
8. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. *J Med Syst* 2016 Mar;40(3):72. [doi: [10.1007/s10916-015-0425-x](https://doi.org/10.1007/s10916-015-0425-x)] [Medline: [26748792](https://pubmed.ncbi.nlm.nih.gov/26748792/)]
9. Abrams LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. *Am J Prev Med* 2011 Mar;40(3):279-285 [FREE Full text] [doi: [10.1016/j.amepre.2010.10.032](https://doi.org/10.1016/j.amepre.2010.10.032)] [Medline: [21335258](https://pubmed.ncbi.nlm.nih.gov/21335258/)]
10. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. *Telemed J E Health* 2011 Dec;17(10):801-803. [doi: [10.1089/tmj.2011.0055](https://doi.org/10.1089/tmj.2011.0055)] [Medline: [22010977](https://pubmed.ncbi.nlm.nih.gov/22010977/)]
11. Rosser BA, Eccleston C. Smartphone applications for pain management. *J Telemed Telecare* 2011;17(6):308-312. [doi: [10.1258/jtt.2011.101102](https://doi.org/10.1258/jtt.2011.101102)] [Medline: [21844177](https://pubmed.ncbi.nlm.nih.gov/21844177/)]
12. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. *Communications of the Association for Information Systems* 2009;24(28):485-500 [FREE Full text]
13. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res* 2012;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](https://doi.org/10.2196/jmir.2104)] [Medline: [23151820](https://pubmed.ncbi.nlm.nih.gov/23151820/)]
14. Drozd F, Lehto T, Oinas-Kukkonen H. Exploring perceived persuasiveness of a behavior change support system: a structural model. 2012 Presented at: Persuasive, 7th International conference; June 6-8, 2012; Linköping, Sweden. [doi: [10.1007/978-3-642-31037-9_14](https://doi.org/10.1007/978-3-642-31037-9_14)]
15. Pinzon O, Iyengar M. Persuasive technology and mobile health: a systematic review. 2012 Presented at: Persuasive, 7th International Conference; June 6-8, 2012; Linköping, Sweden.
16. Guarascio AJ, Ray SM, Finch CK, Self TH. The clinical and economic burden of chronic obstructive pulmonary disease in the USA. *Clinicoecon Outcomes Res* 2013;5:235-245 [FREE Full text] [doi: [10.2147/CEOR.S34321](https://doi.org/10.2147/CEOR.S34321)] [Medline: [23818799](https://pubmed.ncbi.nlm.nih.gov/23818799/)]
17. World Health Organization. 2016. Burden of COPD URL: <http://www.who.int/respiratory/copd/burden/en/> [accessed 2017-03-29] [WebCite Cache ID 6kDzwmtRy]
18. National Institute for Health and Clinical Excellence. 2011 Chronic obstructive pulmonary disease: costing report implementing NICE guidance URL: <https://www.nice.org.uk/guidance/cg101/resources/costing-report-134511805> [accessed 2016-09-02] [WebCite Cache ID 6kE01n7eF]
19. Suter P, Hennessey B, Florez D, Newton Suter W. Review series: Examples of chronic care model: the home-based chronic care model: redesigning home health for high quality care delivery. *Chron Respir Dis* 2011;8(1):43-52. [doi: [10.1177/1479972310396031](https://doi.org/10.1177/1479972310396031)] [Medline: [21339374](https://pubmed.ncbi.nlm.nih.gov/21339374/)]
20. Department of Health. National Archives. 2004. Annual report of the chief medical officer on the state of public health: tobacco and borders, death made cheaper URL: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/AnnualReports/Browsable/DH_4117035 [accessed 2017-03-29] [WebCite Cache ID 6pKHoyupX]
21. Bourbeau J. Activities of life: the COPD patient. *COPD* 2009 Jun;6(3):192-200. [Medline: [19811375](https://pubmed.ncbi.nlm.nih.gov/19811375/)]
22. Troosters T, Casaburi R, Gosselink R, Decramer M. Pulmonary rehabilitation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005 Jul 1;172(1):19-38. [doi: [10.1164/rccm.200408-1109SO](https://doi.org/10.1164/rccm.200408-1109SO)] [Medline: [15778487](https://pubmed.ncbi.nlm.nih.gov/15778487/)]

23. Pitta F, Troosters T, Probst VS, Spruit MA, Decramer M, Gosselink R. Quantifying physical activity in daily life with questionnaires and motion sensors in COPD. *Eur Respir J* 2006 May;27(5):1040-1055 [FREE Full text] [doi: [10.1183/09031936.06.00064105](https://doi.org/10.1183/09031936.06.00064105)] [Medline: [16707399](https://pubmed.ncbi.nlm.nih.gov/16707399/)]
24. Garcia-Aymerich J, Lange P, Benet M, Schnohr P, Antó JM. Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. *Thorax* 2006 Sep;61(9):772-778 [FREE Full text] [doi: [10.1136/thx.2006.060145](https://doi.org/10.1136/thx.2006.060145)] [Medline: [16738033](https://pubmed.ncbi.nlm.nih.gov/16738033/)]
25. Waschki B, Kirsten A, Holz O, Müller K, Meyer T, Watz H, et al. Physical activity is the strongest predictor of all-cause mortality in patients with COPD: a prospective cohort study. *Chest* 2011 Aug;140(2):331-342. [doi: [10.1378/chest.10-2521](https://doi.org/10.1378/chest.10-2521)] [Medline: [21273294](https://pubmed.ncbi.nlm.nih.gov/21273294/)]
26. American Thoracic Society, The European Respiratory Society. Thoracic. 2004. Standards for the diagnosis and management of patients with COPD URL: <http://www.thoracic.org/copd-guidelines/resources/copddoc.pdf> [accessed 2016-09-02] [WebCite Cache ID 6kE06bmm2]
27. National Institute for Health and Care Excellence. 2010. Chronic obstructive pulmonary disease in over 16s: diagnosis and management URL: <https://www.nice.org.uk/guidance/CG101/chapter/1-Guidance> [accessed 2016-09-02] [WebCite Cache ID 6kE08qMXP]
28. McCarthy B, Casey D, Devane D, Murphy K, Murphy E, Lacasse Y. Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2015 Feb 23(2):CD003793. [doi: [10.1002/14651858.CD003793.pub3](https://doi.org/10.1002/14651858.CD003793.pub3)] [Medline: [25705944](https://pubmed.ncbi.nlm.nih.gov/25705944/)]
29. Egan C, Deering BM, Blake C, Fullen BM, McCormack NM, Spruit MA, et al. Short term and long term effects of pulmonary rehabilitation on physical activity in COPD. *Respir Med* 2012 Dec;106(12):1671-1679 [FREE Full text] [doi: [10.1016/j.rmed.2012.08.016](https://doi.org/10.1016/j.rmed.2012.08.016)] [Medline: [23063203](https://pubmed.ncbi.nlm.nih.gov/23063203/)]
30. Dyer CAE, Harris ND, Jenkin E, Langley-Johnson C, Lewis R, Taylor GJ, et al. Activity levels after pulmonary rehabilitation - what really happens? *Physiotherapy* 2013 Sep;99(3):228-232. [doi: [10.1016/j.physio.2012.01.004](https://doi.org/10.1016/j.physio.2012.01.004)] [Medline: [23219642](https://pubmed.ncbi.nlm.nih.gov/23219642/)]
31. Heppner PS, Morgan C, Kaplan RM, Ries AL. Regular walking and long-term maintenance of outcomes after pulmonary rehabilitation. *J Cardiopulm Rehabil* 2006;26(1):44-53. [Medline: [16617228](https://pubmed.ncbi.nlm.nih.gov/16617228/)]
32. Esteban C, Arostegui I, Aburto M, Moraza J, Quintana JM, Aizpiri S, et al. Influence of changes in physical activity on frequency of hospitalization in chronic obstructive pulmonary disease. *Respirology* 2014 Apr;19(3):330-338 [FREE Full text] [doi: [10.1111/resp.12239](https://doi.org/10.1111/resp.12239)] [Medline: [24483954](https://pubmed.ncbi.nlm.nih.gov/24483954/)]
33. Moy ML, Teylan M, Weston NA, Gagnon DR, Garshick E. Daily step count predicts acute exacerbations in a US cohort with COPD. *PLoS One* 2013;8(4):e60400 [FREE Full text] [doi: [10.1371/journal.pone.0060400](https://doi.org/10.1371/journal.pone.0060400)] [Medline: [23593211](https://pubmed.ncbi.nlm.nih.gov/23593211/)]
34. Moy ML, Weston NA, Wilson EJ, Hess ML, Richardson CR. A pilot study of an Internet walking program and pedometer in COPD. *Respir Med* 2012 Sep;106(9):1342-1350 [FREE Full text] [doi: [10.1016/j.rmed.2012.06.013](https://doi.org/10.1016/j.rmed.2012.06.013)] [Medline: [22795984](https://pubmed.ncbi.nlm.nih.gov/22795984/)]
35. Nguyen HQ, Donesky D, Reinke LF, Wolpin S, Chyall L, Benditt JO, et al. Internet-based dyspnea self-management support for patients with chronic obstructive pulmonary disease. *J Pain Symptom Manage* 2013 Jul;46(1):43-55 [FREE Full text] [doi: [10.1016/j.jpainsymman.2012.06.015](https://doi.org/10.1016/j.jpainsymman.2012.06.015)] [Medline: [23073395](https://pubmed.ncbi.nlm.nih.gov/23073395/)]
36. Wang C, Chou P, Joa W, Chen L, Sheng T, Ho S, et al. Mobile-phone-based home exercise training program decreases systemic inflammation in COPD: a pilot study. *BMC Pulm Med* 2014 Aug 30;14:142 [FREE Full text] [doi: [10.1186/1471-2466-14-142](https://doi.org/10.1186/1471-2466-14-142)] [Medline: [25175787](https://pubmed.ncbi.nlm.nih.gov/25175787/)]
37. Moy ML, Janney AW, Nguyen HQ, Matthes KR, Cohen M, Garshick E, et al. Use of pedometer and internet-mediated walking program in patients with chronic obstructive pulmonary disease. *J Rehabil Res Dev* 2010;47(5):485-496 [FREE Full text] [Medline: [20803392](https://pubmed.ncbi.nlm.nih.gov/20803392/)]
38. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. It's Life! Mobile and web-based monitoring and feedback tool embedded in primary care increases physical activity: a cluster randomized controlled trial. *J Med Internet Res* 2015 Jul 24;17(7):e184 [FREE Full text] [doi: [10.2196/jmir.4579](https://doi.org/10.2196/jmir.4579)] [Medline: [26209025](https://pubmed.ncbi.nlm.nih.gov/26209025/)]
39. Moy ML, Martinez CH, Kadri R, Roman P, Holleman RG, Kim HM, et al. Long-term effects of an internet-mediated pedometer-based walking program for chronic obstructive pulmonary disease: randomized controlled trial. *J Med Internet Res* 2016 Aug 08;18(8):e215 [FREE Full text] [doi: [10.2196/jmir.5622](https://doi.org/10.2196/jmir.5622)] [Medline: [27502583](https://pubmed.ncbi.nlm.nih.gov/27502583/)]
40. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: Sage Publications, Inc; 2011.
41. Catani M, Biers D. Usability evaluation prototype fidelity: users and usability professionals. 1998 Presented at: The Human Factors Society 42nd Annual Meeting; October 5-9, 1998; Chicago, IL URL: <http://journals.sagepub.com/doi/abs/10.1177/154193129804201901>
42. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol* 2008 May;27(3):379-387. [doi: [10.1037/0278-6133.27.3.379](https://doi.org/10.1037/0278-6133.27.3.379)] [Medline: [18624603](https://pubmed.ncbi.nlm.nih.gov/18624603/)]
43. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. *Psychol Bull* 1982 Jul;92(1):111-135. [Medline: [7134324](https://pubmed.ncbi.nlm.nih.gov/7134324/)]
44. Fletcher CM, Elmes PC, Fairbairn MB, Wood CH. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *Br Med J* 1959 Aug 29;2(5147):257-266 [FREE Full text] [Medline: [13823475](https://pubmed.ncbi.nlm.nih.gov/13823475/)]

45. Stenton C. The MRC breathlessness scale. *Occup Med (Lond)* 2008 May;58(3):226-227. [doi: [10.1093/occmed/kqm162](https://doi.org/10.1093/occmed/kqm162)] [Medline: [18441368](#)]
46. National Health Service (NHS). Physical activity guidelines for older adults URL: <http://www.nhs.uk/Livewell/fitness/Pages/physical-activity-guidelines-for-older-adults.aspx> [accessed 2017-01-20] [WebCite Cache ID 6nedba4pa]
47. Lund A, Lund M. Lærd Statistics: Cronbach's Alpha. URL: <https://statistics.laerd.com/> [WebCite Cache ID 6kE0DzCs8]
48. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *BMJ* 2010;341:c4587. [Medline: [20851841](#)]
49. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess R, editors. *Analysing Qualitative Data*. London: Routledge; 1994:173-194.
50. Miles MB, Huberman AM. *An expanded sourcebook: Qualitative Data Analysis*, 2nd edition. Thousand Oaks, CA: Sage Publications; 1994.
51. Fetter MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. *Health Serv Res* 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: [10.1111/1475-6773.12117](https://doi.org/10.1111/1475-6773.12117)] [Medline: [24279835](#)]
52. Lehto T, Oinas-Kukkonen H. Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. *J Med Internet Res* 2011;13(3):e46 [FREE Full text] [doi: [10.2196/jmir.1559](https://doi.org/10.2196/jmir.1559)] [Medline: [21795238](#)]
53. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. *JMIR Serious Games* 2014;2(2):e9 [FREE Full text] [doi: [10.2196/games.3413](https://doi.org/10.2196/games.3413)] [Medline: [25654660](#)]
54. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. : ACM; 2011 Presented at: 15th International Academic MindTrek Conference: Envisioning Future Media Environments; September 28 - 30, 2011; Tampere, Finland URL: <http://dl.acm.org/citation.cfm?id=2181040>
55. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games* 2013;1(1):e3 [FREE Full text] [doi: [10.2196/games.3139](https://doi.org/10.2196/games.3139)] [Medline: [25658754](#)]
56. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. *Int J Med Inform* 2016 Dec;96:62-70. [doi: [10.1016/j.ijmedinf.2016.02.006](https://doi.org/10.1016/j.ijmedinf.2016.02.006)] [Medline: [26944611](#)]
57. de Vette F, Tabak M, Dekker-van Weering M, Vollenbroek-Hutten M. Engaging elderly people in telemedicine through gamification. *JMIR Serious Games* 2015 Dec 18;3(2):e9 [FREE Full text] [doi: [10.2196/games.4561](https://doi.org/10.2196/games.4561)] [Medline: [26685287](#)]

Abbreviations

COPD: chronic obstructive pulmonary disease
HCP: health care professional
MRC: Medical Research Council
NHS: National Health Service
PR: pulmonary rehabilitation
PSD: persuasive systems design
PwCOPD: people with chronic obstructive pulmonary disease

Edited by G Eysenbach; submitted 29.10.16; peer-reviewed by C Hanson, KT Win; comments to author 22.12.16; revised version received 02.02.17; accepted 21.02.17; published 20.04.17.

Please cite as:

Bartlett YK, Webb TL, Hawley MS

Using Persuasive Technology to Increase Physical Activity in People With Chronic Obstructive Pulmonary Disease by Encouraging Regular Walking: A Mixed-Methods Study Exploring Opinions and Preferences

J Med Internet Res 2017;19(4):e124

URL: <http://www.jmir.org/2017/4/e124/>

doi: [10.2196/jmir.6616](https://doi.org/10.2196/jmir.6616)

PMID: [28428155](#)

©Yvonne Kiera Bartlett, Thomas L Webb, Mark S Hawley. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 20.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Therapist-Supported Internet-Based Cognitive Behavior Therapy for Stress, Anxiety, and Depressive Symptoms Among Postpartum Women: A Systematic Review and Meta-Analysis

Ying Lau¹, RN, RM, B HealthSc, BN(Hons), MNurs, PhD; Tha Pyai Htun¹, MBBS, MPH; Suei Nee Wong², BSc (Hons), MSc; Wai San Wilson Tam¹, BSc, Mphil, PhD; Piyanee Klainin-Yobas¹, BSN, MSN, PhD

¹Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

²Medical Resource Team, National University of Singapore Libraries, Singapore, Singapore

Corresponding Author:

Ying Lau, RN, RM, B HealthSc, BN(Hons), MNurs, PhD
Alice Lee Centre for Nursing Studies
Yong Loo Lin School of Medicine
National University of Singapore
Level 2, Clinical Research Centre, Block MD11, 10 Medical Drive
Singapore, 117597
Singapore
Phone: 65 66011603
Fax: 65 67767135
Email: nurly@nus.edu.sg

Abstract

Background: A growing number of meta-analyses have supported the application of therapist-supported Internet-based cognitive behavior therapy (iCBT) for psychological disorders across different populations, but relatively few meta-analyses have concentrated on postpartum women.

Objective: This meta-analysis evaluated the efficacy of therapist-supported iCBT in improving stress, anxiety, and depressive symptoms among postpartum women.

Methods: A total of 10 electronic databases were used to search for published and unpublished trials. Cochrane Collaboration tool for assessing risk of bias was utilized to measure methodological quality. Meta-analysis was performed using the RevMan software (Review Manager version 5.3 for Windows from the Nordic Cochrane Centre, the Cochrane Collaboration, 2014). Among the 789 studies identified, 8 randomized controlled trials were selected, involving 1523 participants across 6 countries.

Results: More than half (65%) of the eligible studies had a low risk of bias with no heterogeneity. Results revealed that therapist-supported iCBT significantly improved stress ($d=0.84$, $n=5$), anxiety ($d=0.36$, $n=6$), and depressive symptoms ($d=0.63$, $n=8$) of the intervention group compared with those of the control group at post-intervention.

Conclusions: This review revealed that therapist-supported iCBT significantly improves stress, anxiety, and depressive symptoms among postpartum women with small to large effects. Future effectiveness studies should establish the essential components, format, and approach of iCBT with optimal levels of human support to maximize a long-term effect.

(*J Med Internet Res* 2017;19(4):e138) doi:[10.2196/jmir.6712](https://doi.org/10.2196/jmir.6712)

KEYWORDS

Internet; post-traumatic stress disorders; stress; anxiety; depression; cognitive behavior therapy; meta-analysis

Introduction

Cognitive behavior therapy (CBT) is a form of psychotherapy based on the assumption that all psychological disturbances are caused by dysfunctional thinking [1]. CBT aims to modify thoughts, beliefs, and perceptions, and change behavioral pattern

[2]. Cognitive restructuring is a common psychotherapeutic process in CBT for identifying, evaluating, and changing negative, distorted thoughts and beliefs [3]. Cognitive restructuring is a useful technique to understand wrong automatic beliefs and it helps individuals to reframe their negative perception or distorted thinking in a more positive

frame of mind [3]. Behavioral activation, which is a different form of CBT, is a functional analytical approach for engaging in enjoyable activities frequently to maintain or improve psychological well-being [4]. Behavioral activation is a development of activity scheduling that focuses on the use of avoided activities as a guide for activity scheduling and functional analysis of cognitive processes that involve avoidance [4]. Despite the well-established efficacy of CBT for treating and preventing psychological disorders [5], barriers to the administration of CBT exist, which may include insufficient therapists, stigmatization, geographical remoteness, long waiting times, and high costs [6]. CBT has been suggested suitable for remote delivery because of its structured content [7]. During the past 15 years, the development of Internet-based psychological intervention has progressed significantly [8], particularly in light of the rapid improvements in Internet technologies globally [9]. An innovative administration has been developed in the form of Internet-based CBT (iCBT) to minimize treatment barriers and increase access to care [8,10].

The Internet makes CBT feasible and worth consideration. Implementation of iCBT is able to administer the full course of a CBT treatment using online self-help format, which might or might not be supported by a therapist [8]. A therapist can provide support through phone, email, text, or an interactive computer interface [10,11]. Therapist-supported iCBT is a therapy that is guided by an identified therapist who gives feedback and answers to questions, and which can include interactive features through the Internet to get access to psychological treatment [8]. Recent systematic reviews of therapist-supported iCBT for anxiety [10] and psychiatric and somatic disorder [12] produced equivalent effect compared with face-to-face CBT. Therapist-supported interventions involving higher levels of human support improve outcomes for depression, but do not significantly affect outcomes for stress [13] or anxiety [10]. An Internet-based intervention with instant feedback has achieved adherence and effectiveness similar to that of the same intervention with human support [14]. The potential benefits of therapist-supported iCBT are customizability, cost-effectiveness, time-effectiveness, geographic flexibility, time flexibility, consistency, high availability, and rapid dissemination [15,16].

The postpartum period is a highly challenging time for women because of changes in physical, familial, financial, occupational, and other realms [17]. Changes may affect a woman's psychosocial and physical resources, resulting in stress, anxiety, and depressive symptoms [17,18]. Reported prevalence rates suggest that 14.3% of women suffer from general stress [18], 1%-30% from post-traumatic stress symptoms [19], 24.9% from anxiety [20], and 13%-19% from depressive symptoms [21] during the postpartum period. In addition, evidence indicates that the coexistence of stress, anxiety, and depressive symptoms occur during the postpartum period [17]. Timely and efficacious intervention is important during the postpartum period, especially when considering the adverse short- and long-term maternal health outcomes and child development outcomes [22-24]. Strong evidence supports that CBT is effective for treatment and prevention during the postpartum period [25]. A review revealed that interventions initiated during the

postpartum period were more effective than those initiated during the antenatal period and that the one-to-one therapy was more effective than the group therapy [25]. Notably, the Internet was found to be the preferred source of information among these women [20]. For example, 90% used the Internet to search for health-related information [20] and 69% used the Internet to seek formation about postpartum depression [26]. They expressed interest in use of Web-based resources and greater engagement in ehealth behaviors related to mental health [27]. An increasing number of randomized controlled trials (RCTs) found that iCBT is effective for stress [28,29], anxiety [30,31], and depressive symptoms [32,33] among postpartum women. Given the burgeoning development of iCBT for a broad range of conditions [8] paralleled by the rapid increase in access to instant cyber connectivity, what is the effect of therapist-supported iCBT in improving stress, anxiety, and depressive symptoms among postpartum women?

Meta-analyses are applied in relation to a wide range of study designs; those that address trial designs are considered to represent the strongest evidence, and they have become increasingly prevalent over the years [34]. A growing number of meta-analyses have supported the application of therapist-supported iCBT for stress [13,35], anxiety [7,10], and/or depressive symptoms [36,37] across different populations, but relatively few meta-analyses of therapist-supported iCBT have concentrated on the postpartum population. We are aware of two recent systematic reviews [38,39] that investigate computer- or Web-based interventions for the prevention and treatment of perinatal mental health [38] or mood disorders [39]. The Two reviews suggest that computer- or Web-based intervention is effective for depressive symptoms [38,39], but mixed results were found for stress and anxiety symptoms [38]. However, the two reviews have been limited in using heterogeneous study designs [38,39], different therapeutic approaches [38], nonspecific outcomes [38,39], and a few eligible studies ($n=4$) [39]. Although one review [38] used a meta-analytical approach, their population was varied with antenatal, postpartum women, and their partners.

Evidently, therapist-supported iCBT provides an efficacious, accessible, and economically sound intervention for a broad range of conditions [8] across different populations [7,13]. A systematic review of existing evidence is necessary to determine whether therapist-supported iCBT for postpartum women is efficacious for prevention and treatment of stress, anxiety, and depressive symptoms before therapist-supported iCBT is embedded in routine clinical practice [39]. Further exploration is recommended to address the gaps in the current literature. The findings of this review could guide future studies exploring the next steps of therapist-supported iCBT implementation among the postpartum population. By conducting a systematic review and meta-analysis, we synthesized the best available evidence. This review aimed to systematically assess studies that examined therapist-supported iCBT interventions for improving stress, anxiety, and depressive symptoms among postpartum women.

Methods

This review was conducted according to the standards outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [40]. The protocol is registered in the PROSPERO database (CRD42016039094).

Eligibility Criteria

Studies were selected for the meta-analysis if they fulfilled the PICOS (population, intervention, comparison, outcomes and study) criteria [19]:

- Population: target women with age ≥ 18 years in the postpartum period (≤ 2 years postpartum) [41]
- Intervention: Therapist-supported iCBT was delivered over the Internet through the use of websites, email, phone, or Skype. The iCBT must have included support of therapists through phone, email, text, and an interactive computer interface. The intervention comprised at least one of the elements of CBT, including cognitive restructuring, behavioral activation, or skills training
- Comparison: attention control, waitlist, or treatment as usual (TAU)
- Outcomes: stress, anxiety, and depressive symptoms at postintervention
- Type of studies: RCTs

We excluded studies on teenage pregnancy (< 18 years). We did not include studies that had no CBT component in the intervention group; had active treatment containing a CBT component in the control group; were clinical controlled trials, cross-sectional, cohort, one-group pre- and posttest, and qualitative designs; only had abstracts and were study protocols, reviews, or conference papers.

Search Strategy

Similar systematic review papers were searched from the Cochrane Databases of Systematic Reviews, Joanne Briggs Institute, Centre for Reviews and Dissemination, University of York, PubMed Clinical Queries, Google, and Google Scholar to verify that the papers had not been conducted recently. The search strategy aimed to find published or unpublished studies without time limitation to maximize the search. We did not restrict our search to studies reported in any particular language. However, we conducted searches in English.

A 3-step search strategy was employed from inception until February 9, 2017. The first phase was a comprehensive search using identified keywords and index terms, searching the following 10 electronic databases: EMBASE, PubMed, Cumulative Index to Nursing and Allied Health Literature, Academic Search Completed, PsycINFO, PsycARTICLES, Cochrane Library, Web of Science, Scopus, and ProQuest Dissertations and Theses. Index and keyword terms were used (Multimedia Appendix 1). The index terms and keywords were combined and truncated according to the syntax rules of each database.

The second phase involved searching ClinicalTrials.gov, WHO International Clinical Trials Registry Platform, and International Standard Randomised Controlled Trial for unpublished trials

relevant to the review. When eligible trials were found, unpublished data were requested.

The third phase involved searching the reference lists of the included studies and checking previous reviews relevant to the topic for additional studies. The bibliographical software package EndNote program version X7 (Thomas Reuters) was used to import all the references and remove duplicates. The remaining studies were assessed independently against the inclusion and exclusion criteria by two authors (ie, LY and TP).

Quality Assessment

After identifying full-text articles that fulfilled the selection criteria, the studies were submitted for quality assessment and verified for eligibility. Cochrane Collaboration tool for assessing risk of bias was used by the two authors to independently evaluate the potential for bias in each study [42]. The following indicators of internal validity specific to the methodology were collected: (1) random sequences generation, (2) allocation concealment, (3) blinding of participants and personnel, (4) blinding of outcome assessment, (5) incomplete outcome data, and (6) selective reporting [42]. Assessment related to the risk of bias was assigned with a judgment of “low risk” of bias, “high risk” of bias, or “unclear risk” of bias. Any difference in opinion between the two authors was resolved by consensus.

Data Extraction

The characteristics of trials and elements of therapist-supported iCBT were extracted from each study through structured summaries. Items extracted for the characteristics of trials included authors, year, countries, design, sample with health condition, age, name of iCBT, control group, sample size, outcomes, attrition rate, follow-up, intention-to-Treat (ITT) analysis, and grant support. Items extracted for the descriptions of iCBT were: aim, numbers of sessions, components, therapy, support, provider, peer support, partner support, contact with therapist, communication, interactivity, activities, multimedia, and duration. The summary tables were thoroughly reviewed for accuracy and relevance by the two authors independently. Study authors were contacted for any missing or additional information.

Data Analysis

Data were synthesized using the RevMan software (Review Manager Version 5.3 for Windows from the Nordic Cochrane Centre, the Cochrane Collaboration, 2014). We used the generic inverse variance method in our meta-analysis to combine the continuous outcomes with means and standard deviations (SD) [42]. The weight given to each trial was chosen to be the inverse of the variance of the effect estimate (ie, one over the square of its standard error) [42]. With the inverse variance method, mean difference was used for the scores of stress, anxiety, and depressive symptoms. The standardized mean differences with their corresponding 95% CI were employed to combine studies that measured the same outcomes with different methods. The test of the overall effect was assessed using Z-statistics at $P < .05$. To quantify the efficacy of therapist-supported iCBT on stress, anxiety, and depressive symptoms, we calculated the effect sizes by subtracting the mean value of the iCBT group from the mean value of the control group at posttest and dividing the result by

the pooled SD of the two groups. The effect sizes were expressed as Cohen d or standardized mean difference, which were interpreted as small ($0.2 < d < 0.5$), medium ($0.5 < d < 0.8$), and large ($d \geq 0.8$) [43].

Heterogeneity between studies was evaluated using the Cochrane Q (chi-square test) and I^2 statistics in the RevMan software. The statistical significance for heterogeneity was set as $P > .10$, and estimates of the degree of heterogeneity using I^2 were made by setting 0%–40% as might not be important, 30%–60% as moderate, 50%–90% as substantial, and 75%–100% as considerable [42]. We used fixed- and random-effects models in our meta-analysis according to heterogeneity between studies on different statistical assumptions [44]. The fixed-effects model was used to estimate one true effect in cases without significant heterogeneity ($P < .10$) because we assumed that the true effect size was the same in the eligible trials, whereas the random-effects model (tau-square) was employed to estimate the mean of a distribution of effects in cases with heterogeneity between studies ($P > .10$) and in those with I^2 values of over 40% because we assumed that the true effect size varied from trial to trial [42]. Subgroup analysis was performed to (1) explore the source of heterogeneity and (2) evaluate the effect in a

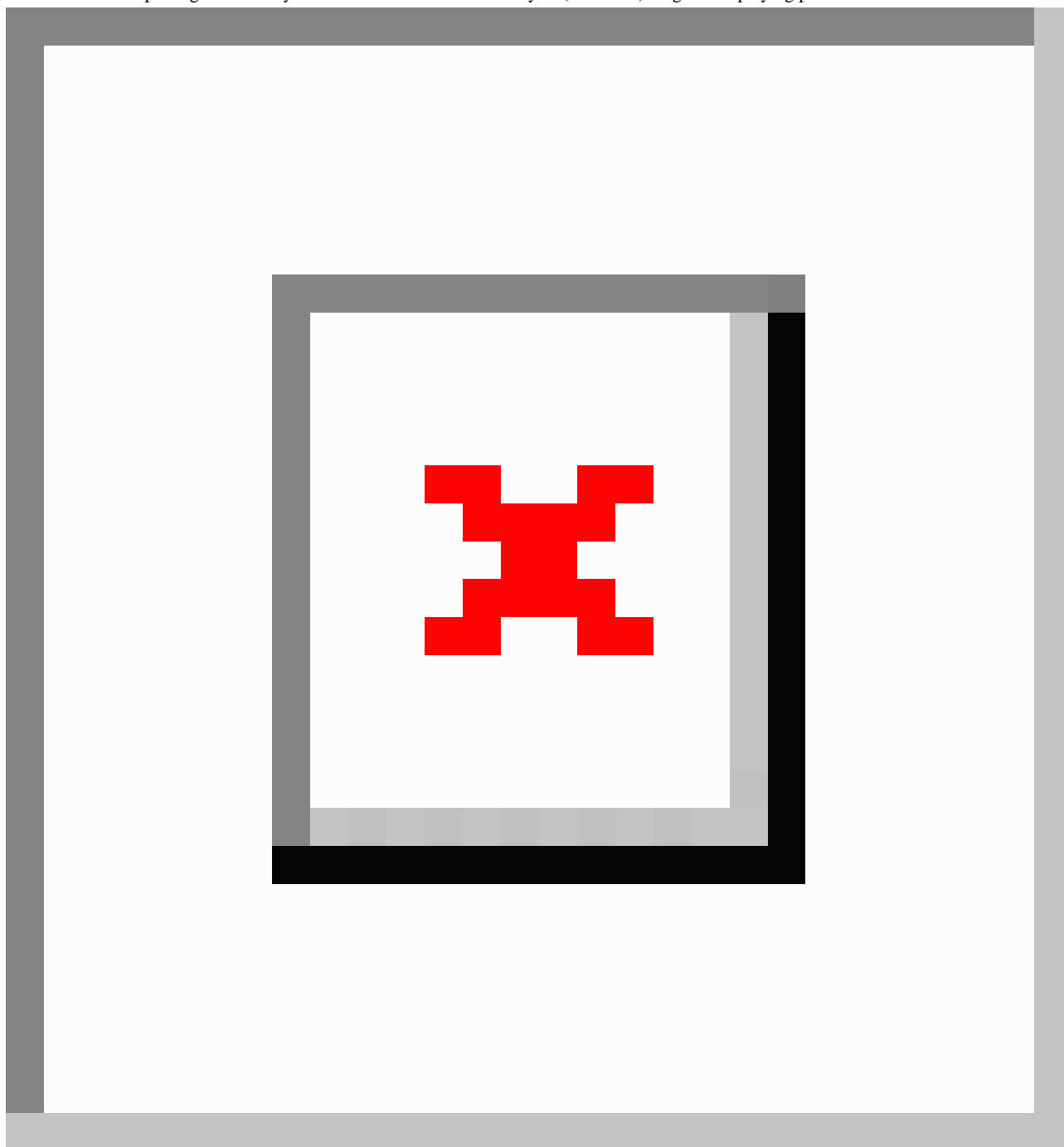
specific subgroup; performing a subgroup analysis provided information about essential elements for maximizing effectiveness of therapist-supported iCBT [45]. The predefined subgroups differed in health condition, control conditions, type of iCBT, number of sessions, and professional support. We planned to construct funnel plots to determine the possible influence of publication biases if the number of eligible trials was more than 10 [42].

Results

Study Selection

The search and selection of articles is illustrated in Figure 1. Using the specified search terms, the searched 10 databases produced a total of 596 records. Among these studies, 200 article duplicates were removed. One additional record was identified from the reference list. On the basis of an analysis of the words in paper titles and abstracts, 332 records were excluded. Full-text articles of the remaining 65 articles were retrieved, reviewed, and selected on the basis of relevance and quality for eligibility. A total of 57 studies were excluded for reasons outlined in Figure 1. Finally, 8 RCTs were selected for meta-analysis.

Figure 1. Preferred reporting items for systematic reviews and meta-analysis (PRISMA) diagram displaying procedure for trials selection.

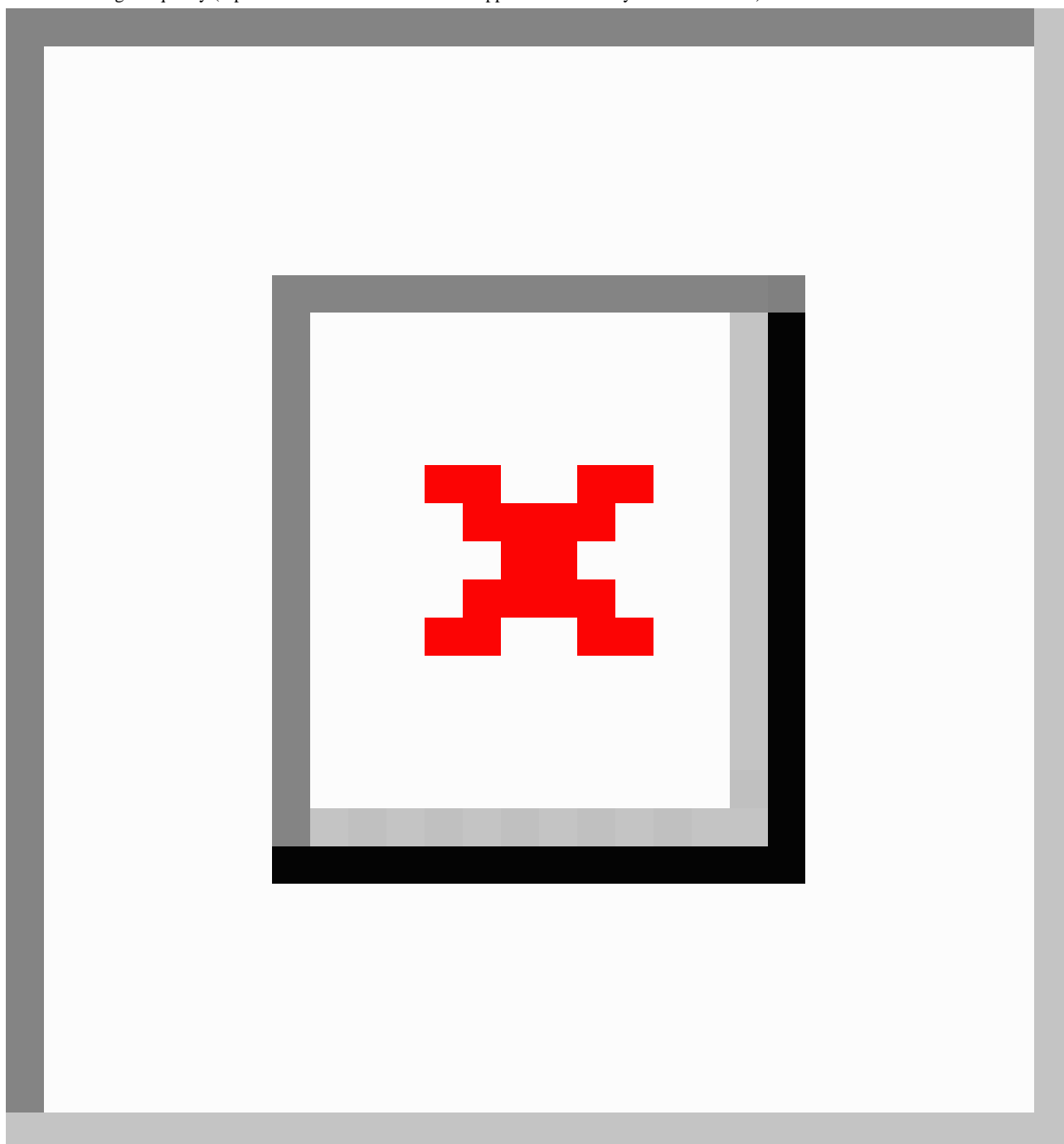


Risk of Bias in Included RCTs

We summarized the findings for the risk of bias graph and summary in [Figure 2](#) and [Multimedia Appendix 2](#), respectively. In all, 8 studies (100%) had adequate sequence generation for randomization. All of them (100%) had adequate allocation concealment. Only one study (13%) had implemented the blinding of participants and personnel and outcome assessment. Whereas 6 trials (75%) were unclear for the blinding of outcome assessment, 1 trial (13%) was unclear to implement the blinding of participants and personnel. Of the 8 studies, 5 (63%)

addressed low risk of bias concerning incomplete outcome data. All of the studies (100%) had low risk of bias for selective reporting. The two authors independently checked for risk of bias. The interrater agreement was 100% for global ratings, with two minor disagreements at the component level, but these disagreements were resolved through discussion. The number of trials ($n < 10$) in this review was insufficient to make a meaningful funnel plot to determine publication bias because the test power was excessively low to distinguish chance from real funnel plot asymmetry [42].

Figure 2. Risk of bias summary. These are authors' judgments of each methodological quality item for each included study. Plus signs (+) indicate high methodological quality (low risk of bias); minus signs (-) indicate low methodological quality (high risk of bias); question marks (?) indicate unclear methodological quality (reported information about what happened in the study was insufficient).



Characteristics of Included Studies

This meta-analysis included 8 studies [28-33,46,47] with a total of 1523 participants conducted across 6 countries (Table 1), including Australia (13%, 1/8) [28], Canada (13%, 1/8) [31], Germany (25%, 2/8) [46,47], Sweden (13%, 1/8), the United Kingdom (25%, 2/8) [30,32], and the United States (13%, 1/8) [33]. All of the studies were conducted between 2011 [47] and 2016 [31], with 2016 having the highest number of publications (38%, 3/8) [28,29,31].

The mean age of the participants in the studies ranged from 31 [33] to 35 years [29]. Participants were postpartum women with

3 different health conditions, namely, depressive symptoms (63%, 5/8) [28,30-33], post-traumatic stress symptoms (13%, 1/8) [29], and pregnancy loss (25%, 2/8) [46,47]. The sample sizes were between 43 [28] and 910 [32]. Comparators were waitlist treatment (50%, 4/8) [29,31,46,47], usual care (38%, 3/8) [28,30,32], and waitlist or usual care (13%, 1/8) [33]. Six of the studies (75%, 6/8) assessed more than one target outcome. Whereas the rest did not report any follow-up, 6 of the studies had a follow-up for 1 [31] to 12 months [46] after intervention. Attrition rates ranged from 3% [33] to 60.8% [32] and from 0% [33] to 63.8% [32] for the intervention and control groups, respectively. Of the all the selected articles, 5 (63%, 5/8) used

ITT analysis, and majority of them (88%, 7/8) were supported by grants.

Table 1. Characteristics of the eligible 8 randomized controlled trials.

Author (year)	Country	Health condition (recruitment)	Age (in years) (mean)	iCBT ^a (I)	Control (C)	Sample size (N)	Outcomes (measures)	Follow-up (in months)
Kersting et al (2011) ^{b,c} [47]	Germany	Pregnancy loss (–)	>18 (34)	Manualized CBT ^d treatment program	Waitlist	I: 48 C: 35	Stress (IES) ^e Anxiety (BSI-GSI) ^f Depression (BSI-GSI) ^f	3
Kersting et al (2013) ^{b,c} [46]	Germany	Pregnancy loss (–)	>18 (34)	Internet-based intervention for parents after parental loss	Waitlist	I: 115 C: 113	Stress (IES-R) ^g Anxiety (BSI-GSI) ^f Depression (BSI-GSI) ^f	3, 12
Milgrom et al (2016) ^{b,c} [28]	Australia	Major or minor depression (< 1 year postpartum)	>18 (32)	MumMood Booster	TAU ^h	I: 21 C: 22	Stress (DASS-21) ⁱ Anxiety (DASS-21) ⁱ Depression (BDI-II) ^j	3
Nieminen et al (2016) ^{b,c} [29]	Sweden	Post-traumatic stress (> 3 months postpartum)	>18 (35)	Internet-based trauma-focused CBT ^d	Waitlist	I: 28 C: 28	Stress (IES-R) ^g Anxiety (BAI) ^k Depression (BDI-II) ^j	(–)
O'Mahen et al (2013) ^c [32]	United Kingdom	Depressive symptoms (< 12 months postpartum)	>18 (32)	Postnatal internet-based behavioral activation	TAU ^h	I: 462 C: 448	Depression (EPDS) ^l	(–)
O'Mahen et al (2014) ^c [30]	United Kingdom	Depressive symptoms (< 12 months postpartum)	>18 (–)	Netmums Helping with Depression	TAU ^h	I: 41 C: 42	Anxiety (GAD-7) ^m Depression (EPDS) ^l	6
Pugh et al (2016) ^b [31]	Canada	Depressive symptoms (< 12 months postpartum)	>18 (–)	Therapist-assisted iCBT ^a	Waitlist	I: 25 C: 25	Stress (DASS-21) ⁱ Anxiety (DASS-21) ⁱ Depression (EPDS) ^l	1
Sheeber et al (2012) ^c [33]	United States	Depressive symptoms (–)	>18 (31)	Mom-Net intervention	Waitlist or TAU ^h	I: 35 C: 35	Depression (BDI-II) ^j	3

^aiCBT: Internet-based cognitive behavioral therapy.^bThese studies used intention-to-treat analysis.^cThese studies had grant support.^dCBT: cognitive behavioral therapy.^eIES: Impact of Event Scale.^fBSI-GSI: Brief Symptom Inventory-Global Severity Index.

^gIES-R: Impact of Event Scale-revised.

^hTAU: treatment as usual.

ⁱDASS-21: Depression Anxiety Stress Scale.

^jBDI-II: Beck Depression Inventory-II.

^kBAI: Beck Anxiety Inventory.

^lEPDS: Edinburgh Postnatal Depression Scale.

^mGAD-7: Generalized Anxiety Disorder Scale.

Descriptions of Internet-Based Cognitive Behavior Therapy

Detailed descriptions of iCBT are presented in [Multimedia Appendices 3](#) and [4](#). All iCBTs were based on cognitive-behavioral strategies, including cognitive restructuring [29,46,47] and behavioral activation [30,33] with additional elements of psychoeducation [31], parenting preparation, or parenting focus [33]. The numbers of sessions ranged from 6 [47] to 12 [30]. All therapies used one-to-one setting. Therapists support all therapies and 3 of them incorporated a self-monitoring or self-help component [28,30,33]. All of the studies had human support coming from psychologists or clinical psychologists (75%, 6/8) [28,29,31,32,46,47], master's or PhD psychology students (25%, 2/8) [29,31], CBT-trained mental health workers (25%, 2/8) [32,33], and a layperson (13%, 1/8) [33].

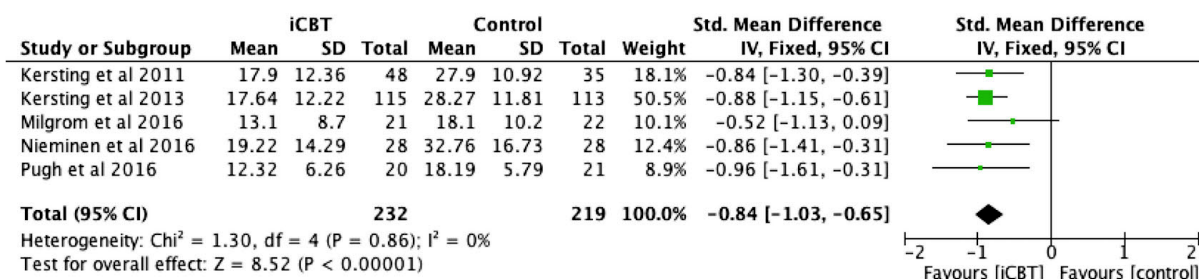
Peer support using chat room discussion was used by 4 of the therapies [28,30,32,33], and only 1 trial used partner support using websites [28]. Majority of the therapies were asynchronous (88%, 7/8) [29-33,46,47] and 1 was synchronous [28] in terms of two-way feedback communication with therapists using websites (88%, 7/8) [28-30,32,33,46,47], phone calls (63%, 5/8) [28,30-33], emails (38%, 3/8) [30-32], and text messages (13%, 1/8) [30]. The delivery modes included multimedia and interactive formats, such as animation, video, audio, or photos [28,31]. The activities of the therapies included text-based reading (63%, 5/8) [28,29,31,46,47], assignment or homework (88%, 7/8) [28-32,46,47], interactive exercise (13%, 1/8) [30],

expressive writing (38%, 3/8) [33,46,47], and online clinics or consultation (25%, 2/8) [32,33]. The duration of the therapies varied among the 8 studies, ranging from 5 [46,47] to 17 weeks [30].

Efficacy of Therapist-Supported iCBT on Stress Symptoms

By comparing the intervention and control groups using stress symptoms at post-intervention as a dependent variable, 5 studies [28,29,31,46,47] assessed the efficacy of therapist-supported iCBT among 451 women. Depression Anxiety Stress Scale-21 (DASS-21) [48], Impact of Event Scale (IES) [49] and IES-Revised [50] were used to measure stress symptoms. [Figure 3](#) shows that therapist-supported iCBT interventions had a large effect size of .84 (95% CI 0.65-1.03) on eliminating stress symptoms. The overall effect ($Z=8.52$, $P<.001$) was significant, but I^2 showed 0%, and the P value of chi-square was .86. A series of subgroup analyses was performed to explore more information about the therapist-supported iCBT intervention, as shown in [Table 2](#). Subgroup differences were found insignificant for stress symptoms based on health condition, control condition, number of sessions, peer support, and professional support. Interestingly, we observed the therapist-supported iCBT with waitlist as comparator and exclusive therapist support were significant ($d=0.88$, $Z=8.43$, $P<.001$) on eliminating stress symptoms, whereas intervention using TAU as comparator and therapist support with self-help component did not ($d=0.52$, $Z=1.66$, $P>.05$). However, only 1 study in the subgroups had limited generalizability and thus further trials were warranted.

Figure 3. Forest plot of standardized mean difference (95% CI) in change of stress symptoms scores for Internet-based cognitive behavior therapy intervention and control group.



Efficacy of Therapist-Supported iCBT on Anxiety Symptoms

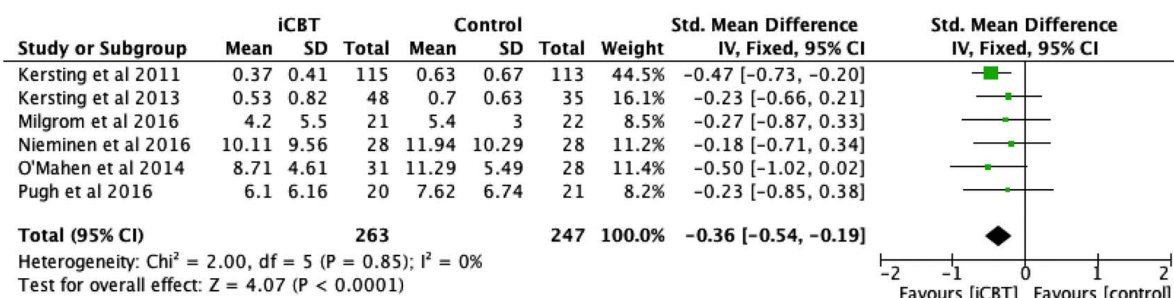
Using anxiety symptoms at post-intervention as a dependent variable in 6 studies [28-31,46,47], the efficacy of therapist-supported iCBT was evaluated among 510 participants by comparing the intervention and control groups. Anxiety

symptoms were measured through the Brief Symptom Inventory-Global Severity Index (BSI-GSI) [51], DASS-21 [48], Generalized Anxiety Disorder Scale [52], and Beck Anxiety Inventory [53]. Therapist-supported iCBT exerted a significant effect on improving anxiety symptoms ($Z=4.07$, $P<.001$) with small to medium effect size ($d=0.36$), as shown in [Figure 4](#). Meta-analysis on anxiety symptoms showed no

heterogeneity ($I^2=0\%$). Significant subgroup differences were not found for anxiety symptoms based on health condition,

control condition, type of iCBT, number of sessions, and professional support (Table 2).

Figure 4. Forest plot of standardized mean difference (95% CI) in change of anxiety symptoms scores for Internet-based cognitive behavior therapy intervention and control group.



Efficacy of Therapist-Supported iCBT on Depressive Symptoms

A total of 8 studies [28-33,46,47] assessed the efficacy of therapist-supported iCBT interventions among 934 participants by comparing the intervention and control groups using depressive symptoms at post-intervention as dependent variable. Depressive symptoms were measured through the Beck Depression Inventory-II [54], Edinburgh Postnatal Depression

Scale [55], and BSI-GSI [51]. iCBT interventions in this meta-analysis exerted a significant effect on improving depressive symptoms ($Z=9.42$, $P<.001$) with medium to large effect size of 0.63 (Figure 5). The meta-analysis of these 8 studies showed no heterogeneity ($I^2=0\%$). No significant subgroup differences were found for depressive symptoms according to health condition, control condition, type of iCBT, number of sessions, and professional support (Table 2).

Figure 5. Forest plot of standardized mean difference (95% CI) in change of depressive symptoms scores for Internet-based cognitive behavior therapy intervention and control group.

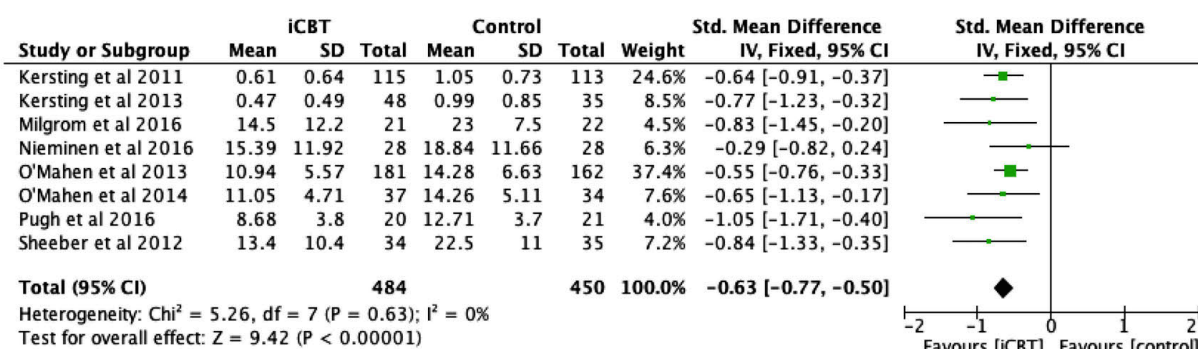


Table 2. Subgroup analyses of Internet-based cognitive behavioral therapy (iCBT) on stress, anxiety, and depressive symptoms.

Subgroup analyses	Number of comparisons	Effect size, <i>d</i>	95% CI	Z-statistics	<i>P</i>	Heterogeneity, <i>I</i> ² (%)	Chi-square, χ^2 (<i>df</i>), comparison
Stress symptoms [28,29,31,46,47]							
Health condition							
Depression [28,31]	2	0.72	0.28-1.17	3.18	.001	0	0.3 (2)
Pregnancy loss [46,47]	2	0.87	0.64-1.10	7.31	<.001	0	
Post-traumatic stress [29]	1	0.84	0.65-1.03	3.06	.002	N/A	
Control condition							
Waitlist [29,31,46,47]	4	0.88	0.64-1.08	8.43	<.001	0	1.2 (1)
Treatment as usual [28]	1	0.52	0.09-1.13	1.66	.10	N/A	
Number of sessions							
< 8 sessions [28,31,46,47]	4	0.84	0.63-1.05	7.95	<.001	0	0.0 (1)
≥ 8 sessions [29]	1	0.86	0.31-1.41	3.06	.002	N/A	
Professional support							
Exclusive therapist support [29,31,46,47]	4	0.88	0.67-1.08	8.43	<.001	0	1.2 (1)
With self-help component [28]	1	0.52	0.09-1.13	1.66	.10	N/A	
Anxiety symptoms [28-31,46,47]							
Health condition							
Depression [28,30,31]	3	0.35	0.02-0.68	2.09	.04	0	0.6 (2)
Pregnancy loss [46,47]	2	0.40	0.18-0.63	3.50	<.001	0	
Post-traumatic stress [29]	1	0.18	0.34-0.71	0.68	.50	N/A	
Control condition							
Waitlist [29,31,46,47]	4	0.35	0.16-0.55	3.54	<.001	0	0.1 (1)
Treatment as usual [28,30]	2	0.40	0.01-0.80	2.01	.04	0	
Type of iCBT^a							
Behavioral activation [30]	1	0.50	0.02-1.02	1.90	.06	N/A	0.3 (1)
Cognitive behavioral therapy [28,29,31,46,47]	5	0.35	0.16-0.53	3.64	<.001	0	
Number of sessions							
< 8 sessions [28,31,46,47]	4	0.37	0.17-0.57	3.63	<.001	0	0.0 (1)
≥ 8 sessions [29,30]	2	0.34	0.02-0.71	1.83	.07	0	
Professional support							
Exclusive therapist support [29,31,46,47]	4	0.35	0.16-0.55	3.54	<.001	0	0.1 (1)
With self-help component [28,30]	2	0.40	0.01-0.80	2.01	.04	0	
Depressive symptoms [28-33, 46,47]							
Health condition							
Depression [28,30-33]	5	0.65	0.48-0.82	7.50	<.001	0	1.8 (2)
Pregnancy loss [46,47]	2	0.67	0.44-0.90	5.76	<.001	0	
Post-traumatic stress [29]	1	0.29	0.24-0.82	1.07	.28	N/A	
Control condition							

Subgroup analyses	Number of comparisons	Effect size, <i>d</i>	95% CI	Z-statistics	<i>P</i>	Heterogeneity, <i>I</i> ² (%)	Chi-square, χ^2 (df), comparison
Waitlist [29,31,46,47]	4	0.65	0.45-0.85	6.39	<.001	16	0.9 (2)
Treatment as usual [28,30,32]	3	0.59	0.40-0.78	6.14	<.001	0	
Waitlist and treatment as usual [33]	1	0.84	0.35-1.33	3.34	<.001	N/A	
Type of iCBT							
Behavioral activation [30,32]	2	0.56	0.37-0.76	5.62	<.001	0	0.9 (1)
Cognitive behavioral therapy [28,29,31,33,46,47]	6	0.69	0.51-0.87	7.62	<.001	0	
Number of sessions							
< 8 sessions [28,31,46,47]	4	0.73	0.52-0.93	6.96	<.001	0	1.4 (1)
≥ 8 sessions [29,30,32,33]	4	0.57	0.40-0.74	6.45	<.001	0	
Professional support							
Exclusive therapist support [29-31,46,47]	5	0.65	0.47-0.84	6.92	<.001	0	0.1 (1)
With self-help component [28,32,33]	3	0.62	0.43-0.80	6.39	<.001	0	

^aiCBT, Therapist-supported Internet-based cognitive behavioral therapy.

Discussion

In this meta-analysis, the efficacy of the therapist-supported iCBT on stress, anxiety, and depressive symptoms during the postpartum period was searched through 10 databases. This study included 8 RCTs involving 1523 postpartum women using iCBT across 6 countries. The results revealed that iCBT significantly improved stress ($d = 0.84$, $n = 5$), anxiety ($d = 0.36$, $n = 6$), and depressive symptoms ($d = 0.63$, $n = 8$) of the intervention group compared with those of the control group at post-intervention.

Quality of the Evidence and Potential Biases

Data were independently extracted, checked, and entered. The methodological quality of the eligible studies was rated to assess the subjective bias. The overall methodological quality of the studies included in the review was mixed. All studies used methods that we judged to have low risk of bias to randomly assign participants to either the intervention or the control group. This result was due to the selection criteria for RCT. Thus, RCTs prevented selection bias and were insured against accidental bias. All studies achieved adequate allocation concealment. Therefore, participants were unlikely to have selection bias.

A potentially important source of bias in this meta-analysis was that only 12.5% (1/8) of the studies achieved the blinding of participants and personnel. Concealing treatment conditions from participants was impossible because of the control conditions used (eg, waitlist). The results might be influenced by performance bias as concealing of treatment conditions was not possible. Only 12.5% (1/8) of the studies achieved effective blinding of outcome assessment, perhaps primarily owing to the characteristics of the interventions. Hence, a high risk of detection bias remained possible for outcomes relying on

self-report or objective outcomes by outcome assessors who were not blinded to treatment allocation. The overall effect of the sample attrition had a low risk of bias in more than half of the studies (62.5%, 5/8), which could improve the generalizability of findings and reduce attrition bias. All studies reported outcomes in a pre-specified manner. Consequently, the eligible studies did not give misleading results because of the selective reporting of outcomes.

In addition, the attrition rates for the intervention and control groups were widely ranged (0%-63.8%). Providing a reminder or tracking system at pre-specified times to intervention users may reduce attrition and ultimately enhance outcomes [29]. Website contact system [29] and phone call [33] may be employed to remind postpartum women to use the site, point to helpful resources, or provide a connection with the treatment team. More than half of the trials (62.5%, 5/8) used ITT analysis, which is an analysis method for solving non-compliance and missing outcomes [56]. Therefore, half of the trials avoided overoptimistic estimates of the efficacy of therapist-supported iCBT by analysing outcomes according to original treatment allocation rather than only for participants completing treatment [56].

Therapist-Supported Internet-Based Cognitive Behavior Therapy

Consistent with a previous review [25], our finding revealed that interventions initiated during the postpartum period were effective. The majority of therapist-supported iCBT was using cognitive restructuring as an essential element to identify, dispute, and correct irrational or maladaptive thoughts [3]. In our review, therapist-supported iCBT provides a feasible, efficacious, accessible, and economically sound intervention for postpartum women with depressive symptoms [28,30-33], post-traumatic stress symptoms [29], and pregnancy loss [46,47].

Therapist-supported iCBT may be particularly useful for postpartum depressive women who struggle with issues of stigma [30,32]. Therapist-supported iCBT helps women with child-related traumatic event to avoid their fearful impulses, thus helping them to decrease the level of fear experienced at a given moment, and eventually help to confront their fears [29]. In addition, therapist-supported iCBT evidently had great practical significance for women with painful memory of the pregnancy loss in order to improve stress, anxiety, and depressive symptoms [46,47].

Our meta-analytic results found no significant difference on anxiety and depressive symptoms using behavioral activation [4] compared with other forms of CBT. This result suggested that scheduling activities could behaviorally activate women and help them to gain a sense of pleasure and improve anxiety and depressive symptoms [4]. Contrary to previous findings on the equivalent effect of support type for stress symptoms in a meta-analytic review [13], our finding showed that exclusive therapist support has a better effect than therapist support with self-care component in the subgroup analysis. A possible interpretation of this result is that the different attrition rates reached 2.9%–60.8% by therapist support with self-care components [28,32,33] compared with the 7.0%–31.3% rates by exclusive therapist support [29–31,46,47]. This finding supported the fact that exclusive therapist support was preferred to reduce stress symptoms, whether by phone, email, text, or interactive computer interface [10,11]. Our review highlighted that support via website reminder or tracking [28], website contact system [29], or phone calls [33] increased retention to therapist-supported iCBT.

Efficacy of Therapist-Supported iCBT on Stress Symptoms

The finding of this meta-analysis provided support for the efficacy of therapist-supported iCBT on improving stress symptoms with large effect size of 0.84. This finding is in line with a recent meta-analysis that demonstrated a medium to large effect of 0.72–0.82 of iCBT for post-traumatic stress [13]. Similar components and strategies of iCBT possibly exist among a homogenous dataset with comparable health care conditions in both reviews. However, our result was contradictory with another review [38] indicating inconsistent effects of –0.32 to 0.98. One possible explanation for this difference might be linked to the different target population and health conditions between our review (postpartum women with pregnancy loss, depression, and post-traumatic stress) and those in the previous review (antenatal, postpartum women, and partners with general mental health, grief, stress, post-traumatic stress) [38]. Subgroup analyses revealed a trend for the waitlist comparator to have higher effect sizes than the TAU comparator for improving stress symptoms. The observation echoed a previous review that different control conditions led to substantively different effect estimates and that waitlist comparator generated larger effect sizes estimate than TAU comparator [57]. This finding could be explained by the possibility of additional stress during waiting period because those among waitlist comparator could not access intervention promptly, whereas those assigned to TAU might actively seek additional support that could possibly improve stress symptoms [57].

Efficacy of Therapist-Supported iCBT on Anxiety Symptoms

Therapist-supported iCBT revealed significantly improved anxiety symptoms with small effect size ($d=0.36$), which was different from the effect size of the three recent meta-analyses among children ($d=0.69$) [7], adults with anxiety disorders ($d=0.79$) [10], perinatal women with mixed results indicating both positive effect ($d=0.51$) and negative effect ($d=-0.61$) [38]. The discrepancy might be due to the difference population with specific health conditions between our review (postpartum women) and those in the previous three reviews (childhood and adults with anxiety disorders [7,10] or combination of antenatal, postpartum women, and their partners [38]). Another possible explanation was the substantial heterogeneity in the previous meta-analysis ($I^2=80\%$) [10] compared with homogeneity in our review ($I^2=0\%$). One possible explanation for the differences could be high normative levels of anxiety during the postpartum period [20] because of the conspicuous changes to the roles, lifestyles, and responsibilities of the expecting mother [17]. Directly comparing the current findings was impossible because the evidence was based on different populations with different health condition. This result encourages future investigations.

Efficacy of Therapist-Supported iCBT on Depressive Symptoms

Therapist-supported iCBT revealed significantly improved depressive symptoms with medium effect size of 0.63 in this review, which was better than the small to medium effect size of 0.41 [36] and 0.56 [37] among adult depression in the other two reviews. This result was consistent with a previous meta-analysis [38] that demonstrated a medium ($d=0.55$) to large ($d=1.03$) effect of intervention for depressive symptoms. The case might be linked to the number of includable trials, types of samples, and variety of therapies in the current review compared with those in the three previous reviews [36–38]. Our review used 8 trials, whereas the previous meta-analytic reviews employed 19 [37] and 12 [36] trials. These two reviews included very heterogeneous therapeutic treatments and samples with substantial to considerable heterogeneity ($I^2=57\%–81\%$) [36,37] compared with homogeneity ($I^2=0\%$) in our review. However, similar effect size between one previous meta-analysis [38] and our review might be explained by overlapping with 4 eligible studies [30,32,46,47] in both reviews. Since the previous review [38] did not report results of heterogeneity, it is hard to give a direct comparison. The number of includable trials is relatively few in the current review. Therefore, this aspect should be further explored in future meta-analytic reviews.

The number of includable studies for the subgroup analysis was low in this review, limiting statistical power. The comparisons were unbalanced. We did not find a significant difference between other subgroups, which could be caused by power problems. Further investigations are required.

Limitations

Several limitations exist in this review. First, a possible sampling bias was evident because the majority of the participants were self-selected on the basis of media-recruited participants rather

than clinical samples. Such recruitment methods often rely on the individual's motivation levels, which potentially correspond to slightly different demographics from those participants who are recruited within community settings. Second, majority of the outcomes were self-reported. Studies typically relied on self-report measures and rarely included formal diagnostic procedures at either recruitment or assessment. As a result, the health conditions or outcomes might be overestimated or underestimated. Third, relatively few trials were included in this review. Therefore, we were underpowered to detect effects for certain contrasts in subgroup analyses. Fourth, different sessions of iCBT tended to report different effect sizes. Fifth, only the short-term benefits of iCBT were investigated on the basis of postintervention outcome measures. The maintenance of benefits in iCBT remains unclear. Wide range of health conditions (especially for stress symptoms, including general and post-traumatic stress), high attrition rate, and broadly defined range of interventions were also considered as limitations of this review. Finally, this review only included studies published in English, and all were conducted in developed regions. The results may not be applicable to marginalized groups in developing regions.

Implications of Practice

This meta-analytic review extends the evidence of efficacy of iCBT on stress, anxiety, and depressive symptoms among postpartum women. As society becomes increasingly comfortable with, and reliant upon, the use of the Internet for routine health care, opportunities to apply iCBT tend to grow continuously [8]. The iCBT features distinct behavioral advice and learnable skills according to one's own pace without stigma, waiting times, or taking time off work [16]. Investigating differential predictors of outcomes for another therapeutic format is important. An enhanced understanding of the effective components is necessary to appropriately inform future evidence-based use of iCBT among postpartum women.

Implications for Future Research

Notably, the evidence of iCBT aims at preventive purpose that employs group-administrated method and exclusive self-help

approach among postpartum women are entirely lacking. Future research is welcome. Further investigation is needed on the relationships between the effects of iCBT and different periods of intervention (ie, antenatal vs postpartum), health conditions, and age groups to inform potential iCBT procedures. Participant characteristics may be directly or indirectly associated with attrition rate through a variety of mechanisms, such as literacy, familiarity with technology, personal preferences, and engagement maintenance. This review revealed that different supportive types might affect the different stress symptoms. Ironically, the concept of support is not fully investigated in the literature [36], and we cannot exclude the possibility that some forms of support can be automated. The concepts of support must be further investigated. Although the effects are superior in exclusive therapist support, determining the optimal levels of support is the next issue. The roles of therapist factors, different types of professionals, and support of professionals or non-professionals require further examination. Further rigorous trials based on clinical populations using accurate diagnostic assessment are required before therapist-supported iCBT can be deemed effective for the postpartum population in accessing regular clinical practice for improving stress, anxiety, and depressive symptoms.

Conclusions

This meta-analytic review supports the efficacy of therapist-supported iCBT for improving stress, anxiety, and depressive symptoms at post-treatment with small to large effects within a range between 0.36 and 0.84. A one-size-fits-all approach is unlikely to succeed considering the complexities and idiosyncrasies of specific health conditions. Future studies should establish the effective components, format, and approach of iCBT with optimal levels of human support. Adequate number of sessions and suitable duration during the postpartum period using appropriate functionality, interactivity, multimedia, and communication modes are important to maximize a long-term effect of iCBT among postpartum women.

Acknowledgments

This work was supported by a start-up grant (Ref: NUHSRO/2013/147/SU/01, WBS No.: R545-000-055-133 and WBS No.: R545-000-055-750) and HSS Seed Fund (1/2016) from the National University of Singapore.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Index and keyword terms for searching in ten databases.

[PDF File (Adobe PDF File), 42KB - [jmir_v19i4e138_app1.pdf](#)]

Multimedia Appendix 2

Risk of bias graph.

[PDF File (Adobe PDF File), 34KB - [jmir_v19i4e138_app2.pdf](#)]

Multimedia Appendix 3

Description of Internet-based cognitive behavioral therapy in 8 trials.

[PDF File (Adobe PDF File), 25KB - [jmir_v19i4e138_app3.pdf](#)]

Multimedia Appendix 4

Interaction of Internet-based cognitive behavioral therapy in 8 trials.

[PDF File (Adobe PDF File), 26KB - [jmir_v19i4e138_app4.pdf](#)]

References

1. Beck J. Cognitive Behavior Therapy: Basics and Beyond. 2nd edition. New York, NY: The Guilford Press; 2011.
2. Beck AT, Haigh EA. Advances in cognitive theory and therapy: the generic cognitive model. *Annu Rev Clin Psychol* 2014;10:1-24. [doi: [10.1146/annurev-clinpsy-032813-153734](#)] [Medline: [24387236](#)]
3. Hope DA, Burns JA, Hayes SA, Herbert JD, Warner MD. Automatic thoughts and cognitive restructuring in cognitive behavioral group therapy for social anxiety disorder. *Cogn Ther Res* 2007 Jun 2;34(1):1-12. [doi: [10.1007/s10608-007-9147-9](#)]
4. Addis M, Martell C. Overcoming Depression One Step at a Time: The New Behavioral Activation Approach to Getting Your Life Back. New York: New Harbinger Press; 2004.
5. Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. *Cognit Ther Res* 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: [10.1007/s10608-012-9476-1](#)] [Medline: [23459093](#)]
6. Paulson JF, Bazemore SD. Prenatal and postpartum depression in fathers and its association with maternal depression: a meta-analysis. *JAMA* 2010 May 19;303(19):1961-1969. [doi: [10.1001/jama.2010.605](#)] [Medline: [20483973](#)]
7. Rooksby M, Elouafkaoui P, Humphris G, Clarkson J, Freeman R. Internet-assisted delivery of cognitive behavioural therapy (CBT) for childhood anxiety: systematic review and meta-analysis. *J Anxiety Disord* 2015 Jan;29:83-92. [doi: [10.1016/j.janxdis.2014.11.006](#)] [Medline: [25527900](#)]
8. Andersson G. Internet-delivered psychological treatments. *Annu Rev Clin Psychol* 2016;12:157-179. [doi: [10.1146/annurev-clinpsy-021815-093006](#)] [Medline: [26652054](#)]
9. ITU. ITU world telecommunication /ICT indicators databases: statistical highlights URL: <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx> [WebCite Cache ID 6krBY0V4e]
10. Olthuis JV, Watt MC, Bailey K, Hayden JA, Stewart SH. Therapist-supported internet cognitive behavioural therapy for anxiety disorders in adults. *Cochrane Database Syst Rev* 2015(3):CD011565. [doi: [10.1002/14651858.CD011565](#)] [Medline: [25742186](#)]
11. Enander J, Andersson E, Mataix-Cols D, Lichtenstein L, Alström K, Andersson G, et al. Therapist guided internet based cognitive behavioural therapy for body dysmorphic disorder: single blind randomised controlled trial. *BMJ* 2016;352:i241 [FREE Full text] [Medline: [26837684](#)]
12. 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](#)]
13. Kuester A, Niemeyer H, Knaevelsrud C. Internet-based interventions for posttraumatic stress: a meta-analysis of randomized controlled trials. *Clin Psychol Rev* 2016 Feb;43:1-16. [doi: [10.1016/j.cpr.2015.11.004](#)] [Medline: [26655959](#)]
14. Kelders SM, Bohlmeijer ET, Pots WT, van Gemert-Pijnen JE. Comparing human and automated support for depression: fractional factorial randomized controlled trial. *Behav Res Ther* 2015 Sep;72:72-80. [doi: [10.1016/j.brat.2015.06.014](#)] [Medline: [26196078](#)]
15. Hedman E, Andersson E, Ljótsson B, Axelsson E, Lekander M. Cost effectiveness of internet-based cognitive behaviour therapy and behavioural stress management for severe health anxiety. *BMJ Open* 2016;6(4):e009327 [FREE Full text] [doi: [10.1136/bmjopen-2015-009327](#)] [Medline: [27113231](#)]
16. 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](#)]
17. Agius A, Xuereb RB, Carrick-Sen D, Sultana R, Rankin J. The co-existence of depression, anxiety and post-traumatic stress symptoms in the perinatal period: a systematic review. *Midwifery* 2016 May;36:70-79. [doi: [10.1016/j.midw.2016.02.013](#)] [Medline: [27106946](#)]
18. Clout D, Brown R. Marital relationship and attachment predictors of postpartum stress, anxiety, and depression symptoms. *J Soc Clin Psychol* 2016 Apr;35(4):322-341. [doi: [10.1521/jscp.2016.35.4.322](#)]
19. Grekin R, O'Hara MW. Prevalence and risk factors of postpartum posttraumatic stress disorder: a meta-analysis. *Clin Psychol Rev* 2014 Jul;34(5):389-401. [doi: [10.1016/j.cpr.2014.05.003](#)] [Medline: [24952134](#)]

20. Osma J, Barrera AZ, Ramphos E. Are pregnant and postpartum women interested in health-related apps? implications for the prevention of perinatal depression. *Cyberpsychol Behav Soc Netw* 2016 Jun;19(6):412-415. [doi: [10.1089/cyber.2015.0549](https://doi.org/10.1089/cyber.2015.0549)] [Medline: [27327069](https://pubmed.ncbi.nlm.nih.gov/27327069/)]
21. O'Hara MW, McCabe JE. Postpartum depression: current status and future directions. *Annu Rev Clin Psychol* 2013;9:379-407. [doi: [10.1146/annurev-clinpsy-050212-185612](https://doi.org/10.1146/annurev-clinpsy-050212-185612)] [Medline: [23394227](https://pubmed.ncbi.nlm.nih.gov/23394227/)]
22. Sanger C, Iles JE, Andrew CS, Ramchandani PG. Associations between postnatal maternal depression and psychological outcomes in adolescent offspring: a systematic review. *Arch Womens Ment Health* 2015 Apr;18(2):147-162. [doi: [10.1007/s00737-014-0463-2](https://doi.org/10.1007/s00737-014-0463-2)] [Medline: [25269760](https://pubmed.ncbi.nlm.nih.gov/25269760/)]
23. Grace SL, Evindar A, Stewart DE. The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature. *Arch Womens Ment Health* 2003 Nov;6(4):263-274. [doi: [10.1007/s00737-003-0024-6](https://doi.org/10.1007/s00737-003-0024-6)] [Medline: [14628179](https://pubmed.ncbi.nlm.nih.gov/14628179/)]
24. de Camps MD, Philipp D, Israel A, Vigod S. Maternal-infant mental health: postpartum group intervention. *Arch Womens Ment Health* 2016 Apr;19(2):243-251. [doi: [10.1007/s00737-015-0551-y](https://doi.org/10.1007/s00737-015-0551-y)] [Medline: [26239582](https://pubmed.ncbi.nlm.nih.gov/26239582/)]
25. Sockol LE. A systematic review of the efficacy of cognitive behavioral therapy for treating and preventing perinatal depression. *J Affect Disord* 2015 May 15;177:7-21. [doi: [10.1016/j.jad.2015.01.052](https://doi.org/10.1016/j.jad.2015.01.052)] [Medline: [25743368](https://pubmed.ncbi.nlm.nih.gov/25743368/)]
26. Maloni JA, Przeworski A, Damato EG. Web recruitment and internet use and preferences reported by women with postpartum depression after pregnancy complications. *Arch Psychiatr Nurs* 2013 Apr;27(2):90-95. [doi: [10.1016/j.apnu.2012.12.001](https://doi.org/10.1016/j.apnu.2012.12.001)] [Medline: [23540519](https://pubmed.ncbi.nlm.nih.gov/23540519/)]
27. Fonseca A, Gorayeb R, Canavarro M. Women's use of online resources and acceptance of e-mental health tools during the perinatal period. *Int J Med Inform* 2016 Oct;94:228-236. [doi: [10.1016/j.ijmedinf.2016.07.016](https://doi.org/10.1016/j.ijmedinf.2016.07.016)] [Medline: [27573331](https://pubmed.ncbi.nlm.nih.gov/27573331/)]
28. Milgrom J, Danaher BG, Gemmill AW, Holt C, Holt CJ, Seeley JR, et al. Internet cognitive behavioral therapy for women with postnatal depression: a randomized controlled trial of MumMoodBooster. *J Med Internet Res* 2016;18(3):e54 [FREE Full text] [doi: [10.2196/jmir.4993](https://doi.org/10.2196/jmir.4993)] [Medline: [26952645](https://pubmed.ncbi.nlm.nih.gov/26952645/)]
29. Nieminen K, Berg I, Frankenstein K, Viita L, Larsson K, Persson U, et al. Internet-provided cognitive behaviour therapy of posttraumatic stress symptoms following childbirth-a randomized controlled trial. *Cogn Behav Ther* 2016 Jun;45(4):287-306. [doi: [10.1080/16506073.2016.1169626](https://doi.org/10.1080/16506073.2016.1169626)] [Medline: [27152849](https://pubmed.ncbi.nlm.nih.gov/27152849/)]
30. O'Mahen HA, Richards DA, Woodford J, Wilkinson E, McGinley J, Taylor RS, et al. Netmums: a phase II randomized controlled trial of a guided Internet behavioural activation treatment for postpartum depression. *Psychol Med* 2014 Jun;44(8):1675-1689 [FREE Full text] [doi: [10.1017/S0033291713002092](https://doi.org/10.1017/S0033291713002092)] [Medline: [24148703](https://pubmed.ncbi.nlm.nih.gov/24148703/)]
31. Pugh NE, Hadjistavropoulos HD, Dirkse D. A randomised controlled trial of therapist-assisted, internet-delivered cognitive behavior therapy for women with maternal depression. *PLoS One* 2016;11(3):e0149186 [FREE Full text] [doi: [10.1371/journal.pone.0149186](https://doi.org/10.1371/journal.pone.0149186)] [Medline: [26930488](https://pubmed.ncbi.nlm.nih.gov/26930488/)]
32. O'Mahen HA, Woodford J, McGinley J, Warren FC, Richards DA, Lynch TR, et al. Internet-based behavioral activation--treatment for postnatal depression (Netmums): a randomized controlled trial. *J Affect Disord* 2013 Sep 25;150(3):814-822. [doi: [10.1016/j.jad.2013.03.005](https://doi.org/10.1016/j.jad.2013.03.005)] [Medline: [23602514](https://pubmed.ncbi.nlm.nih.gov/23602514/)]
33. Sheeber LB, Seeley JR, Feil EG, Davis B, Sorensen E, Kosty DB, et al. Development and pilot evaluation of an internet-facilitated cognitive-behavioral intervention for maternal depression. *J Consult Clin Psychol* 2012 Oct;80(5):739-749 [FREE Full text] [doi: [10.1037/a0028820](https://doi.org/10.1037/a0028820)] [Medline: [22663903](https://pubmed.ncbi.nlm.nih.gov/22663903/)]
34. Carolan-Olah M, Barry M. Antenatal stress: an Irish case study. *Midwifery* 2014 Mar;30(3):310-316. [doi: [10.1016/j.midw.2013.03.014](https://doi.org/10.1016/j.midw.2013.03.014)] [Medline: [23684696](https://pubmed.ncbi.nlm.nih.gov/23684696/)]
35. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. *Clin Psychol Rev* 2009 Jul;29(5):410-420 [FREE Full text] [doi: [10.1016/j.cpr.2009.04.001](https://doi.org/10.1016/j.cpr.2009.04.001)] [Medline: [19403215](https://pubmed.ncbi.nlm.nih.gov/19403215/)]
36. 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](https://doi.org/10.1080/16506070903318960)] [Medline: [20183695](https://pubmed.ncbi.nlm.nih.gov/20183695/)]
37. 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](https://doi.org/10.1016/j.cpr.2012.02.004)] [Medline: [22466510](https://pubmed.ncbi.nlm.nih.gov/22466510/)]
38. Ashford MT, Olander EK, Ayers S. Computer- or web-based interventions for perinatal mental health: a systematic review. *J Affect Disord* 2016 Jun;197:134-146. [doi: [10.1016/j.jad.2016.02.057](https://doi.org/10.1016/j.jad.2016.02.057)] [Medline: [26991368](https://pubmed.ncbi.nlm.nih.gov/26991368/)]
39. Lee EW, Denison FC, Hor K, Reynolds RM. Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review. *BMC Pregnancy Childbirth* 2016;16:38 [FREE Full text] [doi: [10.1186/s12884-016-0831-1](https://doi.org/10.1186/s12884-016-0831-1)] [Medline: [26928898](https://pubmed.ncbi.nlm.nih.gov/26928898/)]
40. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009 Oct;62(10):1006-1012. [doi: [10.1016/j.jclinepi.2009.06.005](https://doi.org/10.1016/j.jclinepi.2009.06.005)] [Medline: [19631508](https://pubmed.ncbi.nlm.nih.gov/19631508/)]
41. Neville CE, McKinley MC, Holmes VA, Spence D, Woodside JV. The effectiveness of weight management interventions in breastfeeding women--a systematic review and critical evaluation. *Birth* 2014 Sep;41(3):223-236. [doi: [10.1111/birt.12111](https://doi.org/10.1111/birt.12111)] [Medline: [24750419](https://pubmed.ncbi.nlm.nih.gov/24750419/)]
42. Higgins JP, Green S. Cochrane. Cochrane handbook for systematic reviews of interventions, Version 5.1.0 URL: <http://handbook.cochrane.org/> [accessed 2017-03-01] [WebCite Cache ID 6oeYb6Ks9]

43. Meades R, Ayers S. Anxiety measures validated in perinatal populations: a systematic review. *J Affect Disord* 2011 Sep;133(1-2):1-15. [doi: [10.1016/j.jad.2010.10.009](https://doi.org/10.1016/j.jad.2010.10.009)] [Medline: [21078523](https://pubmed.ncbi.nlm.nih.gov/21078523/)]
44. Borenstein M, Hedges LV, Rothstein H. *Introduction to Meta-Analysis*. Oxford: Wiley; 2009.
45. Sormani MP. Subgroup analysis in MS trials. *Mult Scler* 2017 Jan;23(1):34-35. [doi: [10.1177/1352458515625808](https://doi.org/10.1177/1352458515625808)] [Medline: [26772964](https://pubmed.ncbi.nlm.nih.gov/26772964/)]
46. Kersting A, Dölemeyer R, Steinig J, Walter F, Kroker K, Baust K, et al. Brief internet-based intervention reduces posttraumatic stress and prolonged grief in parents after the loss of a child during pregnancy: a randomized controlled trial. *Psychother Psychosom* 2013;82(6):372-381. [doi: [10.1159/000348713](https://doi.org/10.1159/000348713)] [Medline: [24061387](https://pubmed.ncbi.nlm.nih.gov/24061387/)]
47. Kersting A, Kroker K, Schlicht S, Baust K, Wagner B. Efficacy of cognitive behavioral internet-based therapy in parents after the loss of a child during pregnancy: pilot data from a randomized controlled trial. *Arch Womens Ment Health* 2011 Dec;14(6):465-477. [doi: [10.1007/s00737-011-0240-4](https://doi.org/10.1007/s00737-011-0240-4)] [Medline: [22006106](https://pubmed.ncbi.nlm.nih.gov/22006106/)]
48. Asselmann E, Wittchen H, Erler L, Martini J. Peripartum changes in social support among women with and without anxiety and depressive disorders prior to pregnancy: a prospective-longitudinal study. *Arch Womens Ment Health* 2016 Dec;19(6):943-952. [doi: [10.1007/s00737-016-0608-6](https://doi.org/10.1007/s00737-016-0608-6)] [Medline: [26846662](https://pubmed.ncbi.nlm.nih.gov/26846662/)]
49. Asselmann E, Wittchen H, Petzoldt J, Martini J. Peripartum changes in partnership quality among women with and without anxiety and depressive disorders prior to pregnancy: a prospective-longitudinal study. *Arch Womens Ment Health* 2016 Apr;19(2):281-290. [doi: [10.1007/s00737-015-0556-6](https://doi.org/10.1007/s00737-015-0556-6)] [Medline: [26264505](https://pubmed.ncbi.nlm.nih.gov/26264505/)]
50. Weiss DS. The impact of event scale-revised. In: Wilson JP, Tang CS, editors. *Cross-Cultural Assessment of Psychological Trauma and PTSD*. Boston, MA: Springer US; 2007:219-238.
51. Martini J, Petzoldt J, Einsle F, Beesdo-Baum K, Höfler M, Wittchen H. Risk factors and course patterns of anxiety and depressive disorders during pregnancy and after delivery: a prospective-longitudinal study. *J Affect Disord* 2015 Apr 1;175:385-395. [doi: [10.1016/j.jad.2015.01.012](https://doi.org/10.1016/j.jad.2015.01.012)] [Medline: [25678171](https://pubmed.ncbi.nlm.nih.gov/25678171/)]
52. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
53. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988 Dec;56(6):893-897. [Medline: [3204199](https://pubmed.ncbi.nlm.nih.gov/3204199/)]
54. Winkel S, Einsle F, Pieper L, Höfler M, Wittchen H, Martini J. Associations of anxiety disorders, depressive disorders and body weight with hypertension during pregnancy. *Arch Womens Ment Health* 2015 Jun;18(3):473-483. [doi: [10.1007/s00737-014-0474-z](https://doi.org/10.1007/s00737-014-0474-z)] [Medline: [25422149](https://pubmed.ncbi.nlm.nih.gov/25422149/)]
55. Sutter-Dallay AL, Giaconne-Marcasche V, Glatigny-Dallay E, Verdoux H. Women with anxiety disorders during pregnancy are at increased risk of intense postnatal depressive symptoms: a prospective survey of the MATQUID cohort. *Eur Psychiatry* 2004 Dec;19(8):459-463. [doi: [10.1016/j.eurpsy.2004.09.025](https://doi.org/10.1016/j.eurpsy.2004.09.025)] [Medline: [15589703](https://pubmed.ncbi.nlm.nih.gov/15589703/)]
56. Alvarado R, Perucca E, Neves E, Rojas M, Monardes J, Olea E, et al. Depressive disorders during pregnancy and associated factors. *Rev Chil Obstet Ginecol* 1993;58(2):135-141. [Medline: [8209041](https://pubmed.ncbi.nlm.nih.gov/8209041/)]
57. Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, et al. Waiting list may be a placebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatr Scand* 2014 Sep;130(3):181-192. [doi: [10.1111/acps.12275](https://doi.org/10.1111/acps.12275)] [Medline: [24697518](https://pubmed.ncbi.nlm.nih.gov/24697518/)]

Abbreviations

BSI: Brief Symptom Inventory
CBT: Cognitive Behavior Therapy
DASS: Depression Anxiety Stress Scale
iCBT: Internet-Based Cognitive Behavior Therapy
IES: Impact of Event Scale
IES-R: Impact of Event Scale-revised
ITT: Intention-to-Treat
IV: inverse variance
PICOS: Patient/Problem Intervention Comparison Outcome Setting
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analysis
PSS: Perceived Stress Scale
RCT: randomized controlled trial
TAU: treatment as usual
WHO: World Health Organization

Edited by G Eysenbach; submitted 30.09.16; peer-reviewed by P Zelkowitz, M Ashford, TR Soron; comments to author 29.12.16; revised version received 26.01.17; accepted 13.02.17; published 28.04.17.

Please cite as:

Lau Y, Htun TP, Wong SN, Tam WSW, Klainin-Yobas P

Therapist-Supported Internet-Based Cognitive Behavior Therapy for Stress, Anxiety, and Depressive Symptoms Among Postpartum Women: A Systematic Review and Meta-Analysis

J Med Internet Res 2017;19(4):e138

URL: <http://www.jmir.org/2017/4/e138/>

doi: [10.2196/jmir.6712](https://doi.org/10.2196/jmir.6712)

PMID: [28455276](https://pubmed.ncbi.nlm.nih.gov/28455276/)

©Ying Lau, Tha Pyai Htun, Sui Nee Wong, Wai San Wilson Tam, Piyanee Klainin-Yobas. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder

Jennifer Nicholas^{1,2}, BSc (hons), BA; Andrea S Fogarty¹, BPsych(Hons), MPH, PhD; Katherine Boydell¹, PhD; Helen Christensen¹, PhD

¹Black Dog Institute, University of New South Wales, Sydney, Australia

²School of Psychiatry, Faculty of Medicine, University of New South Wales, Sydney, Australia

Corresponding Author:

Jennifer Nicholas, BSc (hons), BA
Black Dog Institute
University of New South Wales
Hospital rd, Prince of Wales Hospital
Sydney, 2031
Australia
Phone: 61 293828507
Fax: 61 293828208
Email: J.nicholas@blackdog.org.au

Abstract

Background: The delivery of mobile health (mHealth) services is acceptable to mental health consumers. However, despite the benefits of accessibility, cost-effectiveness, anonymity, and ability to tailor content to individual needs, consumer engagement remains a hurdle for uptake and continued use. This may be unsurprising as few studies have examined app content from the consumer perspective or assessed consumer preferences for the content of apps for mental health management. An opportunity to examine consumer perspectives exists in using naturally generated data that is publically available in the Google Play and Apple app stores. Whereas commercial developers routinely use this data, to date there has been no in-depth evaluation within scientific research.

Objective: The aim of our study was to explore what consumers consider useful content for mental health management apps, identify unmet needs, and understand user expectations of mental health apps within the context of apps for bipolar disorder.

Methods: Publically available English language consumer reviews of 48 apps for bipolar disorder were used as data, providing a total of 2173 reviews. Review text was coded and analyzed using a team approach to qualitative content analysis. Results were presented in 2 forms: (1) a quantitative summary of the 9 major and minor themes and (2) a qualitative synthesis of key thematic findings.

Results: The majority of reviews were for symptom monitoring apps (87.94%, 1911/2173). The qualitative content analysis revealed 5 main themes: (1) laudatory talk, comments regarding the app's benefits including helpfulness and successful design features (74.00% of reviews, 1608/2173); (2) unfavorable feedback, negative reviews largely concerning unmet needs, privacy and technical issues, and potential dangers of app use (25.54%, 555/2173); (3) conceptions of community, referring to both communities of users with mental ill-health accessed via the app and a community created among app users and developers (24.25%, 527/2173); (4) wishlist features, app features requested by users (17.53%, 381/2173); and (5) apps and therapy, app use within clinical care (10.58%, 230/2173). Four minor themes were also identified: (1) app cost, (2) privacy and data security, (3) comparisons with traditional monitoring, and (4) evidence-based mHealth.

Conclusions: Although mostly positive, the proportion of reviews containing wishlist requests indicates consumer needs are not adequately addressed by currently available disorder management apps. Consumers value content that is helpful, supportive, and easy to use, and they are integrating apps into their health management and clinical care without necessarily considering the evidence-base or clinical effectiveness of the tool. User expectations regarding developer responsiveness to their needs has implications for community-based participatory research and integrated knowledge translation. However, this expectation is incompatible with current mHealth funding structures.

(*J Med Internet Res* 2017;19(4):e105) doi:[10.2196/jmir.7273](https://doi.org/10.2196/jmir.7273)

KEYWORDS

mobile applications; bipolar disorder; smartphone; telemedicine; qualitative research

Introduction

The delivery of mobile health (mHealth) services is acceptable to mental health consumers [1,2]. Among the advantages of mHealth cited by consumers are convenience, the ability to identify triggers to mood states, and reducing isolation [1]. Furthermore, preliminary research indicates that the use of mHealth to facilitate symptom monitoring is acceptable to individuals with bipolar disorder [3]. Harnessing these qualities, smartphone apps have the potential to provide a platform for intervention and support for mental health conditions. Since the debut of the app store in 2008, the number of health and mental health apps available has risen exponentially [4]. However, no research has examined consumer perceived usefulness of apps for bipolar disorder management. This is despite the fact that apps may be uniquely suited to the condition; as active self-management is critical for bipolar disorder, with individuals reporting symptom monitoring and personal pattern identification as essential to disorder management [5,6].

A recent assessment of such mHealth tools developed for the disorder [7] found none demonstrated efficacy in reducing symptoms, preventing relapse, or facilitating disorder management. Moreover, few apps provided information grounded in evidence-based practice or adhered to clinical resources for disorder management. This has been further noted in mHealth research generally [8], with poor integration of evidence-based recommendations across apps for health [9,10] and mental health [11-13] conditions.

However, research beyond app evidence-base or efficacy is scant. Few studies have examined mental health apps from the consumer perspective or assessed consumer preferences regarding app content specifically for disorder management. This lack of attention to the consumer perspective may in part contribute to the modest app use statistics reported by several industry surveys. A US consumer survey found that 26% of health apps are abandoned after first use and overall 74% are discontinued within 10 uses [14], with engagement, finding a better app, and usability the main reasons for discontinuation. Similar research suggests only 5% of apps maintain continued use over a month after download [15].

Therefore, while delivery of resources and interventions using mobile technology employs the benefits of accessibility, cost-effectiveness, anonymity, and ability to tailor content to individual needs, consumer engagement remains a hurdle for uptake and continued use. An examination of consumer's perspectives of apps is therefore a critical first step to improving the utility of such resources. The opportunity already exists to examine consumer perspectives using naturally generated data that is publically available in the Google Play and Apple app stores. App users provide information about features they value, need, and dislike when they choose to review the apps they use. App developers ensure they are meeting user expectations by examining review content [16] but to date there has been no comprehensive evaluation within scientific research.

In this paper, we perform a qualitative analysis of user reviews in order to (1) explore what app content users consider useful for bipolar disorder management and identify unmet needs, (2) understand user expectations of disorder management apps, (3) examine how apps for the disorder are being used, and (4) determine if users' perspectives on current apps are consistent with scientific evaluation. Using qualitative methodology, we examine the main themes within user reviews and discuss the implications of these perspectives for app development and clinical practice.

Methods

Data Collection

The study used publicly available data. User review inclusion was restricted to the reviews of 82 apps previously investigated in a study of app content [7]. Detailed methods of app selection and assessment are outlined in Nicholas et al [7]; in brief, apps were identified by a search of the Google Play and Apple app stores using the following bipolar related search terms: bipolar, bipolar, "manic depression," "mood swings," "mania" and "mood," cyclothymia, and cyclothymic. Apps with descriptions stating they were useful for the disorder, targeted at consumers or carers, and available in English were included in the study. All English language reviews submitted up to December 31, 2015 for the 82 apps included in Nicholas et al [7] were included in the analysis.

Android reviews were collected using a data extraction script, which included the text of the review, star rating given, and review date, as well as app details including name, package ID, and cost. Apple apps reviews were transcribed by hand, extracting the same information. Although reviewers were often logged into an Apple or Google Play account when reviewing an app, to maintain anonymity, no data about the review author was collected. For this reason, direct quotes within text and illustrative examples in [Multimedia Appendix 1](#) have been minimally amended to preserve this anonymity, without changing the intent of the review. Reviews were transcribed or exported into Excel, cleaned and formatted, and then imported into the qualitative data management program NVivo11 (QSR International 2016) for analysis. The study received ethical approval from the UNSW Australia Human Research Ethics Advisory (Protocol number: HC14358).

Data Analysis

App and Review Characteristics

Categorization of app function (eg, symptom monitoring) followed categories assigned to apps in Nicholas et al [7], which were determined by an extensive analysis of app content. Descriptive statistics were used to detail app price, number of reviewed apps, and number of reviews per app platform and app function. Average user-star rating by themes was calculated using NVivo11.

Qualitative Analysis

Data from the reviews were described using qualitative content analysis techniques [17,18], following a “conventional” approach, as is indicated for fields of research where existing theory is limited [17,18]. Content analysis was chosen for its strengths in systematically categorizing and summarizing large volumes of text-based data, and the ability to assist in interpreting patterns occurring in the text, with attention given to the context from which sample data is drawn [19]. We used a deliberately broad operationalization of qualitative content analysis, in line with previous research [17,20,21] which asserts that data derived from content analyses can be reported qualitatively without presenting full “counts” or statistical analysis. However, in recognition of more traditional and narrowly defined approaches to content analysis [22,23], we do present a brief quantitative summary of key features in the data.

Our approach to the analysis followed established coding techniques [24] over 3 broad phases: (1) immersion in the data, (2) reduction of the data through systematic coding and generation of themes, and (3) interpretation of the findings. Given the paucity of content analyses of consumer perspectives on app technology, the analysis used an inductive approach to developing a coding framework [25,26].

A preliminary sample of 10 pages of reviews was randomly selected to ensure adequate coverage across the dataset. Three coders (KB, JN, and AF) immersed themselves in the data by reading and rereading extracted reviews, followed by independently generating first stage concepts, or “main ideas” detected in the text. These main ideas were then compared among all 3 coders to check for similarities and differences before tentative agreement was reached regarding appropriate coding categories for the main ideas. These main ideas formed the basis for the quantitative component of the content analysis and are reported in the Results section as “major themes” if they were present in more than 10% of all reviews, or as “minor themes” where they were not [27]. Given the volume of data generated by over 2000 reviews, quantitative reporting of the content analysis was restricted to these major and minor themes.

Two coders (JN and AF) then selected another sample of 60 reviews to code in common, and 60 reviews to code independently. At this stage, the previously identified major and minor themes were tested and potential subthemes relevant to each of the major themes were identified. The team then met to discuss and agree upon a final coding framework to be applied

to the remaining data. Detailed descriptions of the microlevel codes related to major, minor, and subthemes were generated in consultation with the study team. All remaining reviews were then evenly divided and coded line-by-line by 2 coders (JN and AF), who regularly checked in to resolve any coding disputes or discuss new codes detected in the data. Upon completion of coding, the coding framework was then applied to the original samples, which were then reintegrated into the dataset. Coders then met to refine reporting of results, including synthesizing information coded under each of the identified subthemes. Thus, results of the content analysis are reported as both a broad quantitative summary of the 9 major and minor themes [28] and a more descriptive qualitative summary of key findings in the subthemes [17].

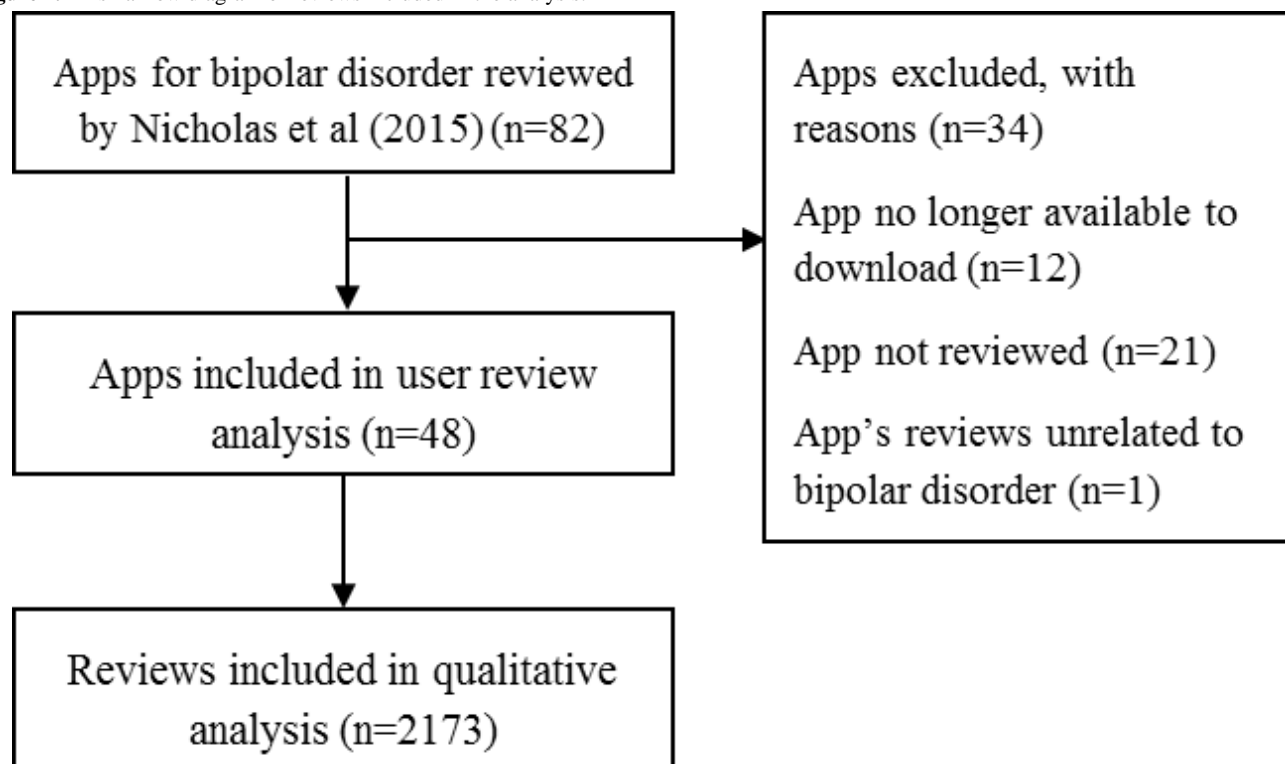
Research Rigor

We attended to research rigor in multiple ways, with specific emphasis on prolonged engagement with the data [29,30]. In our qualitative reporting, we established confirmability of results via consistent team debriefing related to the description and definition of themes and subthemes [30]. A second check involved both coders independently synthesizing key findings in the same subthemes, before comparing results. This process was repeated for three themes, before both coders developed an acceptable level of similarity in their approach. The remaining subthemes were then divided evenly between the 2 primary coders to report. For our quantitative reporting, intercoder reliability was assessed by using a random number generator to select 217 (10.00%, 217/2173) reviews to detect kappa with a null value of .40 at 90% power [31]. Each coder assessed the major and minor themes in each review and agreement between coders was calculated.

Results

Sample

Of the 82 eligible apps, consumer reviews of 48 apps were included in the analysis; 37 Android apps and 11 iOS (see [Multimedia Appendix 2](#)). Of the 34 excluded apps, 12 were no longer available for download, 21 had no reviews, and one app’s reviews were excluded as the app related to general health and reviews were unrelated to bipolar disorder. In total, users reviewed the 48 apps 2173 times ([Figure 1](#)). Reviews were submitted over a 10-year period, from July 13, 2005 to December 31, 2015, with the majority submitted after January 1, 2011.

Figure 1. Prisma flow diagram of reviews included in the analysis.

Description of Apps and Reviews

Thirty-six (75%, 36/48) of the included apps were free, and the average cost of paid apps was Aus \$4.37. Just over half (52%, 25/48) were categorized as symptom monitoring, which accounted for 87.94% (1911/2173) of total reviews. Thirteen apps provided information about the disorder and had 122 reviews (5.61%, 122/2173). Very few apps and reviews related to screening, community support, or treatment provision (See [Table 1](#)). The majority of reviews (89.46%, 1944/2173) were for Android apps.

Five major and four minor (present in less than 10% of reviews) themes were identified in the data ([Table 2](#)). [Table 2](#) also reports the percentage of reviews rated over 3 stars for each theme. Examples of subthemes identified within major themes and minor themes are displayed in [Multimedia Appendix 1](#). Pertinent subthemes are discussed in further detail and were selected where results related to pragmatic implications for: clinical practice, design and development of app technology, and future research.

Table 1. Number of apps and reviews by app function.

App function	Number of apps with reviews (n=48)	Reviews (N=2173), n (%)	Average reviews per app
Symptom monitoring	25	1911 (87.94)	73.50
Information	13	122 (5.61)	9.38
Screening and assessment	6	73 (3.36)	12.17
Community support	3	63 (2.90)	21
Treatment	1	4 (0.18)	4

Table 2. Prevalence of major and minor themes identified in the data (categories not exclusive; N=2173).

Themes	Reviews, n (%)	Reviews above 3 star rating, n (%)
Major themes		
1. Laudatory talk	1608 (74.00)	1501 (93.35)
2. Unfavorable feedback	555 (25.54)	150 (27.03)
3. Community	527 (24.25)	411 (77.99)
4. Wishlist	381 (17.53)	295 (77.43)
5. Apps and therapy	230 (10.58)	214 (93.04)
Minor themes		
1. App cost	149 (6.86)	78 (52.35)
2. Privacy and data security	138 (6.35)	77 (55.80)
3. Comparisons with traditional monitoring	26 (1.20)	25 (96.15)
4. Evidence-based mHealth	5 (0.23)	5 (100.00)

Research Rigor

Intercoder reliability for the 9 major and minor themes was strong, with almost perfect agreement for the themes laudatory talk ($\kappa=.838$, $P<.001$, 95% CI 0.736-0.940), unfavorable feedback ($\kappa=.882$, $P<.001$, 95% CI 0.808-0.956), and app cost ($\kappa=.85$, $P<.001$, 95% CI 0.683-1.017). Substantial intercoder agreement was observed for 3 major themes: wishlist ($\kappa=.741$, $P<.001$, 95% CI 0.629-0.853), apps and therapy ($\kappa=.714$, $P<.001$, 95% CI 0.528-0.900), and privacy and data security ($\kappa=.701$, $P<.001$, 95% CI 0.593-0.989). There were moderate levels of intercoder agreement for community ($\kappa=.493$, $P<.001$, 95% CI 0.350-0.636). Perfect agreement was observed for comparison with traditional monitoring ($\kappa=1$, $P<.001$), and both coders agreed that none of the randomly selected reviews contained any reference to the evidence-based mHealth theme.

Major Theme 1: Laudatory Talk

Almost two thirds of reviews featured positive commentary about the app. However, positive reviews often included a contrasting statement, most commonly a request for additional features, for example, “*I love this, but...I want the ability to type longer notes.*” Positive reviews were often general in nature and did not provide specific details regarding which aspects of the app were highly valued, though “ease of use” heavily featured. Two main laudatory subthemes are discussed: (1.1) helpfulness and (1.2) successful design.

Subtheme 1.1: Helpfulness

Many positive reviews indicated that the app was helpful, but did not specify exactly how. Of the reviews that did elaborate, the app assisted in two interrelated areas—keeping track of and gaining insight into moods. Whereas mood was the main symptom monitored, reviewers used the apps to monitor a range of factors including sleep, medication, and episode triggers. Gaining insight into mood was often attributed to the app’s ability to provide an increased understanding of affect changes, including identification of mood triggers or establishing duration of and variations in mood. For some, this insight was otherwise occluded by mood state, for example, “*It can be hard to see*

when you’re in it.” A few users spoke about gaining control of their mood through app use and identified that understanding and insight brought power and self-efficacy, for example, “*This helps me predict and deal with mood changes. I can prepare for down moods and reassure myself that these feelings are normal for me at the time and that it will pass.*”

Subtheme 1.2: Successful Design

Many reviews were complimentary of the app’s features or overall design. A commonly described positive app feature was ease of use—simple, easy, intuitive, and quick to learn, for example, “*I like this app. It’s easy to use, and that means I’m more likely to use it.*” When specific app features were praised, these were similar to features requested within major theme 4, wishlist (shown in Table 3). Most commonly praised features were symptom monitoring options, provision of graphs and analysis, reminders, and the app’s interface.

Major Theme 2: Unfavorable Feedback

Negative comments were present in just over a quarter of reviews and mostly concerned overall app design and functionality. Here, general comments were less common, with negative reviews concentrating on 3 key areas: (2.1) app content or features not meeting user needs; (2.2) deal breakers, features that prompted users to cease using an app, or technical issues that interfered with app use; and less commonly, (2.3) the potential for apps to precipitate or worsen distress.

Subtheme 2.1: Features Not Meeting User Needs

A recurring complaint concerned the lack of fit between user experience of symptoms and available features. This was particularly evident where reviewers self-identified as having “rapid-cycling bipolar disorder” and reported mood fluctuations throughout the day, not adequately captured by once-daily monitoring options. Besides wanting to track more frequently, many also wanted to select from a more extensive range of feelings, symptoms, or behaviors, for example, “*There is a very very small mood selection. Just moods like ‘happy, sad, depressed’ nothing like ‘disappointed, irritated, proud’ which are important.*” Similarly, though not as common, were complaints about the inability to customize monitoring scales,

for example, “Grr. The scale goes 0-10, but instead of 0 being very anxious and 10 being very calm, its 0-10 anxiety...with the middle marked as ‘normal’.”

Subtheme 2.2: “Deal Breakers” and Technical Issues

The aspects of apps considered deal-breakers were repeatedly mentioned. The majority of “deal breakers” involved concerns about privacy or technology failure. Users also discontinued use when available options were inadequate, the app used too much data or processing power, or when tasks were too burdensome, for example, “Too complicated to do when sad.” Occasionally, reviews referred to inaccurate or unclear app content, lack of instructions, difficulty learning phone gestures (eg, tap and hold), or advertising as reasons for discontinuing use.

Similarly, technical issues that interfered with app use featured prominently in negative reviews. For example, inactive export features; apps that crashed, lost data, froze, did not load content, or would not download; apps that were slow, unresponsive, or unavailable offline; apps without developer support; and apps that were buggy or “glitchy.”

Subtheme 2.3: Potential to Precipitate Distress

Less commonly, reviewers were disparaging of apps they thought could potentially cause or exacerbate distress via provision of poor information or failure to consider vulnerable users. Some noted that app developers lacked necessary clinical expertise, for example, “This person is a quack not a dr...unsafe and irresponsible.” Some did not specify the perceived danger but noted the app had the capacity to harm help-seekers by attracting them with promises that the app failed to deliver. Less commonly, users described feeling more distressed or frustrated after engaging with the app, for example, “Of course it’s because of mood swings that one gets a mood app. So imagine what happens when the first time trying to open the app it doesn’t work!!!”

Major Theme 3: Community

About a quarter of all reviews referred to feeling part of a “community” that extended well beyond the apps. Generally, community referred to (3.1) a wide interactive community of app users (actual or potential) and developers or (3.2) a Web-based community accessed via the app.

Subtheme 3.1: An Interactive Community of App Users and Developers

Within the app use community, reviews often directly addressed app developers, most commonly to thank, encourage, or make requests. At times, requests were “incentivized” with reviewers promising improved star ratings or donations in return for added features or fixes. There was also an emphasis on developer

responsiveness, with reviewers pleased when concerns raised with developers were addressed. Less frequently, reviewers expressed confusion over the apps intended function, or questioned the developers’ expertise in creating a mental health app.

The community of other app users was referenced during interactions with developers, to support and backup requests, for example, “My manic brothers understand the need to write paragraphs at times. Please fix this.” This sometimes extended to referring to another user’s review as evidence of a shared need. The community was not restricted to those with bipolar disorder, extending to any potential users, with many reviews including opinions about the apps usefulness to subgroups of individuals with mental ill-health. Although they were mostly positive about who would benefit, some reviews served as a warning to others about inadequate privacy or potentially harmful content.

Subtheme 3.2: App-Accessed Web-Based Communities

Reviews related to Web-based communities typically referred to their benefits. Some reviewers reported that they gained help from interactions with other users, viewing it as a resource to improve their mood when depressed, for example, “It helps when you’re feeling low.” The major benefit discussed was a normalizing experience, with reviewers gaining a sense of connection and understanding through communicating with others with mental ill-health. This understanding through shared experience was often expressed as something that was difficult to access in other aspects of life, for example, “It helps get your problems out if you can’t tell anyone else.” The direct or immediate support gained through Web-based communities was often reciprocated, with users providing encouragement to others.

However, not all found community interaction positive. Some expressed an unwillingness to have strangers privy to their experience of the disorder, for example, “I really don’t want the world to know how I feel.” The few negative reviews among those eager to engage with a community concerned the unresponsiveness, coldness, or immaturity of other users.

Major Theme 4: Wishlist

Approximately one in five reviews included a wishlist suggestion, that is, features reviewers thought would improve their app experience or health management. Reviewers made both direct requests for app changes and requested the ability to customize, with direct requests more common. Wishlist requests are displayed in Table 3; requests for tracking options were the most prevalent. Of the 34 additional symptoms requested for tracking, sleep was the most common.

Table 3. Requested features of apps for bipolar disorder.

Request frequency	Feature requested
Highly requested features	Customizable symptom tracking
	Multiple entries per day
	Export data
	Graph changes
	Additional tracking categories; sleep, medication, diet, mood triggers, exercise, anxiety, substance use
	Reminders
	Edit or delete inputs
	Longer section for notes
	Mood scale range or options
Minimally requested features	Personal identification number or password security
	Information backup
	User interface changes
	Social interaction capability
	Between app integration
	Sync across devices

Major Theme 5: Apps and Therapy

About one in ten reviews mentioned using apps in partnership with clinical services. Generally, these were (5.1) perspectives of app users and a minority of health care practitioners who saw the potential clinical utility of apps or (5.2) perceived therapeutic effects of app use.

Subtheme 5.1: Using Apps With Clinical Care

Comments were contributed, in the main, by app users, followed by a minority of health care practitioners (HCPs). HCPs indicated they were either recommending apps to clients or evaluating apps with a view to making recommendations in future. For HCPs, apps were another tool that could help their clients develop insights into their moods, triggers, and behaviors; provide feedback; and facilitate between session symptom management.

Consumers agreed that apps were a useful adjunct to clinical care, though sometimes for different reasons. Whereas users also mentioned “insight,” they reported that apps could help their HCPs gain greater insight into their experience of the disorder. For example, apps could relay their experience with mood swings, for example, “*This app has helped me and my clinician both better understand my mood changes,*” responses to medication and general progress to guide treatment decisions based on a more “accurate” record of their symptoms. The concept of accuracy was recurring, with apps taking the onus off memory and in turn improving communication with their HCP, for example, “*I can’t remember what brings on my bipolar episodes and this allows me to accurately tell my psychiatrist how my moods have been.*” To facilitate this communication, many commented on the capacity (existing or desired) for apps to export data to show HCPs. An increased likelihood of participating in between session activities was also expressed, for example, “*Can now do my therapy homework.*” Finally,

there was a sense that in sharing apps with their HCP, they might be helping others in similar situations.

Rarely, reviews identified the ability of apps to support family care. These reviewers used apps to monitor behavior or symptoms in their partners, children, or other family members not restricted to bipolar disorder.

Subtheme 5.2: Therapeutic Effects of App Use

A few users indicated the app itself had a therapeutic effect, positively affecting mood or well-being. Being able to see mood changes was reported as a motivator to do well, particularly if this data was visible to or shared with others. A few reviewers developed a therapeutic alliance or partnership with the app. Typically, these reviews alluded to working on mental health “together” with some reference to receiving support from the app, for example, “*We have made it through thick and thin together.*”

Minor Theme 1: App Cost

The minority of reviews that referenced app cost were varied in nature. Reviews indicated a willingness to pay for one-off donations, “pro” versions, removal of advertising, and overwhelmingly, a good product. “Good products” were easy to use, aesthetically pleasing, worked seamlessly, integrated with other software, or synced to other devices, and importantly, contained features deemed “useful” or “helpful.” Whereas many that had purchased apps indicated that it was worth it, those who disagreed often used strong language, for example, “*I feel like I’ve been swindled.*” Cost influenced app expectations with reviewers using price as a guide, for example, “*For the price it really needs improvement*” or “*Good for a freebie.*” In general, negative reviews related to app prices were largely complaints about paying for features or services not received (eg, poor tech support or apps that crashed or froze).

Minor Theme 2: Privacy and Data Security

Similarly, a minority of reviews related to app privacy or data security. Data privacy and security was generally conceived of in terms of handset access, with most comments related to using a personal identification number, login, or password to secure access to the app and data stored within it. Generally, people supported these approaches. Other comments concerned cyber access, permissions requested by the app in the terms of service use, with main concerns regarding data storage practices or data sharing with third parties, for example, *"I began to feel that this app was a useless little spybot."*

However, most comments about privacy were generated in response to one particular app, which was designed so that the mood journal was public by default and made private by upgrade purchase. That is, user entries could be seen and commented on by other app users, as part of a Web-based community, unless users paid Aus \$1. This approach uncovered conflicting perspectives. Contrasting views surrounded which should be default, public, or private entries, with some stating that privacy should be default, for example, *"Already a stranger commented on my mood...it's no-one's business! I'm uninstalling this app—imagine if my stalker found it!"* Others were open to sharing personal information and highlighted benefits of the community created (eg, finding support). Linking privacy to app cost was also contentious. Some users reported that privacy was a right, and therefore, not for sale, whereas others felt that privacy was worth paying for and, in this case, was inexpensive. For those who objected to linking privacy and price, the language used was notably strong, for example, manipulative, cruel, terrifying, offensive, greedy, unethical, extortion, black mail, morally wrong, and intolerable.

Minor Theme 3: Comparisons With Traditional Monitoring

A small minority of reviews addressed the benefits of app-based tools over traditional paper and pencil resources. Benefits included increased access, ease, and convenience, leading to increased reliability, for example, *"Loving the simplicity compared to paper methods."* Increased reliability was also linked to an app's ability to remind users to complete monitoring, for example, *"I've set a reminder so I'll definitely fill it in daily."* Although collectively positive, a few reviews stated a preference for paper and pencil tools as they were seen as more advanced, or more secure, for example, *"I trust paper and pen a little more."*

Minor Theme 4: Evidence-Based mHealth

Minimal reviews commented on the evidence-base or scientific quality of an app. Those that did referred to research or scientific involvement or reference to well-known tools in development, or health professional endorsement, as indications of app quality.

Discussion

Principal Findings

This is one of the first explorations of consumer perspectives on currently available mHealth tools for mental health using publically available review data. The mixed qualitative method

employed here builds on previous approaches [32] and allowed for both a quantitative summary of the main themes and fine-grained exploration of important subthemes. Our data show that user feedback on mental health apps could generally be summarized by 9 themes, with variation in their distribution across the reviews. Reviews varied in length and specificity, with many providing rich, detailed text with important consumer perspectives on the potential for apps to be helpful tools for mental health management.

The proportion of reviews containing positive and negative themes was similar to previous results in general apps [27]. When compared with reviews of apps that reduce harmful drinking [32], reviews of apps for bipolar disorder were more favorable. The main content of negative reviews reported here also supports previous literature citing functionality issues, lack of features, and crashing as the most common complaints [32,33]. However, reviews were rarely restricted to one theme, indicating that consumer needs and expectations about apps are complex and multidimensional, with no currently available app providing a satisfactory balance.

The predominance of contrasting phrases and "wishlist" requests within the reviews indicate that consumer needs are not adequately addressed by currently available apps for the disorder, even given the generally positive reviews. Many of the most requested features, for example, monitoring additions and reminders, were highlighted previously as areas in which apps failed to conform to evidence-based practice [7]. A comprehensive review of the features and scientific quality of the same apps noted that 65% had inadequate mood scales and reminder features were often absent or not functioning as intended [7]. These examples highlight failure to translate scientific best practice and consumer considerations into satisfactory and functional apps for self-management and likely reflect the high proportion of private, individual, or corporate developers [7]. However, user requests also extended beyond features present in clinically used self-management resources, indicating a need to balance user preference and clinical relevance.

Despite these unmet needs, reviewers were largely positive about the apps, valuing content that was supportive, helpful, and easy to use. However, sustaining such helpfulness and engagement throughout an episodic condition represents a specific challenge. The prevalence of reviews that referred to community demonstrates the importance of connectedness and social support in managing mental health, as well as the potential for Web-based communities to provide that support. Although previous research has been inconclusive about the benefits of Web-based communities for health problems [34,35], our results indicate that consumers perceive benefits, even when the community had not been sought. This concurs with research in which consumers emphasize the importance of support obtained through app communities [32]. Consumers reported that app communities were normalizing, supportive, and mutually beneficial, something that has been established in other health problems [36]. Users who gained understanding and support from community members often spoke of providing similar support to other users. Indeed, reciprocity has been noted as an important aspect of the social nature of self-management

and thus, app communities have the potential to provide this support to individuals who may otherwise not receive it [37]. Furthermore, as mental health related stigma is well documented and interferes with people's willingness to access support and mental health services [38,39], apps with associated Web-based communities may in part remove this barrier to support by being anonymous, ubiquitous, and inexpensive.

Although the overwhelming majority of reviews about community discussed the understanding supportive environment created, the potential for negative experiences requires acknowledgment. Many expressed unwillingness to discuss their mental health with others, emphasizing that community elements should be upfront and engagement optional. Indeed a review of Web-based health communities posited that willing engaged participants were necessary for communities to be beneficial [35]. Active participants may also reduce incidences of disappointment and distress expressed after not receiving support when reaching out. Although, negative effects of studies involving Web-based communities were not reported [35], such lack of support when solicited could potentially enhance feelings of stigma and isolation commonly experienced by those with mental health disorders [40]. Given the potential for supporting those in need, but also possible negative effects, it is crucial that this balance is considered during app development and that apps are derived, where possible, from a clear evidence-base and best clinical practice.

Although this need for established efficacy or evidence-informed content in mHealth tools for mental health is widely acknowledged by mHealth researchers and clinicians [7,41-43], discussion of scientific quality was strikingly absent from consumer reviews. This could represent genuine disinterest in the scientific basis of mental health apps, or could reflect a lack of health app regulation knowledge among consumers. With most apps classified as health and fitness or medical [7], there may be an assumption that they are on the app store because they "work." In this case, a reviewer would discuss scientific quality only if an app obviously violated that assumption. Due to the growing number of apps for mental health [4], it is imperative that consumer "app-literacy," knowledge regarding evidence and data privacy of apps, is understood. This understanding is vital for successful knowledge translation of evidence-based mHealth tools and requires that future research investigate consumer knowledge about app use for mental health. There is also a clear role for researchers to develop better communication and advocacy regarding clinically effective apps built on robust scientific data.

A further user expectation elucidated in the analysis was that developers would address their needs and perspectives. This was an implication of the wider community of other app users (active or potential) and developers that was identified around apps. Previous research reinforces the importance of developer responsiveness, reporting that developers who respond to user reviews receive a 1-point increase in star rating, approximately a third of the time [44]. This conception of community has important implications for community-based participatory research and knowledge translation. Whereas currently this interplay between users and developers is occurring after an app's deployment, acknowledgment by the user base that they

have useful knowledge regarding app features and disorder management, highlights the potential for engaging consumers throughout the development process. Consumer participation in design has obvious advantages for the resulting app's function and uptake and use by the target community [45,46]. Furthermore, consumer involvement in the delivery of services can effectively support recovery, with related benefits in consumer empowerment, social inclusion, satisfaction with services provided, and well-being [47,48]. Capitalizing on this potential in Web-based service delivery is an important avenue for future research, with consumer involvement actively recommended [47]. Such recommendations have recently been echoed by The International Society for Bipolar Disorders working group on consumer research involvement [49].

In keeping with the concept of meaningful partnerships, several reviews contained information related to routine clinical management, as frequently, disorder specific apps were not used independently. Reviewers used apps to inform and support their relationship with their HCP, with emphases on improved recall and feeling better understood during clinical appointments. Many reported that sharing app data with their HCP helped to represent their lived experience more fully, potentially promoting greater therapeutic alliance [50]. Clearly, apps for mental health are not viewed as replacements for clinical care, but rather as useful adjuncts to treatment.

Whereas the potential of technology for scaling mental health interventions and delivering mental health services is recognized for Web-based mental health resources [51], the recency of the field and lack of an established evidence base mean such initiatives have not extended to mHealth. Our results show that apps are already used by consumers looking for tools to manage their health, which supports anecdotal evidence that consumers are driving the introduction of apps into clinical practice. Little research has explored HCPs' attitudes to the use of apps in clinical care, yet current data indicate that some consumers expect HCPs to be open to app use and some HCPs are already attempting to determine the best available tools. Given the current lack of evidence-based mHealth resources, there is a clear need to support both consumers and HCPs with an interest in using apps distinguish those of high quality.

Limitations

This study is not without limitations. The data used were publically available existing text, rather than data gathered through more exploratory methods, using more refined and/or explicit questions. It is possible that there may be salient themes for app development and disorder management that are elided by the data gathering techniques used here, or that important consumer attitudes have not been adequately captured. Similarly, we cannot be clear about the needs or preferences of consumers in the community who have no experience using apps to manage their mental health. However, given richness of the data contained in these reviews, the large number of reviews recorded across multiple apps spanning more than five years, we can be reasonably confident the findings reported here represent real concerns held by consumers in the community. Future research will need to further clarify and confirm that the themes reported resonate with a wider community of consumers and explore the

needs of those who have not previously engaged with mHealth technologies.

Reviews of apps for bipolar disorder were restricted to apps reported in Nicholas et al [7]. However, this methodology had the clear advantage that the function and quality of apps reviewed were known and consumer perspectives could be compared with quality assessment. Nevertheless, we acknowledge that in the time since this review, more apps for the condition are likely to have been made available, therefore potentially limiting study comprehensiveness. Finally, as the app marketplace is dynamic and transforming [52], the apps, their features, and opinions of consumers are subject to change.

Conclusions

This paper provides a unique perspective on consumer attitudes and expectations toward mental health apps, and hence, represents an important contribution to the knowledge base in the expanding research area of mHealth. Consumers value content that is helpful, supportive, and easy to use, and they are

integrating apps into their health management without necessarily considering evidence-base or the clinical effectiveness of the tool. Such consumer insights are vital to our ability to be competitive in the unregulated app store environment but also reveal the need to balance user preferences and clinical relevance. Indeed, integrated knowledge translation strategies involving consumers in all stages of mHealth research may be critical to ensuring uptake and continued use, and results indicate such research strategies would be acceptable to consumers.

The expectation of developer responsiveness has implications for research app development and mHealth resource funding. Currently funding awards do not account for resource sustainability costs, a critical facet of existing and competing in the dynamic app marketplace. To realize the potential of mHealth to support self-management, increase access to care, and provide mental health resources, a change in app development practices and funding structures for such resources is required.

Acknowledgments

The authors would like to thank Mark Erik Larsen for his assistance in extracting consumer reviews from the Google Play store. This study is supported by a grant from the National Health and Medical Research Council (NHMRC) and forms part of research conducted by the NHMRC Centre for Research Excellence in Suicide Prevention (CRESP; APP1042580). Additionally, JN is supported by an Australian Postgraduate Award, and HC by an NHMRC Fellowship 1056964. The NHMRC and APA were not involved in the study design, data collection, data analysis, report writing, and decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of subthemes and minor themes identified.

[PDF File (Adobe PDF File), 37KB - [jmir_v19i4e105_app1.pdf](#)]

Multimedia Appendix 2

List of apps included in the analysis.

[PDF File (Adobe PDF File), 25KB - [jmir_v19i4e105_app2.pdf](#)]

References

1. Proudfoot J, Parker G, Hadzi PD, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. *J Med Internet Res* 2010;12(5):e64 [FREE Full text] [doi: [10.2196/jmir.1475](#)] [Medline: [21169174](#)]
2. Whittaker R. Issues in mHealth: findings from key informant interviews. *J Med Internet Res* 2012;14(5):e129 [FREE Full text] [doi: [10.2196/jmir.1989](#)] [Medline: [23032424](#)]
3. Bardram J, Frost M, Szántó K, Faurholt-Jepsen M, Vinberg M, Kessing L. Designing Mobile Health Technology for Bipolar Disorder: A Field Trial of the MONARCA System. 2013 Presented at: CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 27 - May 2, 2013; Paris, France. [doi: [10.1145/2470654.2481364](#)]
4. Firth J, Torous J, Yung AR. Ecological momentary assessment and beyond: the rising interest in e-mental health research. *J Psychiatr Res* 2016;80:3-4 [FREE Full text] [doi: [10.1016/j.jpsychires.2016.05.002](#)] [Medline: [27236099](#)]
5. Murray G, Suto M, Hole R, Hale S, Amari E, Michalak EE. Self-management strategies used by 'high functioning' individuals with bipolar disorder: from research to clinical practice. *Clin Psychol Psychother* 2011;18(2):95-109. [doi: [10.1002/cpp.710](#)] [Medline: [20572206](#)]
6. Russell SJ, Browne JL. Staying well with bipolar disorder. *Aust N Z J Psychiatry* 2005;39(3):187-193. [doi: [10.1111/j.1440-1614.2005.01542.x](#)] [Medline: [15701069](#)]

7. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. *J Med Internet Res* 2015;17(8):e198 [FREE Full text] [doi: [10.2196/jmir.4581](https://doi.org/10.2196/jmir.4581)] [Medline: [26283290](https://pubmed.ncbi.nlm.nih.gov/26283290/)]
8. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013;15(11):e247 [FREE Full text] [doi: [10.2196/jmir.2791](https://doi.org/10.2196/jmir.2791)] [Medline: [24240579](https://pubmed.ncbi.nlm.nih.gov/24240579/)]
9. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. *BMC Med* 2012;10:144 [FREE Full text] [doi: [10.1186/1741-7015-10-144](https://doi.org/10.1186/1741-7015-10-144)] [Medline: [23171675](https://pubmed.ncbi.nlm.nih.gov/23171675/)]
10. Pandey A, Hasan S, Dubey D, Sarangi S. Smartphone apps as a source of cancer information: changing trends in health information-seeking behavior. *J Cancer Educ* 2013;28(1):138-142. [doi: [10.1007/s13187-012-0446-9](https://doi.org/10.1007/s13187-012-0446-9)] [Medline: [23275239](https://pubmed.ncbi.nlm.nih.gov/23275239/)]
11. Larsen ME, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. *PLoS One* 2016;11(4):e0152285 [FREE Full text] [doi: [10.1371/journal.pone.0152285](https://doi.org/10.1371/journal.pone.0152285)] [Medline: [27073900](https://pubmed.ncbi.nlm.nih.gov/27073900/)]
12. Abroms LC, Lee WJ, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. *Am J Prev Med* 2013;45(6):732-736 [FREE Full text] [doi: [10.1016/j.amepre.2013.07.008](https://doi.org/10.1016/j.amepre.2013.07.008)] [Medline: [24237915](https://pubmed.ncbi.nlm.nih.gov/24237915/)]
13. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. *JMIR Mhealth Uhealth* 2015;3(3):e82 [FREE Full text] [doi: [10.2196/mhealth.4328](https://doi.org/10.2196/mhealth.4328)] [Medline: [26290327](https://pubmed.ncbi.nlm.nih.gov/26290327/)]
14. Consumer Health Information Corporation. Consumer-health. Motivating patients to use smartphones health apps URL: <http://www.consumer-health.com/motivating-patients-to-use-smartphone-health-apps/> [accessed 2017-01-12] [WebCite Cache ID 6nRol8kVJ]
15. Pinch Media. Slideshare. iPhone AppStore Secrets URL: <http://www.slideshare.net/pinchmedia/iphone-appstore-secrets-pinch-media?type=presentation> [accessed 2017-01-12] [WebCite Cache ID 6nRowwNm4]
16. Di Sorbo A, Panichella S, Alexandru C, Shimagaki J, Visaggio C, Canfora G, et al. What Would Users Change in My App? Summarizing App Reviews for Recommending Software Changes. 2016 Presented at: The 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering; November 13-18; Seattle, WA. [doi: [10.1145/2950290.2950299](https://doi.org/10.1145/2950290.2950299)]
17. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
18. Neuendorf K. *The Content Analysis Guidebook*. Thousand Oaks, CA: SAGE Publications; 2002.
19. Tesch R. *Qualitative research: Analysis types and software tools*. Bristol, PA: Falmer; 1990.
20. Mayring P. *Qualitative Content Analysis*. Forum Qual Soc Res 2000;1(2).
21. Patton M. *Qualitative research and evaluation methods* (3rd edition). Thousand Oaks, CA: SAGE Publications; 2002.
22. Sandelowski M. Whatever happened to qualitative description? *Res Nurs Health* 2000;23(4):334-340. [Medline: [10940958](https://pubmed.ncbi.nlm.nih.gov/10940958/)]
23. Sandelowski M. Qualitative analysis: what it is and how to begin. *Res Nurs Health* 1995;18(4):371-375. [Medline: [7624531](https://pubmed.ncbi.nlm.nih.gov/7624531/)]
24. Forman J, Damschroder L. Qualitative Content Analysis. In: Jacoby L, Siminoff L, editors. *Empirical Methods for Bioethics: A Primer* (Advances in Bioethics, Volume 11). Oxford, UK: Emerald Group Publishing Limited; 2007:39-62.
25. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008;62(1):107-115. [doi: [10.1111/j.1365-2648.2007.04569.x](https://doi.org/10.1111/j.1365-2648.2007.04569.x)] [Medline: [18352969](https://pubmed.ncbi.nlm.nih.gov/18352969/)]
26. Gibbs G. *Analysing Qualitative Data*. London, UK: SAGE Publications; 2008.
27. Pagano D, Maalej W. User feedback in the appstore: An empirical study. 2013 Presented at: The 21st IEEE international requirements engineering conference (RE); July 15-19; Rio de Janeiro, Brazil. [doi: [10.1109/RE.2013.6636712](https://doi.org/10.1109/RE.2013.6636712)]
28. Morgan DL. Qualitative content analysis: a guide to paths not taken. *Qual Health Res* 1993;3(1):112-121. [Medline: [8457790](https://pubmed.ncbi.nlm.nih.gov/8457790/)]
29. Boydell KM, Volpe T, Pignatiello A. A qualitative study of young people's perspectives on receiving psychiatric services via televideo. *J Can Acad Child Adolesc Psychiatry* 2010;19(1):5-11 [FREE Full text] [Medline: [20119561](https://pubmed.ncbi.nlm.nih.gov/20119561/)]
30. Erlandson D, Harris E, Skipper B, Allen S. *Doing naturalistic inquiry: a guide to methods*. Newbury Park, CA: Sage Publications; 1993.
31. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85(3):257-268 [FREE Full text] [Medline: [15733050](https://pubmed.ncbi.nlm.nih.gov/15733050/)]
32. Milward J, Khadjesari Z, Fincham-Campbell S, Deluca P, Watson R, Drummond C. User preferences for content, features, and style for an app to reduce harmful drinking in young adults: analysis of user feedback in app stores and focus group interviews. *JMIR Mhealth Uhealth* 2016;4(2):e47 [FREE Full text] [doi: [10.2196/mhealth.5242](https://doi.org/10.2196/mhealth.5242)] [Medline: [27220371](https://pubmed.ncbi.nlm.nih.gov/27220371/)]
33. Khalid H, Shihab E, Nagappan M, Hassan AE. What do mobile app users complain about? *IEEE Softw* 2015;32(3):70-77. [doi: [10.1109/MS.2014.50](https://doi.org/10.1109/MS.2014.50)]
34. Demiris G. The diffusion of virtual communities in health care: concepts and challenges. *Patient Educ Couns* 2006;62(2):178-188. [doi: [10.1016/j.pec.2005.10.003](https://doi.org/10.1016/j.pec.2005.10.003)] [Medline: [16406472](https://pubmed.ncbi.nlm.nih.gov/16406472/)]
35. 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. *Br Med J* 2004;328(7449):1166 [FREE Full text] [doi: [10.1136/bmj.328.7449.1166](https://doi.org/10.1136/bmj.328.7449.1166)] [Medline: [15142921](https://pubmed.ncbi.nlm.nih.gov/15142921/)]
36. Mo PK, Coulson NS. Exploring the communication of social support within virtual communities: a content analysis of messages posted to an online HIV/AIDS support group. *Cyberpsychol Behav* 2008;11(3):371-374. [doi: [10.1089/cpb.2007.0118](https://doi.org/10.1089/cpb.2007.0118)] [Medline: [18537512](https://pubmed.ncbi.nlm.nih.gov/18537512/)]

37. Topor A, Borg M, Di GS, Davidson L. Not just an individual journey: social aspects of recovery. *Int J Soc Psychiatry* 2011;57(1):90-99. [doi: [10.1177/0020764010345062](https://doi.org/10.1177/0020764010345062)] [Medline: [21252359](https://pubmed.ncbi.nlm.nih.gov/21252359/)]
38. 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;45(1):11-27. [doi: [10.1017/S0033291714000129](https://doi.org/10.1017/S0033291714000129)] [Medline: [24569086](https://pubmed.ncbi.nlm.nih.gov/24569086/)]
39. Corrigan P. How stigma interferes with mental health care. *Am Psychol* 2004;59(7):614-625. [doi: [10.1037/0003-066X.59.7.614](https://doi.org/10.1037/0003-066X.59.7.614)] [Medline: [15491256](https://pubmed.ncbi.nlm.nih.gov/15491256/)]
40. Hawke LD, Parikh SV, Michalak EE. Stigma and bipolar disorder: a review of the literature. *J Affect Disord* 2013;150(2):181-191. [doi: [10.1016/j.jad.2013.05.030](https://doi.org/10.1016/j.jad.2013.05.030)] [Medline: [23759420](https://pubmed.ncbi.nlm.nih.gov/23759420/)]
41. Leigh S, Flatt S. App-based psychological interventions: friend or foe? *Evid Based Ment Health* 2015;18(4):97-99. [doi: [10.1136/eb-2015-102203](https://doi.org/10.1136/eb-2015-102203)] [Medline: [26459466](https://pubmed.ncbi.nlm.nih.gov/26459466/)]
42. Chan S, Torous J, Hinton L, Yellowlees P. Towards a framework for evaluating mobile mental health apps. *Telemed J E Health* 2015;21(12):1038-1041. [doi: [10.1089/tmj.2015.0002](https://doi.org/10.1089/tmj.2015.0002)] [Medline: [26171663](https://pubmed.ncbi.nlm.nih.gov/26171663/)]
43. Powell AC, Landman AB, Bates DW. In search of a few good apps. *J Am Med Assoc* 2014;311(18):1851-1852. [doi: [10.1001/jama.2014.2564](https://doi.org/10.1001/jama.2014.2564)] [Medline: [24664278](https://pubmed.ncbi.nlm.nih.gov/24664278/)]
44. McIlroy S, Shang W, Ali N, Hassan AE. Is it worth responding to reviews? A case study of the top free apps in the Google Play store. *IEEE Softw* 2015;99:1-6. [doi: [10.1109/MS.2015.149](https://doi.org/10.1109/MS.2015.149)]
45. Ellis L, McCabe K, Davenport T, Burns J, Rahilly K, Nicholas M, et al. Development and evaluation of an Internet-based program to improve the mental health and wellbeing of young men. *Interactive Technology and Smart Education* 2015;12(1):2-13. [doi: [10.1108/ITSE-05-2014-0009](https://doi.org/10.1108/ITSE-05-2014-0009)]
46. Ellis LA, McCabe KL, Rahilly KA, Nicholas MA, Davenport TA, Burns JM, et al. Encouraging young men's participation in mental health research and treatment: perspectives in our technological age. *Clin Investig* 2014;4(10):881-888. [doi: [10.4155/CLI.14.61](https://doi.org/10.4155/CLI.14.61)]
47. Grey F, O'Hagan M. Mental Health Commission of New South Wales. 2015. The effectiveness of services led or run by consumers in mental health: Rapid review of evidence for recovery-orientated outcomes URL: <https://goo.gl/XjxIDx> [WebCite Cache ID 6nRpAxN10]
48. Clay S, Schell B, Corrigan P, Ralph R. On Our Own, Together: Peer Programs for People With Mental Illness. Nashville, TN: Vanderbilt University Press; 2005.
49. Michalak EE, Jones S, Lobban F, Algorta GP, Barnes SJ, Berk L, et al. Harnessing the potential of community-based participatory research approaches in bipolar disorder. *Int J Bipolar Disord* 2016;4(1):4. [Medline: [26856996](https://pubmed.ncbi.nlm.nih.gov/26856996/)]
50. Clarke J, Proudfoot J, Whitton A, Birch MR, Boyd M, Parker G, et al. Therapeutic alliance with a fully automated mobile phone and web-based intervention: secondary analysis of a randomized controlled trial. *JMIR Ment Health* 2016 Feb 25;3(1):e10 [FREE Full text] [doi: [10.2196/mental.4656](https://doi.org/10.2196/mental.4656)] [Medline: [26917096](https://pubmed.ncbi.nlm.nih.gov/26917096/)]
51. Australian Government Department of Health and Ageing. Health. 2012. e-Mental health strategy for Australia URL: [https://www.health.gov.au/internet/main/publishing.nsf/content/7C7B0BFEB985D0EBCA257BF0001BB0A6/\\$File/emstrat.pdf](https://www.health.gov.au/internet/main/publishing.nsf/content/7C7B0BFEB985D0EBCA257BF0001BB0A6/$File/emstrat.pdf) [accessed 2017-01-12] [WebCite Cache ID 6nRpJBrAF]
52. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: longitudinal tracking of mental health apps. *JMIR Mhealth Uhealth* 2016;4(3):e96 [FREE Full text] [doi: [10.2196/mhealth.6020](https://doi.org/10.2196/mhealth.6020)] [Medline: [27507641](https://pubmed.ncbi.nlm.nih.gov/27507641/)]

Abbreviations

mHealth: mobile health

HCP: health care practitioner

Edited by G Eysenbach; submitted 11.01.17; peer-reviewed by J Milward, S Chan, M Faurholt-Jepsen; comments to author 23.02.17; revised version received 02.03.17; accepted 04.03.17; published 07.04.17.

Please cite as:

Nicholas J, Fogarty AS, Boydell K, Christensen H

The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder

J Med Internet Res 2017;19(4):e105

URL: <http://www.jmir.org/2017/4/e105/>

doi: [10.2196/jmir.7273](https://doi.org/10.2196/jmir.7273)

PMID: [28389420](https://pubmed.ncbi.nlm.nih.gov/28389420/)

©Jennifer Nicholas, Andrea S Fogarty, Katherine Boydell, Helen Christensen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 07.04.2017. This is an open-access article distributed under the terms of the Creative

Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles

Sohrab Saeb¹, PhD; Thaddeus R Cybulski², BSc; Stephen M Schueller¹, PhD.; Konrad P Kording², PhD; David C Mohr¹, MD, PhD

¹Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Northwestern University, Chicago, IL, United States

²Rehabilitation Institute of Chicago, Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL, United States

Corresponding Author:

Sohrab Saeb, PhD

Center for Behavioral Intervention Technologies

Department of Preventive Medicine

Northwestern University

750 N Lake Shore Dr

Rubloff Building, 10th floor

Chicago, IL, 60611

United States

Phone: 1 312 503 4626

Fax: 1 312 908 9588

Email: s-saeb@northwestern.edu

Related Article:

This is a corrected version. See correction statement: <http://www.jmir.org/2017/4/e143>

Abstract

Background: Sleep is a critical aspect of people's well-being and as such assessing sleep is an important indicator of a person's health. Traditional methods of sleep assessment are either time- and resource-intensive or suffer from self-reporting biases. Recently, researchers have started to use mobile phones to passively assess sleep in individuals' daily lives. However, this work remains in its early stages, having only examined relatively small and homogeneous populations in carefully controlled contexts. Thus, it remains an open question as to how well mobile device-based sleep monitoring generalizes to larger populations in typical use cases.

Objective: The aim of this study was to assess the ability of machine learning algorithms to detect the sleep start and end times for the main sleep period in a 24-h cycle using mobile devices in a diverse sample.

Methods: We collected mobile phone sensor data as well as daily self-reported sleep start and end times from 208 individuals (171 females; 37 males), diverse in age (18–66 years; mean 39.3), education, and employment status, across the United States over 6 weeks. Sensor data consisted of geographic location, motion, light, sound, and in-phone activities. No specific instructions were given to the participants regarding phone placement. We used random forest classifiers to develop both personalized and global predictors of sleep state from the phone sensor data.

Results: Using all available sensor features, the average accuracy of classifying whether a 10-min segment was reported as sleep was 88.8%. This is somewhat better than using the time of day alone, which gives an average accuracy of 86.9%. The accuracy of the model considerably varied across the participants, ranging from 65.1% to 97.3%. We found that low accuracy in some participants was due to two main factors: missing sensor data and misreports. After correcting for these, the average accuracy increased to 91.8%, corresponding to an average median absolute deviation (MAD) of 38 min for sleep start time detection and 36 min for sleep end time. These numbers are close to the range reported by previous research in more controlled situations.

Conclusions: We find that mobile phones provide adequate sleep monitoring in typical use cases, and that our methods generalize well to a broader population than has previously been studied. However, we also observe several types of data artifacts when collecting data in uncontrolled settings. Some of these can be resolved through corrections, but others likely impose a ceiling on

the accuracy of sleep prediction for certain subjects. Future research will need to focus more on the understanding of people's behavior in their natural settings in order to develop sleep monitoring tools that work reliably in all cases for all people.

(*J Med Internet Res* 2017;19(4):e118) doi:[10.2196/jmir.6821](https://doi.org/10.2196/jmir.6821)

KEYWORDS

sleep monitoring; mobile phones; decision trees; classification

Introduction

Background

Sleep is intrinsically linked to many aspects of life, including both physical and mental health [1]. The connection between sleep and well-being is bidirectional, where sleep disorders can cause significant burden on a person's life, and underlying disease can manifest itself as disruptions a person's sleep. These links manifest themselves in a number of facets of a person's health, from immune and metabolic effects [2] to disordered sleep patterns as a part of disease processes such as schizophrenia, depression, or post-traumatic stress disorder [3,4]. As such, sleep can provide a unique window into monitoring, tracking, or treating disease processes, and be both a target and outcome of intervention [5]. Thus, monitoring sleep is important.

Indeed, sleep monitoring plays a critical role in current clinical practice. Polysomnography, the "gold standard" for diagnosis of sleep disorders, monitors a variety of signals over the course of several nights, for example, electroencephalogram (EEG), breathing, and muscle and eye movements, to produce a detailed picture of a patient's sleep patterns [6]. Ambulatory polysomnography is a lower cost option than in-clinic assessment and acquiring data from a person's home environment might better represent their typical sleep patterns [7]. However, it is still expensive, time-consuming, and the tools used for assessment might themselves interact with the sleep behavior. Thus, for chronic sleep tracking, clinicians have typically relied on instruments such as sleep diaries, questionnaires, and similar instruments [8]. These approaches have several drawbacks, such as patient adherence and reporting bias [9]. Having a way to monitor sleep that does not suffer from these drawbacks but is easier to perform than polysomnography would be a boon for clinical practice and research.

With the advent of mobile phones, a majority of Americans now carry a multifunctional sensor platform in their pocket [10]. These devices and other wearable activity sensors can be used to monitor a person's behavior and environment, and as a result, can be used to monitor sleep. Previous work has used this mobile-sensor-based approach to predict sleep with relatively high accuracy. One study predicted the sleep or awake state of every 10-min long bin of phone sensor data with 93% accuracy [11]. Another study estimated the sleep duration with an error of 42 min [12], and a subsequent study on a college student population was able to predict bedtimes with an accuracy of 25 min of the ground truth [13]. Finally, a more recent study was able to predict sleep or awake states with 89% accuracy solely based on the users' interactions with their phones [14]. These approaches hold promise for sleep tracking in the future;

however, there remains significant work to do before they can be used more generally.

Several issues impair the ability to apply these findings to the general population. First, much of this work has used small subsets of the population, mostly students [12]. Students tend to be homogeneous in terms of demographics such as age and other patterns such as school schedules, and some evidence suggests that these demographic and life similarities might impact their sleep patterns [9,15]. Second, study participants typically receive instructions, such as placing the mobile phone face-down on the bed with them as they went to sleep [12] or keeping the phone turned on and to keep it in their bedroom while sleeping [11]. Although this increases the reliability of automated sleep assessment, to the degree that people to change their daily habits, it means that attempts to use these assessments in the wild will likely fail. Finally, many studies simply exclude noncompliant participants in the analysis. However, noncompliance may be related to other factors, such as untraditional sleep schedules, that might bias results in ways that reduce generalizability [13]. Therefore, classifiers (algorithms that distinguish sleep from waking states) that do not depend on specific instructions regarding the use or placement of the phone and which are generalizable to a broader population still need to be tested.

Aim of This Study

In this study, we aim to explore the use of mobile devices for sleep tracking in a broad population of participants. Participants are recruited from across the United States without restrictions on age, leading to a substantially more heterogeneous sample than previous work. Participants use their own personal devices and are given no instructions on device use, allowing us to gather data from the natural, daily course of their lives. We will use techniques from machine learning to detect the sleep times of each participant, and will examine whether these techniques will generalize to other participants. Overall, we will assess if, and to what extent, we can scale passive sleep monitoring, from normal everyday phone use, to the more general population.

Methods

Participant Recruitment

We recruited the participants for our study between October 28, 2015 and February 12, 2016. The recruitment was done in collaboration with Focus Pointe Global (FPG), a company that specializes in market and scientific research strategies and participant recruitment and retention. FPG used Internet and qualitative panels of participants as a primary means of recruitment. They sent out emails to these panels with links to the screener questionnaire. Additionally, they used phone calls to potential participants in their in-house registries.

In the screener questionnaire, interested individuals were screened for eligibility. Individuals were eligible for our study if they were at least 18 years old, able to read and understand English, owned a mobile phone with Android 4.4 through 5.1 (excluding 5.0 due to problems that limited reliable access to some sensor data), and had access to WiFi for at least one 3-h period a day. We excluded individuals who were diagnosed with any psychotic disorders, were identified as not being able to walk more than half a mile (4 city blocks), or had positive screens for alcohol abuse (alcohol use disorder identification test, AUDIT [16] score ≥ 16), drug abuse (drug abuse screening test, DAST-10 [17] score ≥ 6), suicidal ideation (patient health questionnaire-9 item, PHQ-9 [18] item 9 rating ≥ 1 ; Beck depression inventory, BDI-II [19] item 9 rating ≥ 2), or bipolar disorder (mood disorder questionnaire, MDQ [20] question 1 score ≥ 7 , an endorsement of question 2, and a response of 2 or 3 for question 3). We also excluded those individuals who shared their phone with others. Eligible participants were consented using procedures approved by the Northwestern University Institutional Review Board, which included descriptions of the data to be gathered along with data security and privacy policies. We selected roughly equal numbers of participants in four groups, such that there were wide ranges of depression and

anxiety symptoms in the sample. We defined the groups as depressed and anxious (PHQ-9 ≥ 10 ; generalized anxiety disorder-7 item, GAD-7 ≥ 10), depressed only (PHQ-9 ≥ 10 ; GAD-7 < 10), anxious only (PHQ-9 < 10 ; GAD-7 ≥ 10), and healthy (PHQ-9 < 10 ; GAD-7 < 10).

Each participant was enrolled for a period of 6 weeks. First, a study ID was assigned to the participant by FPG. Then participants were asked to complete a Web-based questionnaire consisting of demographics (eg, age, gender, race and ethnicity, state of residence) and life aspects (eg, living situation, employment issues, where they keep their phone) that could potentially affect their sleep and phone use. Participants were compensated from US \$25 to US \$270.40 depending on how long they stayed in the study, and how much of the daily questionnaires they answered.

Data Collection

We collected two categories of data: mobile phone sensor data and ecological momentary assessment (EMA) data, which consisted of daily questions sent to participants asking them about their last night sleep times. The sensors used in our study and their attributes are listed in Table 1.

Table 1. List of the mobile phone sensors and their attributes, used in our study.

Sensor	Description
Activity	Physical activity class provided by the Android Activity Recognition API ^a (<i>still, walking, running, tilting, on bike, in vehicle, unknown</i>) and the confidence of the classifier (0-100%)
Light	Light intensity (lux)
Sound	Average sound intensity (dB) and dominant sound frequency (Hz)
Screen	State of the phone screen (on or off)
Battery	State of the battery (<i>not charging, charging via power cable, charging via USB</i> ^b)
GPS ^c location	Geographic latitude and longitude in degrees
WiFi	The MAC ^d address of the access point which the device is currently connected to
Communication events	Contact names, contact numbers, outgoing or incoming calls, outgoing or incoming SMS ^e
Time of day	Time of the day

^aAPI: application program interface.

^bUSB: Universal Serial Bus.

^cGPS: Global Positioning System.

^dMAC: media access control.

^eSMS: short message service.

EMA data was collected on a daily basis. On each day, at 9am local time, the questionnaire was launched on each participant's phone, asking them about the time they went to sleep last night, or *sleep start time*, and the time they woke up, or *sleep end time*. Participant could respond immediately to the questions, or delay the response until later that day. If they did not answer the questions before 12am that night, the questionnaire disappeared; and on the next day, a new questionnaire was launched asking about sleep start and end times of the night before. *Sleep duration* was defined as the time from sleep start time to sleep end time.

We used Purple Robot [21] to collect both sensor and EMA data. Purple Robot is a multipurpose, open-source Android app that is developed for our phone-based behavioral sensing studies [22], and adapted to this study. The app gathers data from the sensors available on the phone, initially stores them locally on the device, and then transmits them as network connectivity becomes available. This allows data collection in a variety of wireless connectivity scenarios with the confidence that intermittent network access does not affect the nature, quality, or quantity of the collected data.

Purple Robot anonymized sensitive information before storage and transmission. Specifically, it used a standard MD5 hashing

algorithm [23] to anonymize the contact names and numbers in the communication events sensor (see Table 1), as well as the participant IDs. Once the data was anonymized, it was locally stored on the device, transmitted to secure data collection server via encrypted, password-protected tunnels, and then deleted from the device. The mobile phone data residing on the server could be linked with other information gathered during the study only if the unique identifiers used by the participants and the study-specific keys used to encrypt the data were known. Furthermore, these were only accessible to individuals with the proper credentials. Overall, these security measures helped to protect the participants' privacy, particularly regarding sensor data such as GPS and MAC addresses, which could risk leakage of personal information.

Initial tests showed that the sound sensor (microphone) was draining battery power to a considerable degree, which could interfere with our data collection and dissatisfy the participants. Thus, we sample the microphone every 5 min for 30 s at a time. The Purple Robot sound sensor then reported the average sound amplitude (dB) and the dominant sound frequency during that 30 s period. The dominant frequency was calculated by taking the Fast Fourier Transform (FFT) of the signal, and finding the frequency at which it was maximum. Using this procedure, we considerably decreased the battery power consumption by Purple Robot.

Feature Extraction

Before using the collected phone sensor data for developing sleep detection algorithms, we extracted their attributes, or *features*. To extract features, we first divided all sensor data into 10-min-long windows. Then, from each window, we extracted 22 distinct features as listed in Table 2. The choice of 10 min was made for consistency with previous research [11]. In our feature set (Table 2), we included features that had previously been shown to be useful [11-13]. For the location features, *location variance*, and *location change*, we converted the GPS coordinates in latitude and longitude degrees to 2D coordinates in kilometers, using the method described in [24], before extracting the features. In addition, we also included time of the day as a feature, since we hypothesized that the time alone is a strong predictor of whether a person is asleep or awake.

To deal with missing sensor data, we used different strategies for different sensors. For the communication events and the screen sensors, we used a 0 value when data was not present, as for these cases absence of data meant no events. For the activity sensor, since the Android's Activity Recognition API (application program interface) does not generate new samples when the phone has been in the same state for a long time, we filled the missing points with the activity sample from the last window which contained data. For the rest of sensors, if the window was empty, the corresponding features were set to "Not a Number" (*NaN*).

Table 2. List of features used in the study.

Feature	Description
Stillness	Percentage of <i>still</i> activity
Light power	Mean of light intensity
Light range	Range of light intensity
Light kurtosis	Kurtosis of light intensity
Light change	$\langle (L(t) - L(t-1))^2 / L(t-1)^2 \rangle$ $L(t)$: light intensity at time t , and $\langle \cdot \rangle$ denotes the average over time.
Audio power	Mean of audio power
Audio freq min	Min. dominant audio frequency
Audio freq max	Max. dominant audio frequency
Screen activity	Number of screen ON or OFF events excluding the ones that last less than 30 s
Location variance	$\sqrt{(\sigma_{\text{lat}}^2 + \sigma_{\text{lng}}^2)}$ where <i>lat</i> and <i>lng</i> are latitude and longitude values in kilometers, respectively.
Location change	Average of change (as defined for light change) between latitude and longitude
Battery charging	1 if mode of battery state is <i>charging</i> ; 0 otherwise
Battery USB ^a	1 if the phone is connected to USB, 0 otherwise
Battery level	Average battery level (0-100)
WiFi	Mode of WiFi MAC ^b address (converted to integer by summing up the characters)
Last name	Last contact name (encrypted) contacted by either call or SMS ^c
Last number	Last phone number (encrypted) contacted by either call or SMS
Call	Number of phone calls
SMS	Number of SMS
Outgoing call	Number of outgoing phone calls
Outgoing SMS	Number of outgoing SMS
Time of day	Time of the day in hours (0-24), defined as the midpoint in the window

^aUSB: Universal Serial Bus.^bMAC: media access control.^cSMS: short message service.

Sleep Detection

Overview

We trained algorithms to detect the sleep start and wake-up times of each participant from the sensor features extracted from their phones. These algorithms, also called *classifiers*, determined whether each feature sample, extracted from 10 min of sensor data, was from a sleep or an awake state as reported by participants. After training, the classifiers were able to predict the state of a given feature sample.

The sleep detection procedure had two stages: first, we used *random forests* to estimate the probability for a feature sample to be from sleep or awake states. Then, a hidden Markov model (HMM) used the sequence of these probabilities to determine whether the participant's state was actually sleep or awake. In the following, we will give a more detailed description of these two stages.

Estimating State Probabilities

To estimate the probability for each feature sample being from an awake or asleep state, we used ensembles of decision trees known as random forests [25]. Each tree in a random forest makes a prediction, or vote, about the class of the feature sample. The random forest calculates the class probability by averaging over the predictions of individual trees. In this study, we used 50 trees.

We trained the random forest to estimate the state (*awake* or *sleep*) probabilities based on the last 5 feature sets extracted from the last 5 windows. For training, we used the bagging method [26], which randomly samples the dataset with replacement to create a training set for each tree. In this way, each tree only observes part of the dataset. In addition, each decision node in the tree randomly samples 5 out of 22 features, and finds the best feature and the best split value based on a Fisher information gain criterion. Therefore, each tree only observes part of both the data samples and the features. This

makes random forests less prone to overfitting and a better candidate for generalization to unseen data [27].

Determining States

Although our random forest classifiers use the last 5 feature samples to provide the class probability of the current sample, they ignore the class probabilities of the surrounding samples. This disregards the fact that sleep and awake states change slowly over time. In fact, transitions from *awake* to *sleep* and vice versa usually happen once in a given 24-h period. Therefore, it is important to consider the class probabilities of the neighbor samples in calculating the class probability of any given sample.

To determine the *sleep* or *awake* states, we first use a median filter in order to reduce the effect of fast changes in the class probabilities. A median filter replaces each sample by the median of w neighboring samples. Here, we set $w=21$, corresponding to 210 min in the data. After recalculating the probabilities, we use the threshold of 0.5 to determine the class of each sample (Probability \leq .5: *awake*; Probability $>$.5: *sleep*). In this way, the median filter captures the slower dynamics of the state probabilities.

After recalculating the state probabilities, the next step is to determine the states. For this, we use a HMM, which is a Bayesian statistical model that infers the states of an unobserved variable, sleep state in our study, given a set of observations, here the set of states estimated by the median filter. The HMM uses a set of parameters called transition probabilities, which represent the probability of transition between the classes. Because there are typically only one sleep-to-awake and one awake-to-sleep transition in each 24-h period, and given that we have 144 feature samples in each 24-h period, we set the transition probabilities as the following:

$$T_{(\text{sleep-awake})} = T_{(\text{awake-sleep})} = 1/144$$

$$T_{(\text{sleep-sleep})} = T_{(\text{awake-awake})} = 143/144$$

Training and Cross-Validation

We train sleep detection models in two different ways: (1) global models and (2) personal models. The former is trained on all data from a number of participants and cross-validated on the rest, whereas the latter is trained and cross-validated on the data from the same participant at different times.

For the global models, we use a subject-wise, 10-fold cross-validation method. We first divide the participants into 10 almost equal, nonoverlapping sets. Then, we train models on all sets except one and cross-validate it on the remaining set. We repeat this procedure 10 times so that all participants are used for cross-validation.

To train personal models, we divide each participant's data into 3 nonoverlapping folds. Then, we train models on 2 folds and validate them on the remaining fold. We repeat this procedure 3 times until all folds have been used for validation. The classification accuracy was averaged across the folds, representing the classification accuracy for the subject.

Results

Participants

In total, 208 eligible participants were recruited for the study. One participant did not install the software on their phone, and therefore was removed from the analysis. Of the 207 participants included in the analysis, 82.6% (171/207) were females and 17.4% (36/207) were males. Their ages ranged between 18 and 66 years old, with a mean of 39.3 (SD 10.3). They represented a geographically diverse sampling of the United States, as shown in Figure 1. Participants did not perfectly represent the racial and ethnic diversity of the United States with 78.7% (163/207) Caucasian, 11.6% (24/207) African American, 2.4% (5/207) Asian, 1.4% (3/207) Native American, and the remaining 4.3% (9/207) of participants were a combination of two or more races. It was found that 1.4% (3/207) of participants preferred not to specify race and 9.2% (19/207) of participants noted Hispanic as their ethnicity. Nevertheless, this is a diverse pool of demographics and locations.

The outcomes on the questionnaires asked during the screening were as follows: the average drug abuse score (DAST-10) was 0.56 (SD 1.06), alcohol abuse score (AUDIT) was 3.66 (SD 3.35), depression score (PHQ-9) was 9.72 (SD 5.10), and anxiety score (GAD-7) was 9.01 (SD 5.41). As expected, the drug and abuse scores were low, since we excluded individuals with high scores. However, there was a wide distribution in depression and anxiety scores as it was intended in the recruitment procedure.

Participants had diverse educational backgrounds: 1.9% (4/207) of participants had some high school education, 12.1% (25/207) had completed high school, 35.3% (73/207) had some college training, 13.5% (28/207) had 2-year college training, 23.6% (49/207) had Bachelor's degree, 11.1% (23/207) had Master's degree, and 2.4% (5/207) had professional Doctorate degree.

Finally, we asked the participants questions about the aspects of their lives that would potentially influence sleep detection. Of the 207 participants, 14.5% (30/207) lived alone, whereas 85.0% (176/207) lived with other people, and 0.5% (1/207) did not specify. In response to the employment status question, 61.4% (127/207) were employed, 20.8% (43/207) were unemployed, 8.2% (17/207) had disability which prevented them from working, 1.9% (4/207) were retired, and 7.7% (16/207) did not specify their employment status. Of the 127 employed participants, 78.0% (99/127) had one job, 18.1% (23/127) had two, 3.1% (4/127) had three, and 0.8% (1/127) had four jobs. It was found that 87.4% (181/207) of participants mentioned that they keep their phones in their bedrooms while sleeping, whereas 12.6% (26/207) keep it in another room. It was also found that 58.5% (121/207) of participants said that they share their bedrooms with someone, whereas 41.5% (86/207) sleep alone in their bedroom. As should be expected, a broad range of life situations occurred.

In addition to understanding the lives of our participants, the purpose of collecting these data was to assist sleep detection algorithms, by adding them to sensor features as inputs. However, our initial tests showed that they were not helpful in

detecting sleep, and therefore we did not use them in later analyses.

Figure 1. Locations of participants on the map, shown as red dots. We added a small random value, drawn from a Gaussian distribution with zero mean and standard deviation of 1.5 km in geographic distance, to each participant's location so that their exact coordinates cannot be extracted from the figure.



Data Characteristics

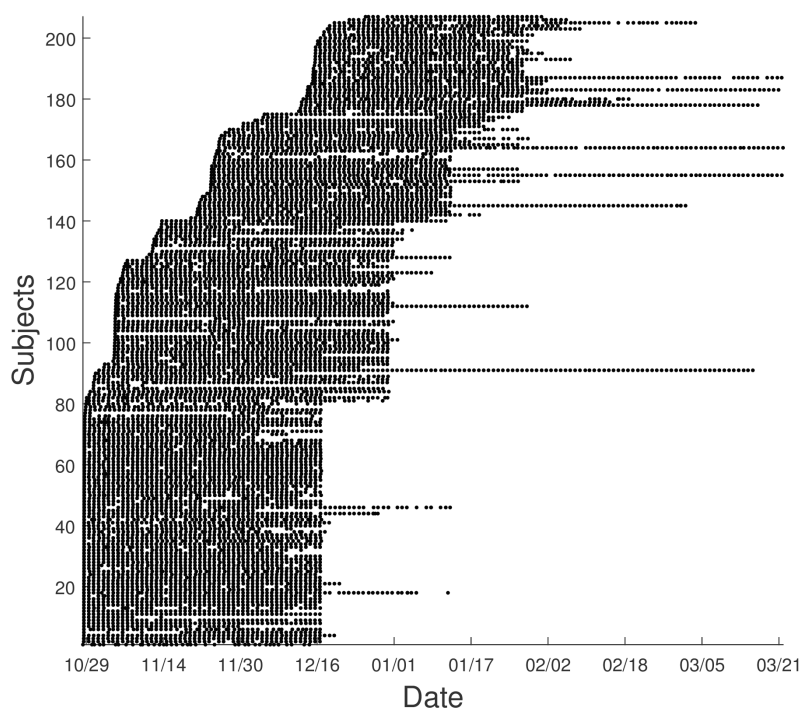
On initial analysis of the data, there were two apparent issues that needed to be resolved. First, some participants had changed their devices during the study, with a number of them reporting on multiple devices at the same time. We detected the change in a participant's mobile phone by tracking their device's MAC address. Out of the 207 participants, 21 changed their phones during the study. When a participant used multiple devices at the same time, we used the data from the first device until there was no EMA data coming from that device, and then switched to the second device.

There were also inconsistent values in sensor and EMA data that needed to be corrected or removed. First, timestamps were stored in different units for some participants, due to difference in phone models. We converted the units of these timestamps to seconds which was used for all other participants. There were

also out-of-range values for sleep times. For example, in some cases, we had negative sleep start or end times; these artifacts were observed in 14 of 207 subjects, with between 1 and 5 erroneous reports for each of these subjects. We removed these instances from the dataset before the analysis. After this processing, our dataset consisted of 207 subjects and a total of 10,649 reports, allowing for a broad characterization of sleep detection.

For the EMA data, there was an extremely high rate of adherence, resulting in little missing data. Of the 207 participants, 10.6% (22) stopped providing labels before the end of the 6-week period. However, many continued to send data after the end of 6 weeks, with 13.0% (27/207) providing more than 60 days of data. The participants' enrollment in the study is depicted in [Figure 2](#). It was surprisingly doable to recruit this large number of subjects over the extended period of time of our study.

Figure 2. Participants' enrollment in the study, sorted based on the enrollment time. Each dot represents an ecological momentary assessment (EMA) report sample we received from the participant. The 4 recruitment waves are evident in the 4 clusters of starting times. Vertical white stripes reflect the time of day when people were less likely to complete their EMA reports (eg, night time). The number of days ranged from 11 to 137 days, with an average of 52.9 days for each participant.



Sleep Detection Results

The average prediction accuracy of the model trained only on sensor features is about 81.8% (95%CI 81.12-82.48), and the addition of time of the day to the feature set increases this accuracy to 88.8% (95%CI 88.41-89.19; [Figure 3](#)). This accuracy, however, is only slightly better than that of the model that has only been trained on time (86.9%; 95%CI 86.68-87.12). These accuracies vary considerably across the subjects, ranging from 65.1% to 97.3% ([Figure 3](#)). Importantly, these results are

in line with those of well-controlled studies for some subjects and dramatically worse for others.

We also compared personal models (those trained with the same participant is predicting) with global models (those trained with other participants' data and predicting a single participant). [Figure 4](#) displays the correlation between the accuracy of personal and global models. Personal models fared better for 80.2% (166/207) of participants; however, the difference between personal and global models was relatively small.

Figure 3. Sleep detection results. (a) Prediction accuracy (error) for global and personal models trained on time feature only, sensor-based features, and all features (see Table 2). Bars show the mean, and error bars show 95% CI. (b) Distribution of the accuracy of global and personal models trained on all features across the participants.

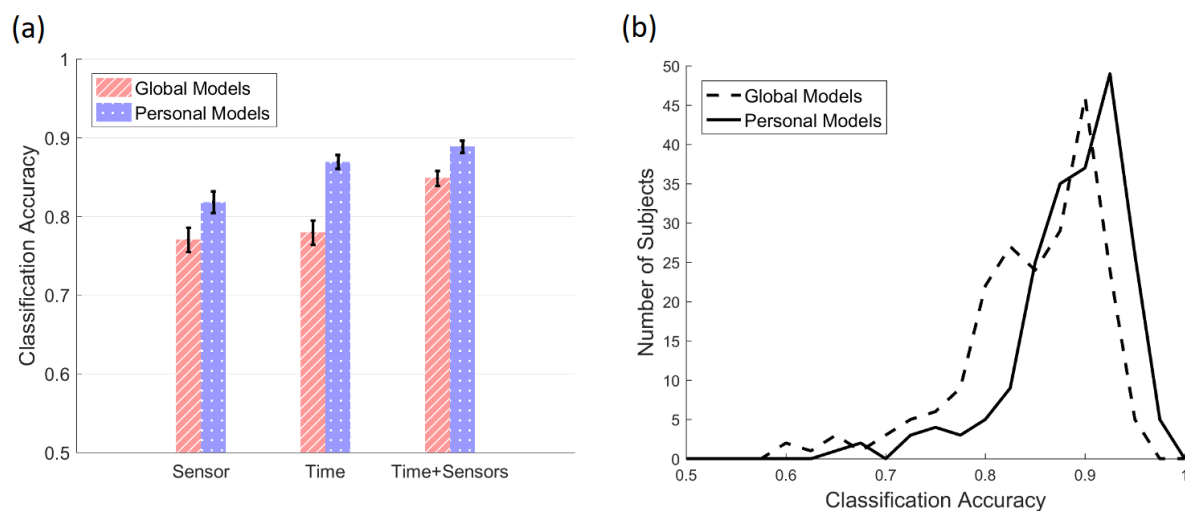
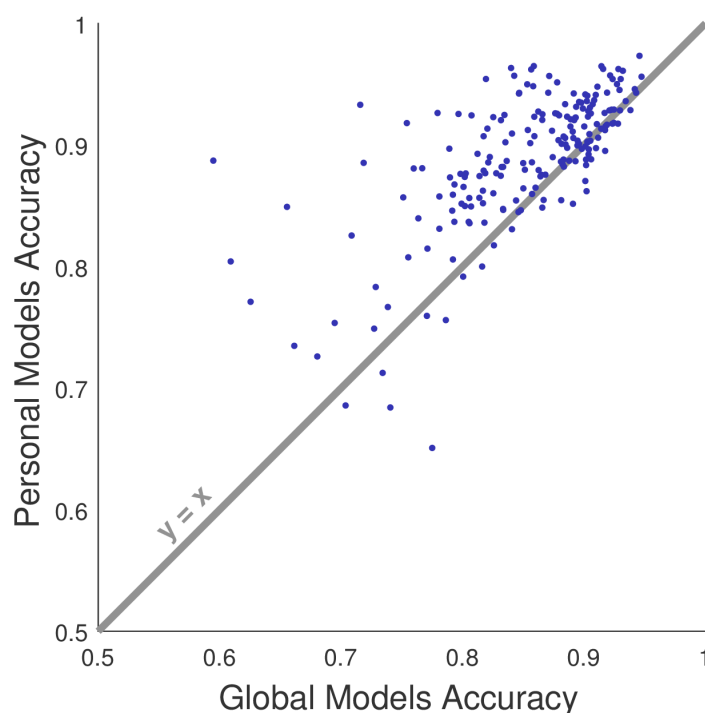


Figure 4. The accuracy of global and personal models across the participants. Each dot is one participant. The dots above the gray line ($y=x$) indicate participants for which personal model performed better than global model, and the dots below indicate the other way around. The correlation between the personal and global model accuracies is high ($r=.685$; $P<.001$).



Where Do Classifiers Fail?

The large variability of prediction accuracies across the participants led us to further explore why prediction fails for specific participants. Here, we looked into various metrics of data quality and investigated their relationship to the classification accuracy. The aim was to find out whether there

are specific data quality issues that caused classifiers to fail, and whether we are able to improve the classification accuracy by resolving those problems.

We found two major data quality issues: missing data and misreports. In the following, we investigate each of these issues.

Missing Data

We estimated the proportion of missing data points in both sensor and EMA data for all participants, and we evaluated the relationship between these and classification accuracy. Although this relationship is complex [Figure 5](#), we found that generally participants with larger proportions of missing sensor or EMA data had lower classification accuracies. Therefore, it seems that missing data is one primary cause for the classifiers' failure.

If missing data is the major cause, the next question is when the missing samples in sensor data occurred. We estimated the proportion of sleep-state samples that were missing, as well as the awake-state samples, and calculated their ratio. As [Figure 6](#) shows, this ratio is significantly higher than 1 for each individual sensor as well as all sensors together. For all sensors, the proportion of missing samples during sleep is almost twice compared with awake.

Figure 5. Dependence of classification accuracy on missing data. (a-b) Accuracy versus the proportion of missing sensor data for global (a) and personal (b) models. Here, we excluded the activity, communication events, and screen state sensors as their absence did not indicate missing data. (c-d) Accuracy versus the proportion of missing ecological momentary assessment (EMA) data for global (c) and personal (d) models. In all four cases, there is a weak but significant, inverse relationship between the classification accuracy and the proportion of missing data. ρ is Spearman rank correlation coefficients, with negative values indicating inverse relationships. One star indicates significance at $P < .05$, two at $P < .01$, and three at $P < .001$.

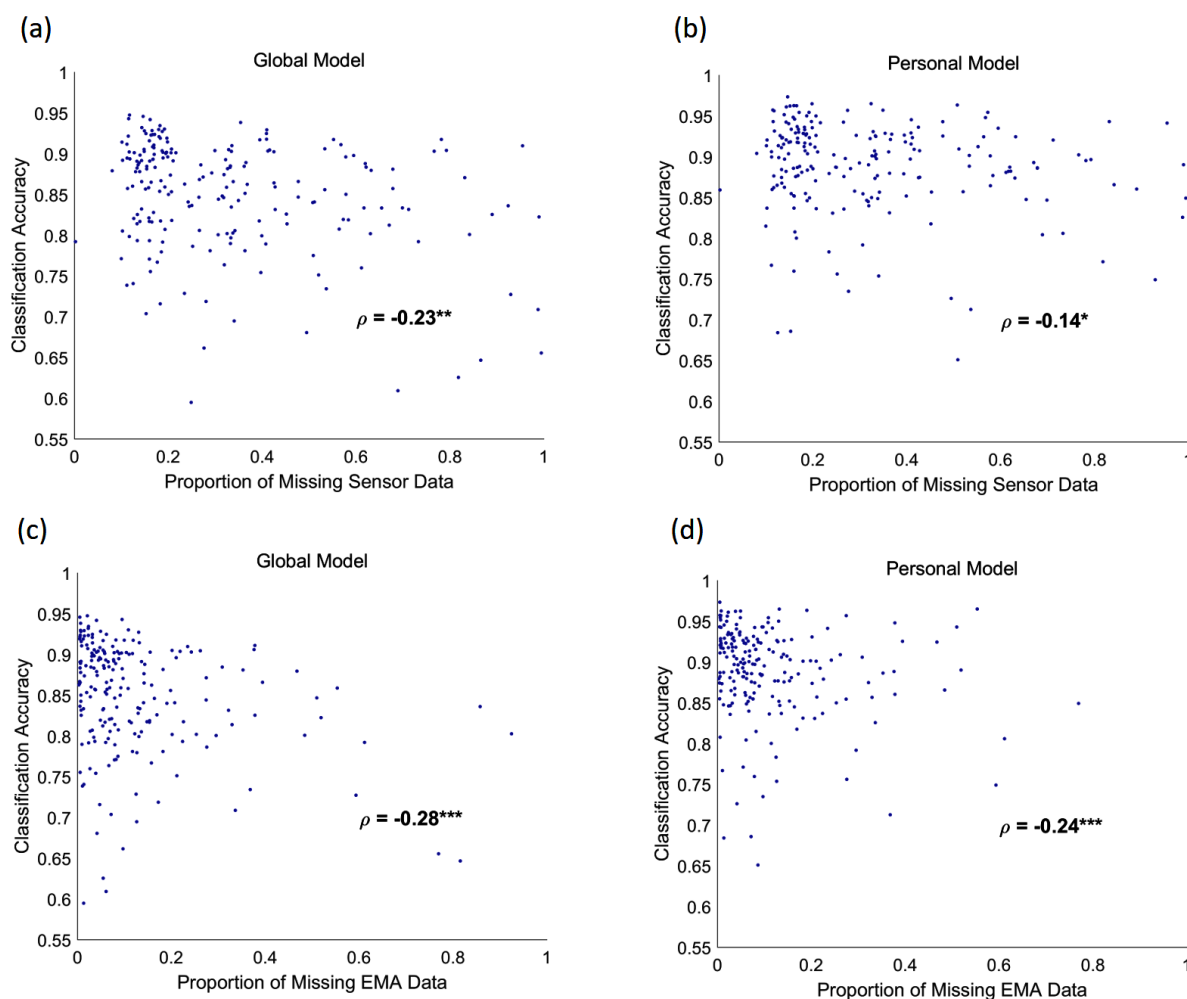
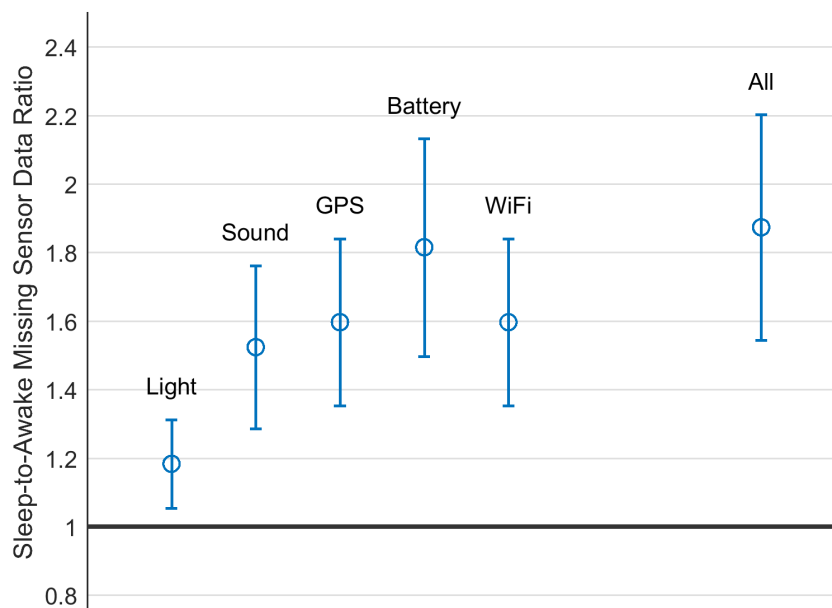


Figure 6. Proportion of missing sensor data during sleep states divided by the proportion of missing sensor data during awake states, across all participants.



Misreports

To investigate the possibility of misreports, we looked at the distribution of sleep start and end times. Although the distribution of sleep end times looks almost normal, sleep start times seem to have an anomaly between 12pm and 3pm (Figure 7). One possible scenario was that participants mislabeled “am” and “pm” times, especially at midnight (12am). Alternatively, these could be short mid-day naps reported instead of previous night’s sleep. To investigate which scenario was more likely,

we also plotted sleep start times versus sleep duration (Figure 7). As evident in this plot, there is a distinct cluster of sleep start times between 12pm and 3pm which is associated with abnormally long (>15 h) sleep duration. Therefore, these data points could not represent mid-day naps, but they are more likely to have been caused by a confusion between “am” and “pm” in reporting sleep start times.

A summary of the data quality issues and their likely causes is shown in Table 3.

Table 3. Summary of the causes for low data quality which likely made the classifiers fail.

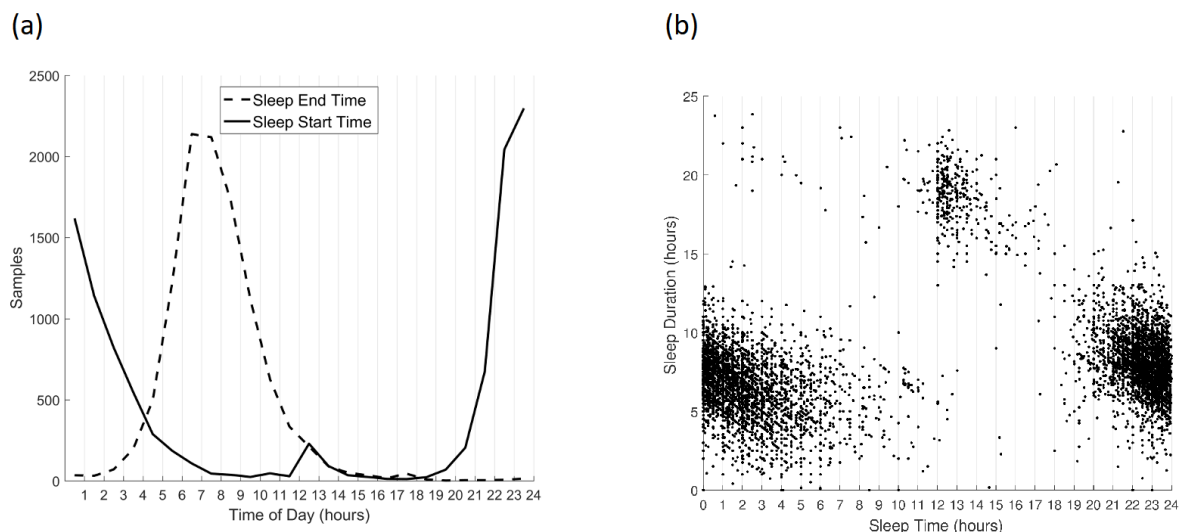
Data source	Issue	Possible causes
Sensors	Missing samples	Mobile phone off, low battery level
Sensors	Missing samples	Purple Robot, operating system, or hardware failure
Sensors	Out of range values	Device model and operating system differences
EMA ^a reports	Missing samples	Participants not reporting
EMA reports	Abnormal values	Participants misreport

^aEMA: ecological momentary assessment.

In addition to missing data and misreports, we also investigated whether the classification accuracy was different between participants with symptoms of depression or anxiety and the ones with no symptoms. We compared four groups of participants: nondepressed and nonanxious, depressed and

nonanxious, nondepressed and anxious, and depressed and anxious. We did not find any significant difference in classification accuracy, for both global and personal models, between any of these groups.

Figure 7. (a) The distribution of sleep start and end times. For sleep start times, there is an anomaly between 12pm and 3pm, which are likely due to the participants confusing “am” and “pm” times. (b) Sleep start times versus sleep duration, shows a distinct cluster (middle top) associated with sleep start times between 12pm and 3pm and abnormally long sleep durations (>15 h).



Improving the Quality of Data

After investigating a number of data quality issues that were likely causing the classifiers to fail in certain situations, we attempted to fix these issues and observed the effects on classification performance. Specifically, we took two steps:

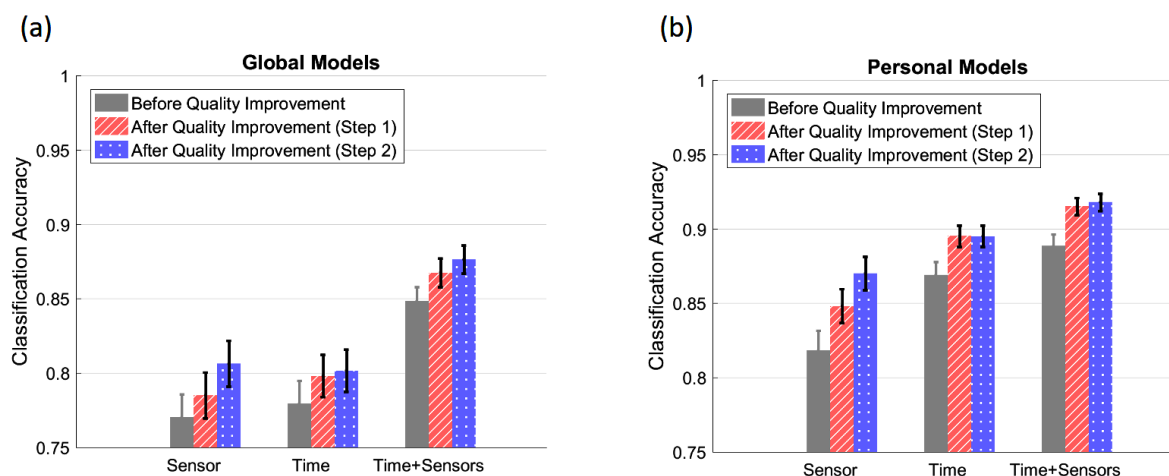
1. When the reported sleep start times were between 12pm and 3pm and their associated sleep duration was longer than 15 h, we changed “pm” to “am.”
2. We removed participants for whom, on average, more than 50% of sensor samples were missing. This consisted of 20.8% (43/207) of participants.

To estimate the proportion of missing sensor data, we excluded the communication events and the screen state sensors, as their absence did not necessarily imply missing samples. After each of these steps, we trained and cross-validation both global and personal sleep prediction models.

The results of the classifier’s performance after improving the data quality is shown in Figure 8. As the figure shows, correcting reported times considerably increased the accuracies of both global and personal classifiers, to 86.7% and 91.5%. Removing participants with large amounts of missing data further increased these accuracies, to 87.6% and 91.8%, respectively, albeit to a slightly smaller degree. It is also interesting to note that the global time-only and sensor-only models have similar performances, which is considerably lower than the performance of the global all-feature model. However, for the personal models, the accuracy of the time-only model is only slightly less than the accuracy of the model trained on both sensor features and time.

Since the amount of missing sensor data was inversely correlated with the classification accuracy, we speculated that adding an extra feature encoding the amount of missing sensor data could be beneficial. However, including these additional features did not improve the accuracy of the classifiers.

Figure 8. (a-b) Sleep detection results after quality improvement, for global (a) and personal (b) models. First, we corrected the reported sleep start and end times, which resulted in significantly higher accuracies (red) for all models. Then, we removed participants for whom the sensor data was missing for more than 50% of the time. This consisted of 43 participants. The resulting accuracies (blue) significantly improved for the sensor-only model, but did not change for the rest.



Prediction of Sleep Start and End Times

Using our predictions of sleep state, we can calculate values for sleep start and end times as well as sleep duration, which can be useful for monitoring clinical processes [5]. We find the closest predicted sleep period to each reported sleep period (from personal models), and examine the bin-indexed errors in predicting the start and end of that sleep period, as well as the total duration of the sleep period. These errors are all calculated on binned data, thus our minimum resolution is the bin size (10 min). We are able to estimate both sleep start and end times with approximately equal accuracy, with an average median absolute deviation (MAD) across participants of 43 min and 38 min, respectively (Figure 9). We are also able to predict sleep duration with similar accuracy, with an average MAD across participants of 58 min (Figure 9). The distribution of these errors are all relatively skew-right, which suggests that poor prediction of a small number of participants substantially affects performance.

Looking at these errors in terms of sleep characteristics can help further elucidate where we make errors. We find that participants

with more extreme, that is, longer or shorter, average sleep durations have larger errors in estimating sleep duration (Figure 10). Specifically, we tend to over-estimate the duration of short sleep periods, and underestimate the duration of long sleep periods. That this occurs even with individual models suggests that, rather than a regression to a global mean, there may be something intrinsically difficult in estimating the durations of extreme sleep periods (or the sleep of those that report extreme sleep periods). We examine the per-participant performance for “outlier” (duration greater or less than two standard deviations (SDs) from the participant’s average sleep duration) and “typical” sleep periods (Figure 10). We find that, for 89% of participants, we can estimate the duration of typical sleep periods within an hour. Interestingly, we can do the same for 38.2% of participants even on their outlier sleep periods, and can estimate outlier sleep periods within 2 h for 62% of participants, suggesting that, while outlier periods are more difficult to predict regularly than most, we do not perform poorly on all outliers as a rule. This suggests difficulties in estimating the sleep duration for particular participants, which may speak to the unique challenges in estimating behavior in large, heterogeneous populations.

Figure 9. (a) Distribution of median absolute deviation (MAD) for predicted sleep start times from true sleep start times over all participants with less than 50% missing data. (b) Distribution of MAD of predicted sleep end times from true sleep end times over all participants with less than 50% missing data. Black line indicates the average MAD over those participants. (c) Distribution of MAD of predicted sleep duration from true sleep duration over all participants with less than 50% missing data. Black lines in (a)-(c) indicate the average MAD over all participants.

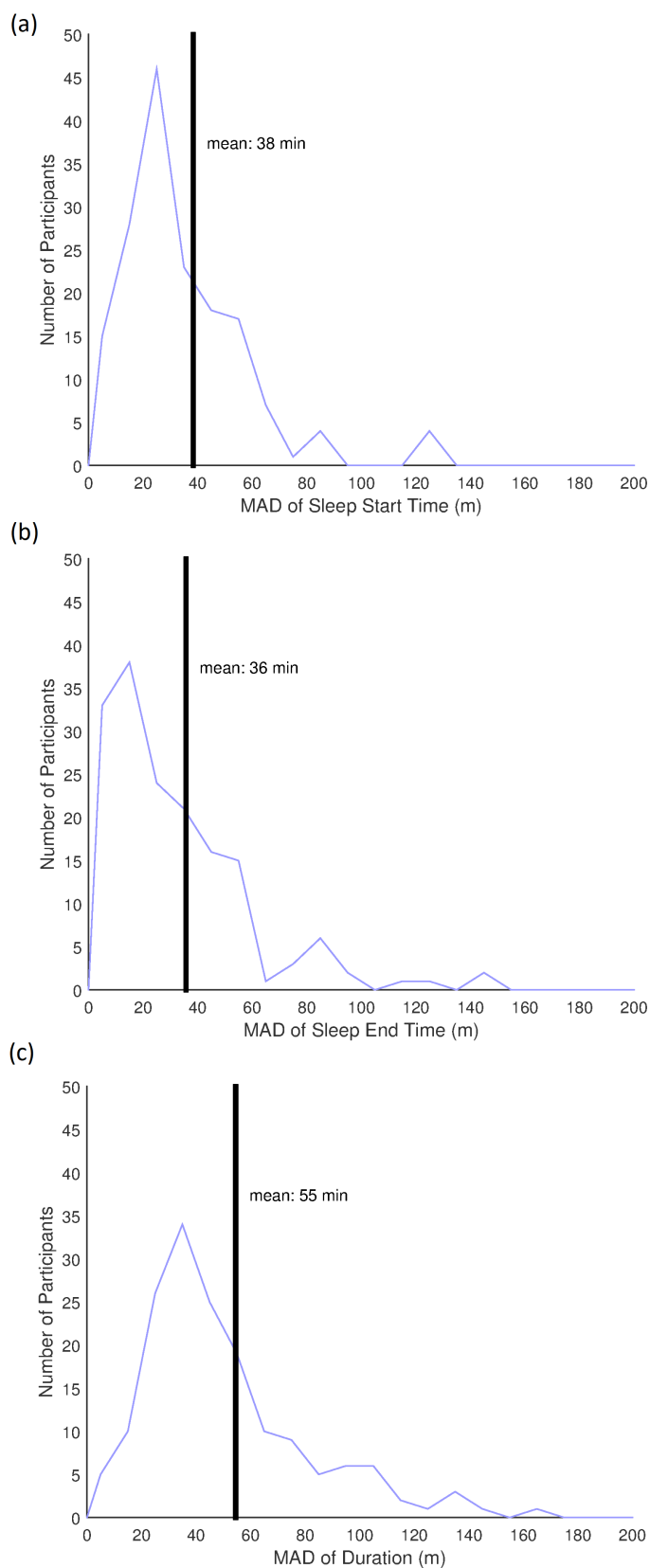
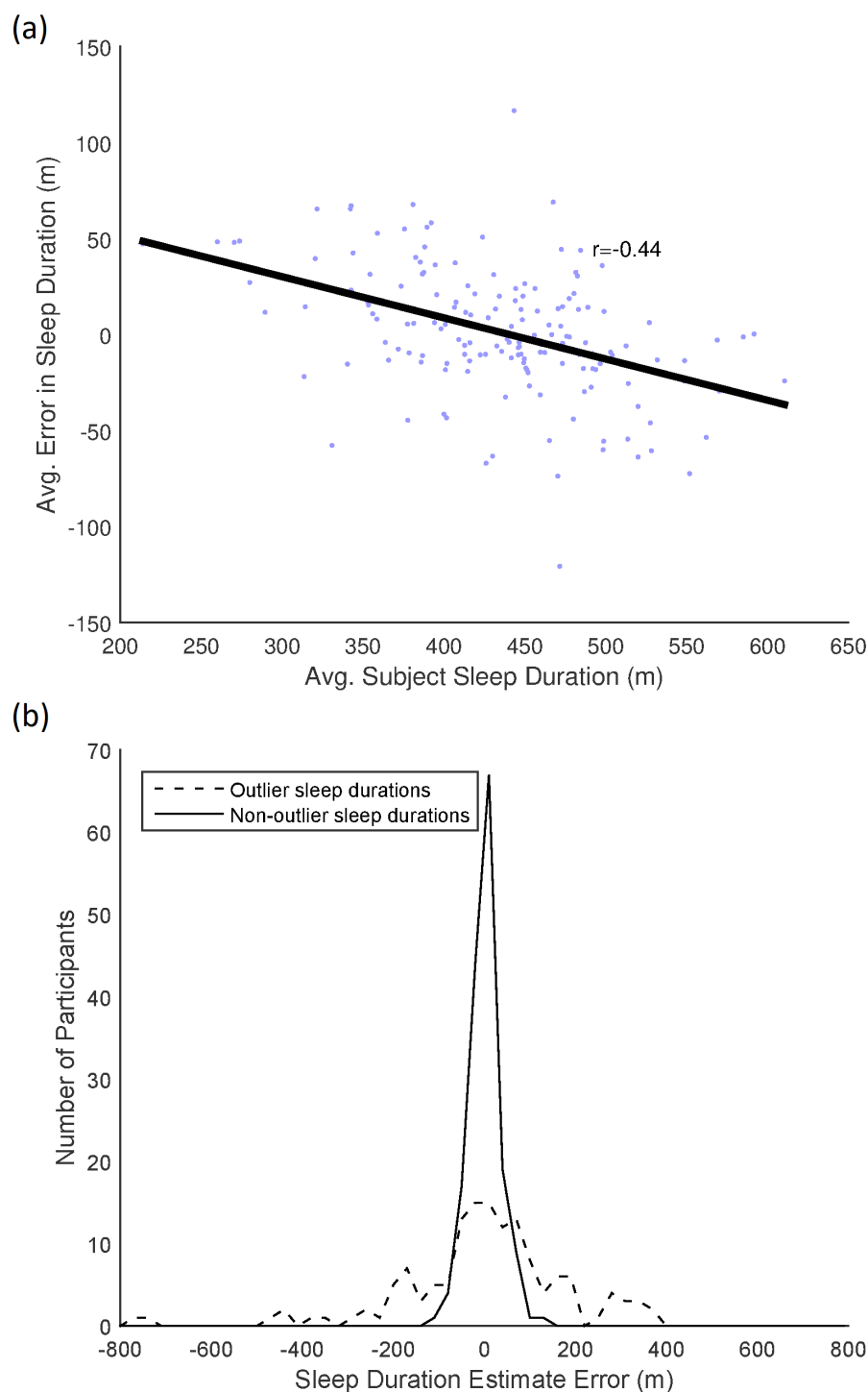


Figure 10. (a) Relationship between average sleep duration and average errors in estimates of sleep duration. Points reflect individual participants with less than 50% missing data, black line represents least-squares regression. (b) Distribution of average sleep duration estimation error over participants with less than 50% missing data for “outlier” (blue) and “nonoutlier” (red) sleep durations. Outlier sleep periods are defined as periods that are two standard deviations shorter or longer than the participant’s average sleep duration, and nonoutlier periods fall within those bounds.



Discussion

Principal Findings

This study was a first step toward bridging initial proof-of-principle studies showing the feasibility of mobile phone-based sleep detection technology with implementation

for a general population in their natural daily-life settings. We divided phone sensor data into 10-min-long windows, and calculated a number of features from them. Then, we trained our models, composed of random forests and HMMs, to predict the state of each window (*sleep* or *awake*). Although the global classifiers trained on all features were able to predict sleep state with 87.6% accuracy, personal models which were trained

separately on each participant had a (significantly) higher accuracy of 91.8%. These numbers were close to the range reported by previous research in more controlled settings [11-13]. Thus our study confirms that sleep tracking via mobile devices is a viable paradigm, and that it can generalize to broad populations when used in daily-life settings.

It is interesting that the performance of personal models trained solely on the time of the day was only slightly lower than the ones trained on all features. This suggests that an individual's sleep patterns do not drastically change day-by-day, and that whether they are asleep or awake at a specific hour can be predicted with good accuracy by time alone. This is an important result, as it shows that the baseline performance, defined by the time-of-day model, is significantly higher than the chance level of 67-71%, calculated by assuming that the average sleep duration across the individuals is between 6 h and 7 h. Therefore, it is necessary that when we report the accuracy of sleep detection algorithms, we compare them to the accuracy of a model only trained on time of the day. This comparison makes the assessment of future sleep detection algorithms easier.

Limitations

There are a number of limitations that should be considered when interpreting the results of our study. First, the self-reported sleep times are not necessarily accurate themselves. In fact, we observed that a number of participants misreported their sleep start times by a substantial amount; when we fixed these reports, the accuracy of the sleep detection algorithm increased substantially. Apart from directly addressable issues like this, there are many other ways in which self-reports might have been inaccurate. Self-reported sleep start times are in general biased, and people tend to over-estimate their sleep duration [28]. Therefore, what we calculate as accuracy is relative to an inaccurate measure. It may be difficult, both here and in other sleep detection work, to calculate the true accuracies of the algorithms.

Second, the parameters of the HMM were adjusted under the assumption that going to sleep and waking up occur only once in 24 h. Although this assumption is true for most people, there are a number of cases for which it is violated. First, most elderly suffer from fragmented sleep [29], during which they can stay awake for a few hours before going to sleep again. Some sleep pattern disorders, such as insomnia or sleep apnea, cause both segmentation of sleep at night and sleepiness during the day [30], likely affecting daytime behavior as well as sleep patterns. In more extreme cases, such as sleepwalking, patients manifest night-time behaviors that resemble daytime, routine activities [31]. Second, we did not ask whether any of our participants worked in different shifts across days, and some of the anomalies seen in the reported times could have been due to shift-working. Finally, a good number of people, almost one-third of Americans, take day-time naps [32]. Therefore, in many cases, the assumption that a person goes to sleep only once in a 24-h period is incorrect, and further understanding of both population and individual sleep habits will be necessary to create more accurate models.

Third, we do not know if any, or which of these participants had a sleep disorder. People with sleep disorders can be

significantly different from the healthy populations in many aspects of their life, which can influence the relationship between mobile phone sensor features and sleep patterns. For example, individuals with disturbed sleep report lower quality of physical functioning, social functioning, vitality, and general health [33]. These differences would likely result in differences in how individuals interact with their mobile phones, thus affecting the data and algorithms for sleep detection. Thus, caution must be applied in generalizing these results to those with atypical sleep patterns.

Finally, our participants were not a perfect sample of the general population in the United States. First, close to 82.6% of our participants were female. Second, we only recruited participants who had WiFi Internet access on their mobile phones. This was important, as the high frequency sensor data can quickly accumulate on the phone and reach the storage limits. Using WiFi to off-load data is energy-efficient and free, unlike using cellular connectivity, which can drain the battery and incur data use fees. For this reason, we recruited participants who had reliable Internet access on their phones. However, with this restriction, it is likely that participants with lower incomes are excluded from our study, who might have different sleep patterns and behavior. Third, 21 participants (10%) changed their phones during the study. Although this may be due to chance, it may also be related to the holiday season, during which people may have received phones as gifts. Finally, we specifically excluded participants with positive screens for several severe psychiatric conditions, which may alter sleep patterns. Thus, it is possible that any or all of these biases reduce the generalizability of these results.

Comparison With Prior Work

We extended previous research in two important ways. First, our sample size was large relative to previous studies and the study participants were more diverse in age, education level, employment, and location. Although a more diverse sample potentially provides a better training dataset for machine learning, it introduces a few problems. First, diversity means more variability in behavior. Unlike college students who have been the participants of a number of previous studies [11,13], participants from the general population do not necessarily use their mobile phones in ways that can help the sleep detection algorithms. For example, mobile phone usage is one feature that is very useful in detecting sleep states since most people use their phones frequently throughout the day. However, phone usage patterns are different across different age groups. Whereas 22% of Americans aged between 18 and 29 years use their phones every few minutes, this number for an older age group of 50-64 years is only 6% [34]. Therefore, a large and diverse sample introduces new challenges to sleep detection algorithms.

The second way in which we extended the previous research was that we did not give participants any instructions regarding the placement of the mobile phone. This meant that participants, for example, could turn their phones off during sleep, or leave it unplugged so that it runs out of battery. As a result, we found that there were many more missing data points during sleep than during awake states. This, however, was not the only scenario that challenged the sleep detection algorithms.

Participants could also leave their phones unattended during the day, or put it in another room when sleeping. Despite all these, the performance of the classifiers is close to, albeit slightly worse than, what has been reported by previous research in more controlled settings.

Conclusions

As mobile phone technology advances, we expect many of the issues we encountered in this study will vanish. For instance, several of the technical problems we experienced will be ameliorated by longer battery life, standardized hardware, and improved app design. Many other limitations, however, will

not be solved by advancing underlying technology. Here we encountered several obstacles, from behaviors that misled algorithms, to sleep patterns unaccounted for by typical models, to inaccurate ground truth data that were due to errors and biases in self-reports rather than technology. Although these obstacles are typically not encountered during demonstrations of sleep detection algorithms, they will likely prove to be impediments to generalized sleep tracking. We believe that mobile phone-based sleep detection technology must tackle these problems in order to become a reliable tool in people's natural life settings.

Acknowledgments

This study was supported by the following National Institute of Health grants: 5R01NS063399, P20MH090318, and R01MH100482. Author SMS was supported by the research grant K08MH102336 from the National Institute of Mental Health.

Conflicts of Interest

None declared.

References

1. Krystal AD. Sleep and psychiatric disorders: future directions. *Psychiatr Clin North Am* 2006 Dec;29(4):1115-30; abstract xi. [doi: [10.1016/j.psc.2006.09.001](https://doi.org/10.1016/j.psc.2006.09.001)] [Medline: [17118285](https://pubmed.ncbi.nlm.nih.gov/17118285/)]
2. Bryant PA, Trinder J, Curtis N. Sick and tired: does sleep have a vital role in the immune system? *Nat Rev Immunol* 2004 Jun;4(6):457-467. [doi: [10.1038/nri1369](https://doi.org/10.1038/nri1369)] [Medline: [15173834](https://pubmed.ncbi.nlm.nih.gov/15173834/)]
3. Gregory AM, Rijdsdijk FV, Lau JY, Dahl RE, Eley TC. The direction of longitudinal associations between sleep problems and depression symptoms: a study of twins aged 8 and 10 years. *Sleep* 2009 Feb;32(2):189-199 [FREE Full text] [Medline: [19238806](https://pubmed.ncbi.nlm.nih.gov/19238806/)]
4. Germain A, Buysse DJ, Nofzinger E. Sleep-specific mechanisms underlying posttraumatic stress disorder: integrative review and neurobiological hypotheses. *Sleep Med Rev* 2008 Jun;12(3):185-195 [FREE Full text] [doi: [10.1016/j.smrv.2007.09.003](https://doi.org/10.1016/j.smrv.2007.09.003)] [Medline: [17997114](https://pubmed.ncbi.nlm.nih.gov/17997114/)]
5. Dahl RE, Lewin DS. Pathways to adolescent health sleep regulation and behavior. *J Adolesc Health* 2002 Dec;31(6 Suppl):175-184. [Medline: [12470913](https://pubmed.ncbi.nlm.nih.gov/12470913/)]
6. Penzel T, Conradt R. Computer based sleep recording and analysis. *Sleep Med Rev* 2000 Apr;4(2):131-148. [doi: [10.1053/smrv.1999.0087](https://doi.org/10.1053/smrv.1999.0087)] [Medline: [12531163](https://pubmed.ncbi.nlm.nih.gov/12531163/)]
7. McCall WV, Erwin CW, Edinger JD, Krystal AD, Marsh GR. Ambulatory polysomnography: technical aspects and normative values. *J Clin Neurophysiol* 1992 Jan;9(1):68-77. [Medline: [1552010](https://pubmed.ncbi.nlm.nih.gov/1552010/)]
8. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep* 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: [10.5665/sleep.1642](https://doi.org/10.5665/sleep.1642)] [Medline: [22294820](https://pubmed.ncbi.nlm.nih.gov/22294820/)]
9. Lauderdale DS, Knutson KL, Yan LL, Rathouz PJ, Hulley SB, Sidney S, et al. Objectively measured sleep characteristics among early-middle-aged adults: the CARDIA study. *Am J Epidemiol* 2006 Jul 01;164(1):5-16. [doi: [10.1093/aje/kwj199](https://doi.org/10.1093/aje/kwj199)] [Medline: [16740591](https://pubmed.ncbi.nlm.nih.gov/16740591/)]
10. Rainee L, Zickuhr K. Pew Research Center.: Pew Research Center; 2015. Americans' views on mobile etiquette URL: <http://www.pewinternet.org/2015/08/26/americans-views-on-mobile-etiquette/> [accessed 2017-03-30] [WebCite Cache ID 6pMDeeivt]
11. Min JK, Doryab A, Wiese J, Amini S, Zimmerman J, Hong JI. Toss'n'turn: smartphone as sleepsleeper quality detector. 2014 Presented at: ACM SIGCHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto, Canada p. 477-486.
12. Chen Z, Lin M, Chen F, Lane ND, Cardone G, Wang R, et al. Unobtrusive sleep monitoring using smartphones. 2013 Presented at: 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops. IEEE; 2013; Venice, Italy p. 145-152.
13. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. Studentlife: assessing mental health, academic performancebehavioral trends of college students using smartphones. 2014 Presented at: Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM; 2014; Seattle, Washington, DC p. 3-14.
14. Cuttone A, Bækgaard P, Sekara V, Jonsson H, Larsen J, Lehmann S. Sensible sleep: a Bayesian model for learning sleep patterns from smartphone events. *PLOS One* 2017;12(1):e0169901.

15. Ohayon M, Carskadon M, Guilleminault C, Vitiello M. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep* 2004 Nov 01;27(7):1255-1273. [Medline: [15586779](#)]
16. Bohn MJ, Babor TF, Kranzler HR. The Alcohol Use Disorders Identification Test (AUDIT): validation of a screening instrument for use in medical settings. *J Stud Alcohol* 1995 Jul;56(4):423-432. [Medline: [7674678](#)]
17. Skinner HA. The drug abuse screening test. *Addict Behav* 1982;7(4):363-371. [Medline: [7183189](#)]
18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [Medline: [11556941](#)]
19. Beck AT, Steer RA, Brown GK. Beck Depression Inventory-II. In: *The Fourteenth Mental Measurements Yearbook*. San Antonio, TX: The Psychological Corporation; 2001.
20. Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck PE, et al. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *Am J Psychiatry* 2000 Nov;157(11):1873-1875. [Medline: [11058490](#)]
21. CBITs TECH. 2015. Purple Robo URL: <https://tech.cbitts.northwestern.edu/purple-robot/> [accessed 2017-03-31] [WebCite Cache ID 6mrcvFT0D]
22. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res* 2015;17(7):e175 [FREE Full text] [doi: [10.2196/jmir.4273](#)] [Medline: [26180009](#)]
23. Heron S. Advanced encryption standard (AES). *Network Security* 2009;2009:8-12.
24. Barnett I, Onnela JP. arXiv. 2016. Inferring mobility measures from GPS traces with missing data URL: <https://arxiv.org/pdf/1606.06328.pdf> [accessed 2017-04-10] [WebCite Cache ID 6pcCh5qu9]
25. Breiman L. Random Forests. *Machine Learning* 2001;45(5):5-32.
26. Breiman L. Bagging predictors. *Machine Learning* 1996;24(2):123-140.
27. Criminisi A, Shotton J, Konukoglu E. Microsoft. 2011. Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning URL: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/decisionForests_MSR_TR_2011_114.pdf [accessed 2017-04-10] [WebCite Cache ID 6pcCyKEsG]
28. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Sleep duration: how well do self-reports reflect objective measures? the CARDIA sleep study. *Epidemiology* 2008;19(6):838-845. [doi: [10.1097/EDE.0b013e318187a7b0](#)]
29. Carskadon MA, Brown ED, Dement WC. Sleep fragmentation in the elderly: relationship to daytime sleep tendency. *Neurobiol Aging* 1982;3(4):321-327. [Medline: [7170049](#)]
30. Stepanski E, Lamphere J, Badia P, Zorick F, Roth T. Sleep fragmentation and daytime sleepiness. *Sleep* 1984;7(1):18-26. [Medline: [6718922](#)]
31. Ohayon MM, Guilleminault C, Priest RG. Night terrors, sleepwalking, and confusional arousals in the general population: their frequency and relationship to other sleep and mental disorders. *J Clin Psychiatry* 1999 Apr;60(4):268-76; quiz 277. [Medline: [10221293](#)]
32. Taylor P. Pew Research Center. 2009. Nap time URL: <http://www.pewsocialtrends.org/2009/07/29/nap-time/> [accessed 2017-03-30] [WebCite Cache ID 6mrdEseqh]
33. Reimer MA, Flemons WW. Quality of life in sleep disorders. *Sleep Med Rev* 2003 Aug;7(4):335-349. [Medline: [14505600](#)]
34. Newport F. Gallup. 2015. Most U.S. smartphone owners check phone at least hourly URL: <http://www.gallup.com/poll/184046/smartphone-owners-check-phone-least-hourly.aspx> [accessed 2017-03-31] [WebCite Cache ID 6mrd7tDQA]

Abbreviations

API: application program interface
AUDIT: alcohol use disorder identification test
BDI: Beck depression inventory
DAST: drug abuse screening test
EEG: electroencephalogram
EMA: ecological momentary assessment
FFT: Fast Fourier Transform
FPG: Focus Pointe Global
GAD: generalized anxiety disorder
HMM: hidden Markov model
MAC: media access control
MAD: median absolute deviation
MDQ: mood disorder questionnaire
PHQ: patient-health questionnaire
USB: Universal Serial Bus

Edited by R Calvo; submitted 14.10.16; peer-reviewed by YS Bin, Y Ma, B Price; comments to author 17.11.16; revised version received 03.01.17; accepted 04.03.17; published 18.04.17.

Please cite as:

Saeb S, Cybulski TR, Schueller SM, Kording KP, Mohr DC

Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles

J Med Internet Res 2017;19(4):e118

URL: <http://www.jmir.org/2017/4/e118/>

doi: [10.2196/jmir.6821](https://doi.org/10.2196/jmir.6821)

PMID:

©Sohrab Saeb, Thaddeus R Cybulski, Stephen M Schueller, Konrad P Kording, David C Mohr. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 18.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach

Jennifer K Carroll¹, MPH, MD; Anne Moorhead², MSc, MA, MICR, CSci, FNutr (Public Health), PhD; Raymond Bond³, PhD; William G LeBlanc¹, PhD; Robert J Petrella⁴, MD, PhD, FCFP, FACSM; Kevin Fiscella⁵, MPH, MD

¹Department of Family Medicine, University of Colorado, Aurora, CO, United States

²School of Communication, Ulster University, Newtownabbey, United Kingdom

³School of Computing & Maths, University of Ulster, Newtownabbey, United Kingdom

⁴Lawson Health Research Institute, Family Medicine, Kinesiology and Cardiology, Western University, London, ON, Canada

⁵Family Medicine, Public Health Sciences and Community Health, University of Rochester Medical Center, Rochester, NY, United States

Corresponding Author:

Jennifer K Carroll, MPH, MD

Department of Family Medicine

University of Colorado

Mail Stop F496

12631 E. 17th Ave

Aurora, CO, 80045

United States

Phone: 1 303 724 9232

Fax: 1 303 724 9747

Email: jennifer.2.carroll@ucdenver.edu

Abstract

Background: Mobile phone use and the adoption of healthy lifestyle software apps (“health apps”) are rapidly proliferating. There is limited information on the users of health apps in terms of their social demographic and health characteristics, intentions to change, and actual health behaviors.

Objective: The objectives of our study were to (1) to describe the sociodemographic characteristics associated with health app use in a recent US nationally representative sample; (2) to assess the attitudinal and behavioral predictors of the use of health apps for health promotion; and (3) to examine the association between the use of health-related apps and meeting the recommended guidelines for fruit and vegetable intake and physical activity.

Methods: Data on users of mobile devices and health apps were analyzed from the National Cancer Institute’s 2015 Health Information National Trends Survey (HINTS), which was designed to provide nationally representative estimates for health information in the United States and is publicly available on the Internet. We used multivariable logistic regression models to assess sociodemographic predictors of mobile device and health app use and examine the associations between app use, intentions to change behavior, and actual behavioral change for fruit and vegetable consumption, physical activity, and weight loss.

Results: From the 3677 total HINTS respondents, older individuals (45-64 years, odds ratio, OR 0.56, 95% CI 0.47-68; 65+ years, OR 0.19, 95% CI 0.14-0.24), males (OR 0.80, 95% CI 0.66-0.94), and having degree (OR 2.83, 95% CI 2.18-3.70) or less than high school education (OR 0.43, 95% CI 0.24-0.72) were all significantly associated with a reduced likelihood of having adopted health apps. Similarly, both age and education were significant variables for predicting whether a person had adopted a mobile device, especially if that person was a college graduate (OR 3.30). Individuals with apps were significantly more likely to report intentions to improve fruit (63.8% with apps vs 58.5% without apps, $P=.01$) and vegetable (74.9% vs 64.3%, $P<.01$) consumption, physical activity (83.0% vs 65.4%, $P<.01$), and weight loss (83.4% vs 71.8%, $P<.01$). Individuals with apps were also more likely to meet recommendations for physical activity compared with those without a device or health apps (56.2% with apps vs 47.8% without apps, $P<.01$).

Conclusions: The main users of health apps were individuals who were younger, had more education, reported excellent health, and had a higher income. Although differences persist for gender, age, and educational attainment, many individual sociodemographic factors are becoming less potent in influencing engagement with mobile devices and health app use. App use was associated with intentions to change diet and physical activity and meeting physical activity recommendations.

KEYWORDS

smartphone; cell phone; Internet; mobile applications; health promotion; health behavior

Introduction

As of 2015, nearly two-thirds (64%) of the American public owned a mobile phone, which is an increase from 35% in 2011 [1]. It is estimated that 90% of the worldwide population will own a mobile phone by 2020 [1]. Current UK data reveals that mobile phone usage is increasing as 66% adults aged more than 18 years owned a mobile phone in 2015, up from 61% in 2014 [2]. Mobile phone ownership is higher among younger people, with 77% ownership for those aged 16-24 years [3]. Although mobile phone ownership is especially high among younger persons and those with higher educational attainment and income [4], those with lower income and educational attainment are now likely to be “mobile phone dependent,” meaning that they do not have broadband access at home and have few other options for Web-based access other than via mobile phone.

As mobile phone ownership rapidly proliferates, so does the number of mobile phone software apps grown in the marketplace [5]. Apps focused on health promotion are quite common: more than 100,000 health apps are available in the iTunes and Google Play stores [6]. This staggering number speaks to both the huge market and ongoing demand for new tools to help the public manage their diet, fitness, and weight-related goals, and the limitations of the current health care system to provide such resources. A recent study found that 53% of cell phone users owned a smartphone—this translates to 45% of all American adults—and that half of those (or about 1 in 4 Americans) have used their phone to look up health information [7]. There is increasing usage of health apps among health care professionals, patients and general public [8], and apps can play a role in patient education, disease self-management, remote monitoring of patients, and collection of dietary data [9-12]. Using mobile phones and apps, social media also can be easily accessed, and increasing numbers of individuals are using social media for health information with reported benefits and limitations [8].

Despite the massive uptake in mobile phone ownership and health app usage and their potential for improving health, important limitations of health apps are the lack of evidence of clinical effectiveness, lack of integration with the health care delivery system, the need for formal evaluation and review, and potential threats to safety and privacy [6,13-17]. Although previous studies have described the sociodemographic factors associated with mobile health and app use [7,18,19], it is a rapidly changing field with the most recent published reports reflecting data at least four to five years old. Additionally, there is a lack of information on the users of health apps in terms of their sociodemographic and health characteristics and health behaviors. Furthermore, to our knowledge, there have been no previous publications reporting on the association between the use of health apps, behavioral or attitudinal factors (ie, readiness or intentions to change), and health outcomes. This information is important for future health-improving initiatives and for

identifying appropriate use of health apps among population groups.

Therefore, the aim for our study was 3-fold: (1) to describe the sociodemographic characteristics associated with health app use in a recent US nationally representative sample; (2) to assess the attitudinal and behavioral predictors of the use of health apps for health promotion; and (3) to examine the association between the use of health-related apps and meeting the recommended guidelines for fruit and vegetable intake and physical activity. Given the increasing focus on new models for integrating technology into health care and the need to expand the evidence base on the role of health apps for health and wellness promotion, these research questions are timely and relevant to inform the development of health app interventions.

Methods**Data Source**

The National Cancer Institute's Health Information National Trends Survey (HINTS) is a national probability sample of US adults that assesses usage and trends in health information access and understanding. HINTS was first administered in 2002-2003 as a cross-sectional survey of US civilians and noninstitutionalized adults. It has since been iteratively administered in 2003, 2005, 2008, 2011, 2012, 2013, and 2014. We used data from HINTS 4 Cycle 4 data released in June 2015, which corresponded to surveys administered in August-November, 2014. Publicly available datasets and information about methodology are available at the HINTS website [20]. The 2014 iteration reported herein contained questions about whether participants used mobile phone or tablet technology and software apps for health-related reasons. The overall response rate was 34.44%. This study was reviewed and qualified for an Exemption by the American Academy of Family Physicians Institutional Review Board.

Participants

A total of 3677 individuals completed the 2014 HINTS survey. From this sample, 148 respondents were considered partial completers, in that they completed 50%-79% of the questions in Sections A and B. We included all 3677 respondents in our analysis. We used sampling weights from the HINTS dataset that were incorporated into the regression analyses.

Measures**Demographics**

We used participants' self-report of their age, sex, race, ethnicity, income, level of education, English proficiency, height, and weight. We converted height and weight into body mass index (BMI), using weight (kg)/height (m²)×10,000, and classified participants as obese (≥30), overweight (29.9-26), or normal weight or underweight (<26).

Usage of Mobile Devices and Health Apps

We used participants' responses to the 3 questions to characterize the distribution of subjects who used health-related software apps on their mobile devices. The participants were asked whether they had a tablet computer, smartphone, basic cell phone only, or none of the above. We examined factors for those with and without mobile devices, since previous studies have shown differences in seeking health information on the Internet related to access (eg, availability of a computer) [21,22], HINTS dataset is a nationally representative sample, and we wished to put our findings on app use in the larger population context. We categorized participants who had a mobile phone or a tablet device under the label "Device+." Similarly, participants who did not report having a mobile phone or a tablet device were labeled "Device-." Of the Device+ group, we also categorized them according to whether they had health apps on their device (Device+/App+) or did not have health apps on their device (Device+/App-).

Fruit and Vegetable Intake

We assessed fruit and vegetable intake using the 2 questions: amount of fruit consumed per day and amount of vegetables consumed per day (7 response options for each ranging from none to >4 cups per day). We reclassified the response options for both questions into a single dichotomous outcome variable, that is, the subject either (1) meets recommendations for fruit or vegetables (4 or more cups for each) or (2) does not meet recommendations for fruit or vegetables (all other response options). Fruit and vegetable scores were analyzed separately.

Physical Activity

We assessed physical activity using the 2 questions: (1) in a typical week how many days do you do any physical activity or exercise of at least moderate intensity, such as brisk walking, bicycling at a regular pace, and swimming at a regular pace? (8 response options ranging from none to 7 days per week) and (2) on the days that you do any physical activity or exercise of at least moderate intensity how long do you do these activities? (2 response options for minutes and hours). We reclassified the response options into a single dichotomous outcome variable for physical activity, that is, whether the subject (1) met physical activity recommendations (≥ 150 minutes per week) or did not meet the physical activity recommendations (< 150 minutes per week).

Intentions to Change Behavior

We examined participants' intentions to change behavior based on the 5 questions (all with yes or no responses): At any time

in the last year, have you intentionally tried to (1) increase the amount of fruit or 100% fruit juice you eat or drink, (2) increase the amount of vegetables or 100% vegetable juice you eat or drink, (3) decrease the amount of regular soda or pop you usually drink in a week, (4) lose weight, and (5) increase the amount of exercise you get in a typical week?

Statistical Analysis

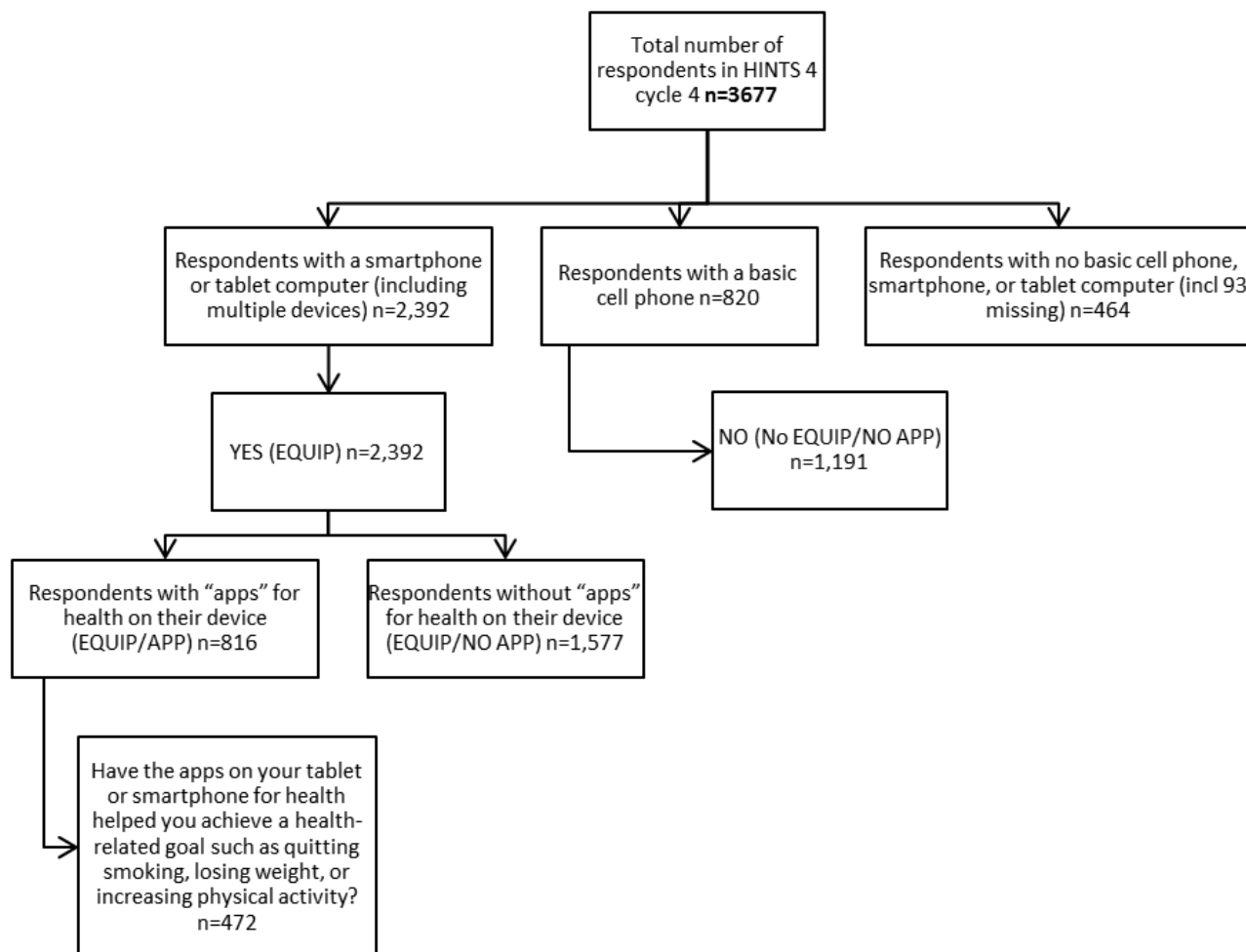
The outcome variable (OUTCOME) was a composite derived from 3 survey variables: (1) *own a smartphone (an Internet-enabled mobile phone "such as iPhone android BlackBerry or Windows phone" differentiated from a "basic cell phone," hereafter referred to as "mobile phone") or device*, (2) *have health apps on mobile phone or device*, and (3) *use of health apps*. *Own a mobile phone or device* was a system-supplied derived variable to categorize responses given to question B4 (possession of a mobile phone or tablet device). *Have health apps on mobile phone or device* (question B5) asked about health apps on a tablet or mobile phone. *Use of health apps* (question B6a) asked whether the apps on a mobile phone or tablet helped in achieving a health-related goal. OUTCOME consisted of 3 levels: Device-/App- (33.2% of respondents), Device+/App- (44% of respondents), and Device+/App+ (22.77% of respondents). Device referred to having a tablet or mobile phone, and App referred to having a health-related app that ran on a tablet or mobile phone. A total of 93 of 3677 respondents were unable to be classified due to missing data. These people were not used in the analyses. To assess the relationship between OUTCOME and the demographic or health behavior variables, simple unweighted 2-way crosstab tables were generated and tested with a chi-square test of association. We used a cutoff of $P < .05$ to determine statistical significance for all analyses.

We used the R programming language (R-Studio) and SPSS (SPSS Inc) for all data modeling and analysis carried out in this study.

Results

Principal Findings

From the 3677 total HINTS respondents, 3584 answered questions about whether or not they had a tablet computer or mobile phone, or used apps. Figure 1 shows the participants in this study.

Figure 1. Health Information National Trends Survey (HINTS) respondents' use of mobile phones, tablets, and apps.

Demographic Variables Associated With App Use

Table 1 compares respondents grouped into Device+/App+, Device+/App-, and Device-, according to sociodemographic characteristics. As shown in Table 1, those who used health apps (compared with those who either did not have apps or did

not have the necessary equipment) were more likely to be younger, live in metropolitan areas, have more education, have higher income, speak English well, be Asian, and report excellent health. There was no significant association between both BMI and smoking status and app use.

Table 1. Demographic variables associated with app usage.

Demographic variables	Device+/App+ n ^{b,c} (%) ^d	Device+/App- n (%)	Device- n (%)	P value
Sex (female vs male; n ^{a,c} =3519)	808 (51.62)	1555 (50.23)	1156 (55.29)	.39
Age (18-44 years vs 45+ years; n=3415)	782 (65.62)	1552 (52.25)	1111 (21.92)	<.001
Education (high school or less vs some college or college graduate, n=3444)	788 (12.72)	1535 (27.95)	1121 (51.82)	<.01
Income (US \$0-49,999 vs 50,000 or greater; n=3530)	808 (31.72)	1560 (42.20)	1162 (75.12)	<.001
Race or ethnicity (white vs other; n=3273)	763 (71.85)	1453 (78.52)	1057 (83.68)	<.01
BMI (normal vs overweight, obese; n=3420)	782 (33.71)	1524 (36.98)	1114 (33.82)	.49
Metro vs nonmetro (n=3584)	816 (92.10)	1577 (85.67)	1191 (78.93)	<.001
Speak English (very well or well vs not well or not at all; n=3584)	759 (99.37)	1497 (97.13)	1089 (90.37)	<.001
Self-rated health (excellent, very good, good vs fair or poor; n=3477)	795 (92.85)	1544 (89.74)	1138 (74.99)	<.001

^aThe sample sizes (n's) listed for each variable in the far left column represent the total number of respondents across all app-usage categories (Device+/App+, Device +/App-, Device-) who answered that question.

^bThe sample sizes (n's) listed for each variable within each cell represent the total number of respondents within a given app-usage category (either Device+/App+, Device +/App-, or Device-) who answered that question.

^cSample sizes vary for each variable due to missing values.

^dPopulation estimates were used for the numerators and denominators in the calculation of percentages. Row percentages do not add to 100%, as the table shows percentages within a given app-usage category (Device+/App+, Device +/App-, or Device-).

Association Between the Use of Apps and Intentions to Change Diet, Perform Physical Activity, and Lose Weight

Table 2 shows the association between the use of apps (versus Device+/App- or Device-) with intentions to change diet,

perform physical activity, or lose weight. As Table 2 shows, participants with apps were significantly more likely to report intentions to improve fruit ($P=.01$) and vegetable consumption ($P<.01$), physical activity ($P<.01$), and weight loss ($P<.01$) compared with those in the Device+/App- or Device- groups.

Table 2. Association between the usage of apps for health-related goal and intentions to change diet, physical activity, or lose weight.

Health-related intention	Device+/App+ n (%)	Device+/App- n (%)	Device- n (%)	P value ^a
Increase fruit	545 (63.76)	885 (58.50)	654 (48.94)	.01
Increase vegetables	621 (74.92)	1023 (64.26)	717 (50.02)	<.01
Decrease soda	630 (84.96)	1135 (82.76)	754 (77.36)	.06
Increase physical activity	707 (82.99)	1237 (65.42)	769 (49.94)	<.01
Lose weight	692 (83.36)	1259 (71.75)	881 (60.02)	<.01

^aSignificance between participants with apps (Device+/App+) compared with those not using apps or devices (Device+/App- or Device- groups).

Association Between the Use of Apps and Meeting Recommendations for Fruit and Vegetable Intake and Physical Activity

Table 3 shows the association between the use of apps (versus Device+/App- or Device-) and meeting the recommendations

for fruit and vegetable intake and physical activity. Participants in the Device+/App+ group were not significantly more likely to meet recommendations for fruit and vegetables compared with those in the Device+/App- or Device- groups; however, they were significantly more likely to exercise more than 2 hours per week.

Table 3. Association between the use of apps for health-related goal and meeting recommendations for fruit and vegetables and physical activity.

Percent respondents meeting recommendations	Device+/App+ n (%)	Device+/App- n (%)	Device- n (%)	P value ^a
Fruit	804 (8.87)	1560 (7.96)	1161 (5.43)	.25
Vegetables	809 (4.81)	1557 (3.01)	1155 (3.48)	.27
Physical activity	801 (56.23)	1552 (47.79)	1144 (37.69)	<.01

^aSignificance between participants with apps (Device+/App+) compared with those not using apps or devices (Device+/App- or Device- groups).

Predicting Health App Adoption Only (Binary Classification)

Table 4 presents the statistically significant odds ratios (ORs) as derived using multivariate logistic regression when applied to the entire dataset. As expected, those aged 45-64 years (OR 0.56) or 65+ years (OR 0.19) had a reduced likelihood of having adopted health apps relative to younger persons. It also showed that males were slightly less likely (OR 0.80) to have a health

app compared with females. The most significant finding was the confirmation that graduates had significantly higher odds (OR 2.83) of having a health app especially when compared with those who had attained an education that was considered “less than high school” (OR 0.43). The results also indicated that the category “completed high school only” had no predictive ability for estimating whether a person had adopted a health app.

Table 4. Statistically significant odds ratios derived using multivariate logistic regression when applied to the entire dataset for predicting health app adoption only.

Variable	Odds ratio (95% CI)	P value
Age (45-64 years)	0.56 (0.47-0.68)	<.001
Age (65+ years)	0.19 (0.14-0.24)	<.001
Sex (male)	0.80 (0.66-0.94)	<.01
Education (college graduate or higher)	2.83 (2.18-3.70)	<.001
Education (less than high school)	0.43 (0.24-0.72)	<.01
Education (some college)	1.70 (1.30-2.26)	<.01
Race (black)	1.25 (0.99-1.55)	.05

Predicting Mobile Technology Adoption Only (Binary Classification)

Table 5 presents the statistically significant ORs that increased or decreased the likelihood that a person had adopted mobile technology (tablet or mobile phone). Interestingly, there were no statistically significant ORs for gender or racial categories.

However, similar to predicting health app adoption, both age and education were significant variables for predicting whether a person had adopted a mobile device, especially if that person was a college graduate (OR 3.30). In addition, the results indicated that the category “completed high school only” had no predictive ability for estimating whether a person had adopted a mobile device.

Table 5. Statistically significant odds ratios derived using multivariate logistic regression when applied to the entire dataset for predicting mobile device adoption only.

Variable	Odds ratio (95% CI)	P value
Age (45-64 years)	0.35 (0.28-0.45)	<.001
Age (65+ years)	0.09 (0.07-0.12)	<.001
Education (college graduate or higher)	3.30 (2.65-4.11)	<.001
Education (less than high school)	0.51 (0.37-0.70)	<.001
Education (some college)	1.87 (1.50-2.32)	<.001

Discussion

Principal Findings

Our first objective was to describe the sociodemographic and health behavior characteristics associated with health app use in a recent US nationally representative sample. Consistent with

previous findings [7], we found that those who were younger, had more education, reported excellent health, and had a higher income were more likely to use health apps. Our predictive modeling using multivariate logistic regression showed that education, sex, gender, and race were only mildly to moderately potent in predicting mobile technology adoption.

Our second objective was to assess the behavioral and attitudinal predictors of the use of health apps for health promotion. We found that participants with apps were also more likely to report intentions to improve fruit and vegetable consumption, physical activity, and weight loss. Finally, the third objective was to examine the association between the use of health-related apps and meeting the recommended guidelines for fruit and vegetable intake and physical activity. We found that participants in the health apps group were significantly more likely to meet recommendations for physical activity compared with those without a device or health apps.

Comparison With Prior Work

This study shares some similarities with previous HINTS analyses. For example, McCully et al [19] reported that users of the Internet for diet, weight, and physical activity tended to be younger and more educated and that Internet use for these purposes was more likely to be associated with higher fruit and vegetable intake and moderate exercise. However in that study, women were no more likely than men to use the Internet for diet, weight, and physical activity, which was different from our findings. In that study, minorities were more likely to use the Internet; in our study, we found no such association. Consistent with our findings, Kontos et al found that males, those with lower education, and older US adults were less likely to engage in a number of eHealth activities [18]. Similar to their findings 3 years ago, our findings pointed to differences by education for app use for health promotion.

The association between app use, intention to change lifestyle behaviors, and actually meeting recommendations for healthy lifestyle factors is interesting and could be due to several reasons. First, it is possible that there are preexisting differences in individuals who engage with health apps compared with those who do not. Users of health apps may have greater motivation and interest in changing their diet, weight, or physical activity. A recent review found that very few available apps provided evidence-based support to meet lifestyle recommendations [13]. It could also be that app users are engaging with health apps to help them simply track or self-manage differently than their counterparts; thus, there could be differences in preferences or needs. Due to the correlational nature of the data, we cannot draw conclusions about the relationships or causal pathways. Similar observations have been reported in a study of users of the Internet for diet, weight, and physical activity promotion [19].

The prevalence of app usage in our study was 22% (816/3677). This is a doubling from the Kontos study in which 11.7% downloaded info onto a mobile device. Although the questions in these 2 HINTS datasets were worded differently (eg, “downloaded” is broader and not referring exclusively to downloading an app), it suggests that demand for apps continues to rise and offers potential for reaching a growing segment of the US population.

Our findings provide evidence for educational, age, and gender differences in the use of mobile devices and health apps. Educational attainment, age, and gender have been previously shown to be important predictors of adoption of mobile devices and apps [18]. Educational attainment appears more important than other variables commonly used as proxies for socioeconomic position (eg, income, race or ethnicity). The reasons for the educational differences are unclear, but may reflect skills and confidence with the use of devices and possibly social norms related to perceived value. Similarly, age likely reflects both social norms and cohort effects, that is, exposure during younger ages to these devices and apps. The reasons for gender differences are less clear, but may reflect differences in health-seeking behavior, and interest and participation in healthy lifestyle interventions generally.

Limitations

This study had limitations that should be kept in mind when interpreting results. First, HINTS is a cross-sectional survey; although it is a nationally representative cohort of individuals, we were not able to evaluate the trends in an individual's health app use over time. There is the possibility of unmeasured confounding, that is, unidentified factors that might be associated with app use and intentions or health behaviors, which could influence the interpretation of results. Although the results showed association, it did not indicate a causal relationship. This study could not answer the question of whether more motivated individuals sought out apps, or whether app use improved motivation and health outcomes. Furthermore, some of the cells for subgroups were small, thereby limiting the generalizability of some of the subanalyses. As with all cross-sectional surveys, this was a study of association, not causation. Finally, we were limited by the questions that were asked in the HINTS survey. For example, we did not have details about specific health apps or features of apps used, the intensity of use, whether the apps were interactive and linked to other health promotion supports (eg, telehealth), and other strategies used for health behavior change. Despite these limitations, the results did identify areas for future research and add to the knowledge base about predictors of the use of health apps.

Conclusions

Compared with previous studies, many individual sociodemographic factors are becoming less important in influencing engagement with mobile devices and health app use; however, differences persist for gender, age, and educational attainment. As health care undergoes technological transformation with its electronic health records systems and individuals' access to their records, there are many opportunities for clinical care models to be expanded and improved, perhaps through the use of apps as a means for sharing data, although this remains an unanswered question. This study contributes to the literature by providing up-to-date information on populations most and least likely to use health apps to guide clinical interventions, commercial developers, and public health programs when designing eHealth technology.

Conflicts of Interest

None declared.

References

1. Ericsson. 2015. Ericsson Mobility Report: On the pulse of the networked society URL: <http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf> [accessed 2016-01-22] [WebCite Cache ID 6ejSFiiicz]
2. OfCom. 2015. Smartphone usage URL: <http://media.ofcom.org.uk/facts/> [accessed 2016-01-22] [WebCite Cache ID 6ejSSqrN7]
3. OfCom. Belfast: OfCom; 2014. Telecommunications facts and figures URL: <https://www.ofcom.org.uk/> [accessed 2017-02-27] [WebCite Cache ID 6ob5yNLrU]
4. Smith A. Pew Research Center. 2015. The Smartphone Difference URL: <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/> [accessed 2017-02-28] [WebCite Cache ID 6ejS9bn6M]
5. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Transl Behav Med* 2014 Dec;4(4):363-371 [FREE Full text] [doi: [10.1007/s13142-014-0293-9](https://doi.org/10.1007/s13142-014-0293-9)] [Medline: [25584085](https://pubmed.ncbi.nlm.nih.gov/25584085/)]
6. Research2guidance. Research2guidance URL: <http://research2guidance.com> [accessed 2016-01-22] [WebCite Cache ID 6ejS4C09X]
7. Fox S, Duggan M. Pew Research Center. Mobile Health 2012: Half of smartphone owners use their devices to get health information one-fifth of smartphone owners have health apps URL: <http://www.pewinternet.org/2012/11/08/mobile-health-2012/> [accessed 2017-02-27] [WebCite Cache ID 6ob6C9mfG]
8. 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](https://doi.org/10.2196/jmir.1933)] [Medline: [23615206](https://pubmed.ncbi.nlm.nih.gov/23615206/)]
9. Zhu F, Bosch M, Woo I, Kim S, Boushey CJ, Ebert DS, et al. The use of mobile devices in aiding dietary assessment and evaluation. *IEEE J Sel Top Signal Process* 2010 Aug;4(4):756-766 [FREE Full text] [doi: [10.1109/JSTSP.2010.2051471](https://doi.org/10.1109/JSTSP.2010.2051471)] [Medline: [20862266](https://pubmed.ncbi.nlm.nih.gov/20862266/)]
10. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
11. O'Malley G, Dowdall G, Burls A, Perry IJ, Curran N. Exploring the usability of a mobile app for adolescent obesity management. *JMIR Mhealth Uhealth* 2014;2(2):e29 [FREE Full text] [doi: [10.2196/mhealth.3262](https://doi.org/10.2196/mhealth.3262)] [Medline: [25098237](https://pubmed.ncbi.nlm.nih.gov/25098237/)]
12. O'Malley G, Clarke M, Burls A, Murphy S, Murphy N, Perry IJ. A smartphone intervention for adolescent obesity: study protocol for a randomised controlled non-inferiority trial. *Trials* 2014;15:43 [FREE Full text] [doi: [10.1186/1745-6215-15-43](https://doi.org/10.1186/1745-6215-15-43)] [Medline: [24485327](https://pubmed.ncbi.nlm.nih.gov/24485327/)]
13. Knight E, Stuckey MI, Prapavessis H, Petrella RJ. Public health guidelines for physical activity: is there an app for that? A review of android and apple app stores. *JMIR Mhealth Uhealth* 2015;3(2):e43 [FREE Full text] [doi: [10.2196/mhealth.4003](https://doi.org/10.2196/mhealth.4003)] [Medline: [25998158](https://pubmed.ncbi.nlm.nih.gov/25998158/)]
14. Pagoto S, Schneider K, Jovic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. *Am J Prev Med* 2013 Nov;45(5):576-582. [doi: [10.1016/j.amepre.2013.04.025](https://doi.org/10.1016/j.amepre.2013.04.025)] [Medline: [24139770](https://pubmed.ncbi.nlm.nih.gov/24139770/)]
15. Breton ER, Fuemmeler BF, Abrams LC. Weight loss—there is an app for that! But does it adhere to evidence-informed practices? *Transl Behav Med* 2011 Dec;1(4):523-529 [FREE Full text] [doi: [10.1007/s13142-011-0076-5](https://doi.org/10.1007/s13142-011-0076-5)] [Medline: [24073074](https://pubmed.ncbi.nlm.nih.gov/24073074/)]
16. Abrams LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. *Am J Prev Med* 2011 Mar;40(3):279-285 [FREE Full text] [doi: [10.1016/j.amepre.2010.10.032](https://doi.org/10.1016/j.amepre.2010.10.032)] [Medline: [21335258](https://pubmed.ncbi.nlm.nih.gov/21335258/)]
17. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. *Pediatr Diabetes* 2013 Jun;14(4):231-238 [FREE Full text] [doi: [10.1111/pedi.12034](https://doi.org/10.1111/pedi.12034)] [Medline: [23627878](https://pubmed.ncbi.nlm.nih.gov/23627878/)]
18. 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](https://doi.org/10.2196/jmir.3117)] [Medline: [25048379](https://pubmed.ncbi.nlm.nih.gov/25048379/)]
19. McCully SN, Don BP, Updegraff JA. Using the Internet to help with diet, weight, and physical activity: results from the Health Information National Trends Survey (HINTS). *J Med Internet Res* 2013;15(8):e148 [FREE Full text] [doi: [10.2196/jmir.2612](https://doi.org/10.2196/jmir.2612)] [Medline: [23906945](https://pubmed.ncbi.nlm.nih.gov/23906945/)]
20. HINTS. 2015. Health Information National Trends Survey URL: <http://hints.cancer.gov/> [accessed 2016-01-22] [WebCite Cache ID 6ejRyxj9Y]
21. Finney RL, Hesse BW, Moser RP, Ortiz MA, Kornfeld J, Vanderpool RC, et al. Socioeconomic and geographic disparities in health information seeking and Internet use in Puerto Rico. *J Med Internet Res* 2012 Jul;14(4):e104 [FREE Full text] [doi: [10.2196/jmir.2007](https://doi.org/10.2196/jmir.2007)] [Medline: [22849971](https://pubmed.ncbi.nlm.nih.gov/22849971/)]

22. Kontos EZ, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. *J Med Internet Res* 2007;9(4):e31 [FREE Full text] [doi: [10.2196/jmir.9.4.e31](https://doi.org/10.2196/jmir.9.4.e31)] [Medline: [17951215](https://pubmed.ncbi.nlm.nih.gov/17951215/)]

Abbreviations

HINTS: Health Information National Trends Survey

Edited by G Eysenbach; submitted 04.02.16; peer-reviewed by J Updegraff, A Burls, B Fuemmeler; comments to author 06.04.16; revised version received 18.05.16; accepted 21.06.16; published 19.04.17.

Please cite as:

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;19(4):e125

URL: <http://www.jmir.org/2017/4/e125/>

doi: [10.2196/jmir.5604](https://doi.org/10.2196/jmir.5604)

PMID: [28428170](https://pubmed.ncbi.nlm.nih.gov/28428170/)

©Jennifer K Carroll, Anne Moorhead, Raymond Bond, William G LeBlanc, Robert J Petrella, Kevin Fiscella. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 19.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis

Cheng K Fred Wen^{1*}, MPH; Stefan Schneider^{2,3*}, PhD; Arthur A Stone^{2,3,4*}, PhD; Donna Spruijt-Metz^{1,3,4,5*}, MFA, PhD

¹Department of Preventive Medicine, University of Southern California, Los Angeles, CA, United States

²Center for Self-Report Science, University of Southern California, Los Angeles, CA, United States

³Center for Economic and Social Research, University of Southern California, Los Angeles, CA, United States

⁴Department of Psychology, University of Southern California, Los Angeles, CA, United States

⁵mHealth Collaboratory, University of Southern California, Los Angeles, CA, United States

*all authors contributed equally

Corresponding Author:

Cheng K Fred Wen, MPH

Department of Preventive Medicine

University of Southern California

Soto Bldg 1, 3rd Fl

2001 N Soto St

Los Angeles, CA, 90033

United States

Phone: 1 5626822468

Email: chengkuw@usc.edu

Abstract

Background: Mobile device-based ecological momentary assessment (mobile-EMA) is increasingly used to collect participants' data in real-time and in context. Although EMA offers methodological advantages, these advantages can be diminished by participant noncompliance. However, evidence on how well participants comply with mobile-EMA protocols and how study design factors associated with participant compliance is limited, especially in the youth literature.

Objective: To systematically and meta-analytically examine youth's compliance to mobile-EMA protocols and moderators of participant compliance in clinical and nonclinical settings.

Methods: Studies using mobile devices to collect EMA data among youth (age ≤ 18 years old) were identified. A systematic review was conducted to describe the characteristics of mobile-EMA protocols and author-reported factors associated with compliance. Random effects meta-analyses were conducted to estimate the overall compliance across studies and to explore factors associated with differences in youths' compliance.

Results: This review included 42 unique studies that assessed behaviors, subjective experiences, and contextual information. Mobile phones were used as the primary mode of EMA data collection in 48% (20/42) of the reviewed studies. In total, 12% (5/42) of the studies used wearable devices in addition to the EMA data collection platforms. About half of the studies (62%, 24/42) recruited youth from nonclinical settings. Most (98%, 41/42) studies used a time-based sampling protocol. Among these studies, most (95%, 39/41) prompted youth 2-9 times daily, for a study length ranging from 2-42 days. Sampling frequency and study length did not differ between studies with participants from clinical versus nonclinical settings. Most (88%, 36/41) studies with a time-based sampling protocol defined compliance as the proportion of prompts to which participants responded. In these studies, the weighted average compliance rate was 78.3%. The average compliance rates were not different between studies with clinical (76.9%) and nonclinical (79.2%; $P=.29$) and studies that used only a mobile-EMA platform (77.4%) and mobile platform plus additional wearable devices (73.0%, $P=.36$). Among clinical studies, the mean compliance rate was significantly lower in studies that prompted participants 2-3 times (73.5%) or 4-5 times (66.9%) compared with studies with a higher sampling frequency (6+ times: 89.3%). Among nonclinical studies, a higher average compliance rate was observed in studies that prompted participants 2-3 times daily (91.7%) compared with those that prompted participants more frequently (4-5 times: 77.4%; 6+ times: 75.0%). The reported compliance rates did not differ by duration of EMA period among studies from either clinical or nonclinical settings.

Conclusions: The compliance rate among mobile-EMA studies in youth is moderate but suboptimal. Study design may affect protocol compliance differently between clinical and nonclinical participants; including additional wearable devices did not affect participant compliance. A more consistent compliance-related result reporting practices can facilitate understanding and improvement of participant compliance with EMA data collection among youth.

(*J Med Internet Res* 2017;19(4):e132) doi:[10.2196/jmir.6641](https://doi.org/10.2196/jmir.6641)

KEYWORDS

ecological momentary assessment; compliance; youth; mHealth

Introduction

Background

There is a growing interest in studying the dynamic relationship among individuals' experiences, social or physical environments, and behaviors. The assessment of these dynamic relationships is enhanced by the development of momentary data collection strategies, such as experience sampling methods (ESM) and ecological momentary assessment (EMA) [1]. Studies using these strategies usually require that their participants provide self-report ratings of their current or near-current experiences, environments, and behaviors. As summarized by Shiffman et al [1], these repeated "in the moment" measurements offer numerous methodological advantages over traditional assessment strategies. First, the momentary assessment of participants' current or immediate past experiences or behaviors reduces the recall bias inherent in traditional retrospective survey methods. Second, the "in the moment" and "in the context" measurements collected in natural settings provide data that is more relevant to the current social or physical environments, thereby providing more ecologically valid data. Third, the daily intensive repeated measurements allow for examinations of immediate antecedents and consequences of behavior in real-time, capturing within-day, within-person behavior, and experience variations across time.

EMA studies can be broadly categorized into (1) time-based and (2) event-based designs. These strategies provide different insights about the study participants. The time-based strategy usually aims to acquire representative characteristics and patterns of behaviors and experiences across time, whereas a study using an event-based strategy aims to examine antecedents and consequences of specific behaviors or experiences [1]. On the basis of the study rationales, variations of time-based (eg, prompting participants at random times and within a window of time) and event-based (eg, participant self-initiated self-report in response to occurrence of specific events or prompted by sensed events, such as location via Global Positioning System (GPS), or bouts of physical activity via accelerometer) strategies have been used, either on their own or in combinations. Technology innovations have transformed and enhanced momentary data collections in natural settings during the past decade. Features specific to mobile technology have provided solutions for many challenges that early EMA researchers faced. For example, some noncompliant behaviors, such as backfilling, or completing all the assigned diaries in bulk at the same time [2], can be effectively addressed by disabling prompt access on a mobile-EMA platform after a specific time window. Furthermore, mobile technologies provide researchers with

time-stamped data on participant compliance that allow for a more reliable and objective measurement of participant compliance as compared with traditional paper-and-pencil recall methods, which have been shown to produce an inflated compliance rate [3]. These noncompliant behaviors (eg, backfilling), as well as other sources of bias (eg, missing assessments due to engagement in other activities that compete for participants' attention), can yield biased data that no longer corresponds to the moment when the behaviors or experiences of interest occurred, thereby reducing ecological validity of the collected data. A major strength of mobile technologies is the readily available features that can validate the timeliness of participants' response (eg, built-in sensors, phone usage data, and automatic timestamp) that can objectively measure response compliance. These can enhance the validity of momentary data collected in mobile technology-based EMA studies. Various emerging mobile technologies have been incorporated into EMA studies throughout the past decade; for example, personal digital assistants (PDAs), palmtop computers, and more recently, smartphones. Among participants across all age groups, mobile device-based EMA studies offer promising opportunities for researchers to study behavior and experience, especially in the youth population. Youth are "digital natives" [4] and are considered adept in and comfortable with technology in their day-to-day life activities, for example, for communication [5,6], and for receiving health-related intervention materials [5]. The high acceptability and ubiquity of digital and mobile devices, along with the methodological advantages, presents valuable opportunities for researchers to engage young study participants in EMA studies. During the recent decade, mobile device-based EMA has been widely utilized to assess and understand the dynamic relationship among youth's behaviors, experiences, and pertinent contextual information in the youth population.

Although collecting momentary data using mobile technologies offers many advantages, these advantages depend on the quality of the collected data. Incorporation of mobile technologies in EMA studies can facilitate momentary data collection with an improved measurement of compliance and possibly in a higher frequency than using more conventional collection techniques. Although this provides an opportunity to understand behavior on a more granular level, systemic missing data (eg, participant noncompliance or engagement in competing activities) still threatens data quality. As stated above, several features of mobile technologies can minimize the impacts of some types of noncompliance behavior (eg, backfilling) on data quality. Nonetheless, as EMA study protocols usually involve participants being repeatedly interrupted and asked to provide self-reported information, these demands on study participants can lead to high perceived participant burden, and to

noncompliance [2]. In the context of mobile-EMA studies, possible sources of participant burden include, but are not limited to, the use of technology (eg, familiarity with the reporting platform and incorporation of additional wearable devices), technological issues (eg, problems with the reporting platforms), study design (length of monitoring and daily sampling frequency), and quality, complexity, and the duration of prompts. Nonetheless, there are a limited number of studies that systemically review participant compliance to EMA protocols [1,3,7-9] and only one review specifically focuses on youth populations [9]. In the review, Liao et al [9] included 13 studies and observed an average compliance rate of 71%. This review, although an important contribution to the literature, was limited to obesity-related behaviors and did not quantitatively examine relationship between participant compliance and study design factors. This systematic review and meta-analysis, therefore, expands upon Liao et al findings by examining compliance to EMA protocols in youth involving a variety of self-reported behaviors and experiences, and by quantitatively assessing the relationships between compliance and some aspects of study design. We restrict the inclusion to studies that use digital momentary assessment techniques where compliance is electronically time stamped on the momentary level.

Aims of This Study

The first aim of this review is to describe the characteristics of EMA protocols conducted among pediatric populations across a wide spectrum of health behaviors. The second aim is to quantify overall compliance rates and to examine the association between study design factors (length of monitoring period and daily sampling frequency) and reported compliance using a meta-analytic approach. Studies using clinical and nonclinical samples were both included in the review; however, given that study populations and objectives often differ quite substantially for these types of studies, they were examined separately and the results were compared. The exploratory aim of this study is to examine the association between participant compliance and other pertinent study design variables (eg, inclusion of additional wearable devices and incentive structure) on a post hoc basis. Finally, this study will also provide recommendations for future research that incorporates mobile devices in collecting real-time self-reported data to maximize the advantages of EMA methodologies.

Methods

Data Acquisition

A comprehensive literature search was conducted using the publicly accessible academic literature search engines (PubMed, PsycINFO, *Journal of Medical Internet Research*, and Google Scholar) from inception to March 28, 2016. The search terms employed in this review were composed of two components: (1) terms related to EMA and (2) terms related to participants aged 18 years or younger. Terms related to EMA were “ecological momentary assessment,” “ecological momentary intervention,” “momentary,” “experience sampling methods,” “event sampling methods,” and “daily diary methods.” Terms related to participant age were “adolescent,” “child,” “children,”

and “youth.” Additional empirical studies were identified from the citations of the articles.

Inclusion and Exclusion Criteria

Criteria for Inclusion in the Qualitative Systemic Review

Abstracts and full articles of the retrieved titles were screened for relevance. The article selection strategies, inclusion criteria, and exclusion criteria were determined in consensus meetings among authors. In order to be included in the systematic review, studies were required to (1) be an empirical study; (2) employ EMA strategies, including diary methods with more than one entry per day, ESM, and event sampling methods; (3) utilize mobile technologies for EMA data collection (cell phones, PDAs, smartphones, and so on); and (4) include children or adolescent (age ≤ 18 years) participants. Studies that involved adult participants (age > 18 years), in addition to children and adolescents, were only included in the review if separate analytic or descriptive results were presented for the children or adolescents subgroups. Studies with any of the following 5 exclusion criteria were excluded: (1) did not utilize any electronic, wearable, or mobile technology; (2) utilized paper-based diaries to collect momentary data; (3) collected momentary or diary data once or less than once per day during the monitoring period; (4) utilized call-based (phone interview) data collection; and (5) data collection did not take place in free-living natural settings.

Criteria for Inclusion in the Meta-Analysis

A subset of studies meeting the criteria for the systematic review was included in the meta-analysis portion of the study. To meet the criteria for meta-analysis, studies were required to report (1) sufficient information that permitted the calculation of an average compliance rate (eg, percentage of EMA prompts answered) to be used as effect size (ES), (2) number of participants in the study, and (3) daily prompting frequency and length of monitoring period.

Meta-Analysis Procedures

Random effects meta-analyses were conducted to (1) examine the average rate of compliance with EMA protocols pooled across all included studies and then across studies with clinical participants and nonclinical participants separately, and (2) to explore potential between-study differences in compliance rates based on daily prompting frequency and length of monitoring. A post hoc analysis was conducted to examine whether there is a difference in compliance rates among studies (1) with and without wearable devices *in addition to* the mobile-EMA platform and (2) with a fixed and incremental incentive strategy. In meta-analyses, study level averages are synthesized rather than individual participant data. Accordingly, compliance rates were operationalized as the average proportion of prompts completed by a participant in a given study (ie, the actual number of prompts completed divided by the number of prompts specified by the study protocol). To calculate an ES adequate for the analysis of proportions [10], the compliance rates were logit-transformed and standard errors were calculated accordingly as shown in Figure 1, where p is the proportion of completed prompts and n is the effective sample size in the study.

In this review, an adequate calculation of the standard errors (and hence the inverse variance weights used in meta-analysis) is complicated by the fact that EMA studies involve a nested study design, with multiple observations (prompts) clustered in participants. In this case, the effective sample size of each study needs to account for the clustered design. Following the methods recommended in the Cochrane Handbook for Systemic Review of Interventions [11], this can be achieved by adjusting the total sample size (ie, total number of prompts in a study) by the intraclass correlation coefficient (ICC) representing the variation of compliance between- and within-study participants. Eight studies reported the variance in participant compliance, which prevented this review from estimating ICCs separately for each included study. In the 8 studies that reported this information, the range of participant-level standard deviations in compliance was 5.96-29.98%. Accordingly, the sample sizes used to calculate the standard errors were adjusted by ICCs reflecting

this range: the meta-analyses were conducted using an ICC calculated assuming lower (SD 5.96%), intermediate (SD 15.00%), and higher (SD 29.98%) values of the ICC, and results were compared in sensitivity analyses. The sensitivity analyses showed that the results were not affected by different ICC values, and only results estimated using the intermediate ICC value are presented here.

The Q statistic was computed by summing the squared deviations of each study's ES from the average pooled ES and was used to test the statistical significance of between-study heterogeneity in compliance rates. The I^2 statistic, which estimates what proportion of the between-study variance is due to actual differences rather than chance, was used to quantify the magnitude of between-study heterogeneity, with values of 25%, 50%, and 75% representing low, medium, and high heterogeneity, respectively [12].

Figure 1. Calculation of effect size.

$$ES_p = \log_e \left[\frac{p}{(1-p)} \right] \text{ and } SE_p = \sqrt{\frac{1}{np} + \frac{1}{n(1-p)}}$$

Moderator Analyses

The second goal of this meta-analysis was to examine the association between EMA study design characters (average daily prompting frequency and length of monitoring period) and participant compliance rates. Both average daily prompting frequency and length of monitoring were coded based on information described in the reviewed publications. For studies that employed different frequencies for weekends and weekdays, an average daily prompting frequency was calculated by dividing the total number of times participants were prompted by the number of study days.

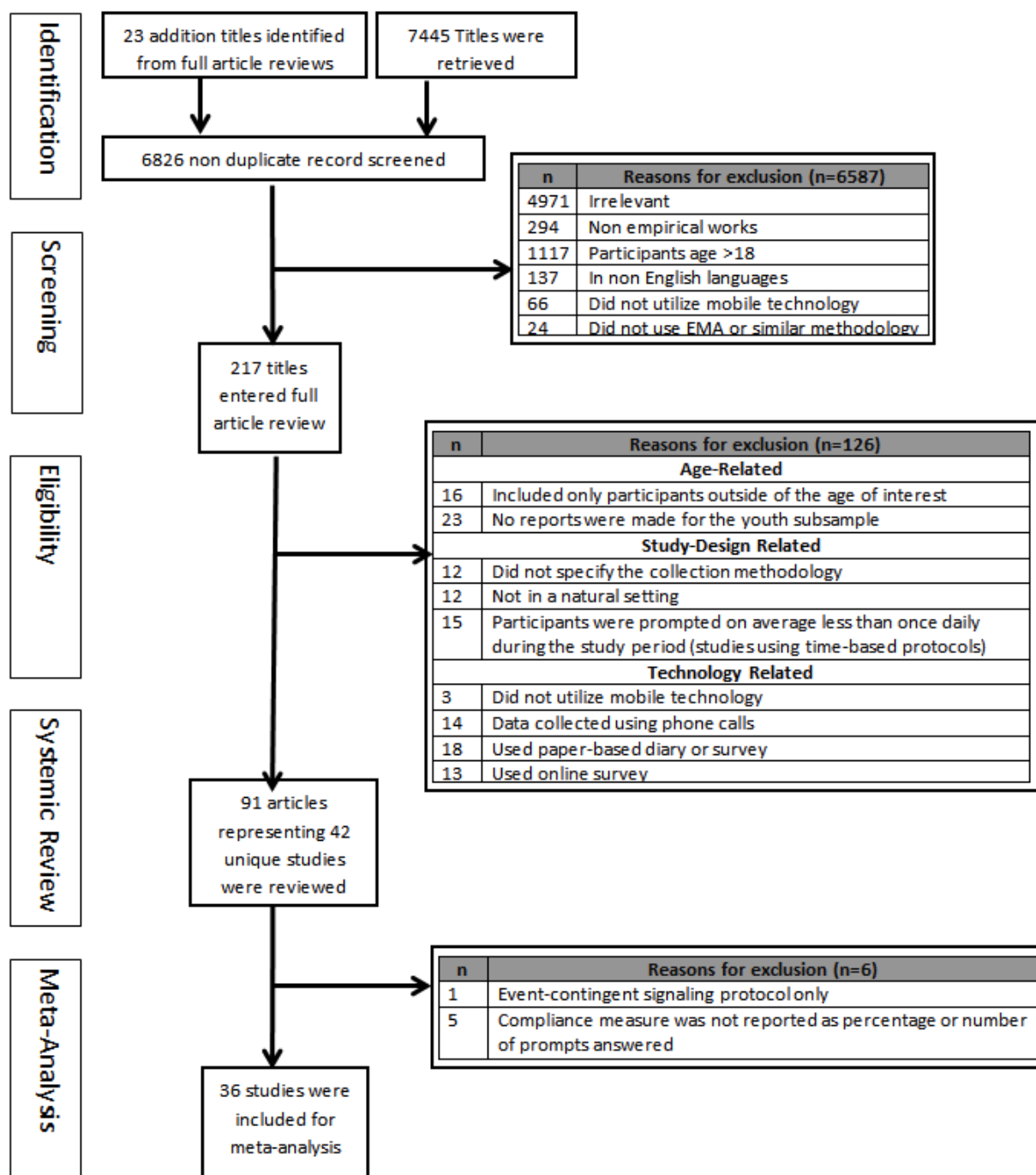
The associations between study design variables (ie, length of EMA protocols and sampling frequency) and reported compliance were examined using random effects analysis of variance (ANOVAs) with inverse variance weights. Models that examine the association of compliance with (1) study length and (2) daily prompting frequency were estimated separately for studies with participants from nonclinical or clinical populations. The length of study protocol was operationalized as study length ≤ 1 week, > 1 week and ≤ 2 weeks, and > 2 weeks, and the prompting frequency was operationalized as prompting

frequency of 2-3 times per day, 4-5 times per day, and ≥ 6 times per day to ensure that each category included a sufficient amount of studies for the purposes of comparison. We considered testing the interaction term of study length and prompting frequency, but did not conduct this analysis because there would have been no or very few studies in several categories comprising the interaction effect. All meta-analysis procedures were conducted using Comprehensive Meta-Analysis (version 3, Englewood NJ, USA).

Results

Study Selection

A total of 6826 nonduplicate titles were identified. Of these, 6803 were identified using search engines and 23 were identified through cited work from the articles screened. After reviewing abstracts and full-text articles for inclusion and exclusion criteria, 91 empirical articles representing 42 unique studies were included in the qualitative systematic review and 36 studies were included in the meta-analysis portion of the study. The detailed study selection process is outlined in Figure 2, the preferred reporting items for systemic review and meta-analysis (PRISMA) diagram [13].

Figure 2. The preferred reporting items for systemic review and meta-analysis (PRISMA) diagram.

Participant Characteristics

Among studies included in the systemic review, the average number of participants in the analytic sample across studies was 98.81 (SD 130.66; range 5-562). Across all studies, the average proportion of female participants was 56.4%, where 3 (7.1%) recruited only female participants. Excluding these 3 studies, the average proportion of female participants was 52.7% (SD 18.7%; range 7.6-86.7%). A majority of the included studies (n=26, 61.9%) recruited only participants from community or nonclinical settings. The 16 studies with clinical populations focused on youths with various health conditions: attention

deficient/hyperactivity disorder [14-16] (25%, 4/16), juvenile idiopathic arthritis (JIA) [17-19] (25%, 4/16), asthma (13%, 2/16), type 1 diabetes (T1D) [20,21] (13%, 2/16), high function autism and Asperger's syndrome (HFASD) [22] (6%, 1/16), concussion patients [23] (6%, 1/16), neurology clinic patients [24] (6%, 1/16), and recovery patients [25] (6%, 1/16). A detailed outline of the study participants can be found in [Multimedia Appendix 1](#).

Study Characteristics

Study Length

The length of EMA protocols ranged from 2 to 42 days (13.27 [SD 9.08]). The average length of monitoring was not statistically different between studies with nonclinical participants (mean 11.4 days [SD 6.9; range 4-30 days]) and clinical participants (mean 16.3 days [SD 11.2; range 2-42 days]; $t_{24,5} = -1.598$, $P = .12$).

Sampling Strategy

A total of 33 (78.6%) studies utilized only time-based sampling protocols, 1 study (2.4%) used only an event-based sampling protocol, and 8 studies (19.0%) used a combination of both time- and event-based sampling protocols.

Among the 41 studies with a time-based sampling component, prompting schedules included random prompts during predetermined time intervals ($n=31$, 75.6%; eg, one random prompt for each 2 h interval), prompts at a fixed schedule ($n=8$, 19.5%; eg, every 30 min during waking hours), prompts at a personalized time ($n=1$, 2.4%; eg, participant's own blood glucose check schedule), and one study did not report the prompting scheme. Prompting schemes of which participants were prompted is shown in [Multimedia Appendix 2](#). Studies assessed a wide range of behaviors (eg, current activity, meal consumption, alcohol consumption, self-injury behavior, unprotected sexual behavior, and disease management), subjective experience (pain, mood, stress, appetite, attention, functional limitations, level of enjoyment, and level of control), and contextual information (current physical location, presence of social companion, presence of smoking cue, and food availability).

Among the 9 studies with an event-based protocol, participants in 8 studies were asked to initiate self-report after occurrence of certain thoughts or emotions such as positive feelings [26], negative feelings [26,27], self-injury thoughts [28]), physical symptoms [27], behaviors (ie, drinking [29], smoking [30,31], eating occasion [32], or self-injury behavior [28]), or exposure to smoking or alcohol-related media campaigns [33]. One study [34] automatically sent an EMA prompt approximately 5 min after using a Bluetooth-enabled inhaler. [Multimedia Appendix 3](#) lists the experiences and/or behaviors that the included studies asked their participants to self-report for the event-based protocol.

Sampling Frequency

A majority of the included studies (95%, 39/41) prompted their participants 2-9 times during each day of EMA data collection. It was found that 2 (4.9%) studies prompted participants more than 25 times each day and 10 (24.3%) studies reported prompting participants in different frequencies on weekdays versus weekend days. Excluding the two studies that prompted participants more than 25 times each day, the average prompting frequency was 4.2 times per day (range 2-9) for studies with nonclinical participants and 3.6 times per day (range 2-7) for studies with clinical samples.

Sampling Devices

EMA data was collected using electronic diaries ($n=1$, 2.3%), wearable platforms ($n=1$, 2.3%), iPods ($n=2$, 4.6%), PDAs ($n=10$, 23.8%), palmtop computers ($n=12$, 28.4%), and mobile phones ($n=16$, 48.1%). A small proportion of studies ($n=5$, 11.9%) reportedly used participants' own phone or mobile phones to implement EMA data collection. Four of these studies sent text messages to participants' own phones or mobile phones for EMA data collection and one allowed participants to choose between using a mobile phone provided by the study and their own smart devices. A small proportion of studies ($n=5$, 11.9%) used wearable devices in addition to the EMA data collection platform. Devices utilized in addition to the EMA data collection platform included accelerometers [34-37], heart rate monitors [35,36], Bluetooth-enabled inhalers [34], and glucometers [21].

Incentive for Participants

A majority of the reviewed studies ($n=28$, 66.67%) reported the strategy used for incentivizing study participants. Among these studies, most of them ($n=26$, 92.86%) provided monetary incentive to their participants. Two studies reported using other nonmonetary incentive strategies, for example, raffle [26] and level-up (promotion) in the EMA software platform [18]. Among studies that provided study participants with monetary incentive, study participants were compensated either (1) in a fixed amount ($n=16$, 57.14%; ranged from US \$40 to US \$200) or (2) an incremental amount of monetary incentive ($n=10$, 35.71%). In studies that used the latter approach, participants received a base amount of compensation (ranged from US \$20 to US \$50) for participation with additional incentive in various rate based on author-specified compliance thresholds. Nonetheless, there is no clear common rationale for determining the level of incentive observed among the reviewed studies. Detailed information about the incentive structure used in these studies can be found in [Multimedia Appendix 1](#).

Compliance as Reported in the Original Studies

Definition of Reported Compliance

Among studies with a time-based sampling protocol component ($N=41$), the majority ($n=36$, 87.8%) defined participant compliance as percentage of prompts to which participants responded. Two studies included response latency, or the time difference between a prompt and participant's response to that particular prompt, as part of the definition of compliance (eg, percentage of prompts responded within 30 min of the first notification [27]). One study provided the percentage of participants who reached a predetermined compliance cutoff (ie, percentage of participants completed 28 entries [28]). The definition of compliance among time-based sampling protocols is listed by study in [Multimedia Appendix 2](#). Among studies that reported compliance in the format of proportion of prompts completed ($n=36$), the reported compliance rates ranged from 54.6% to 96.21%. Approximately 31% of the reviewed studies ($n=13$) reportedly excluded participants from the analytic sample because they were considered dropouts or did not meet a minimal compliance threshold.

Among studies with an event-based sampling protocol component ($n=9$), the majority ($n=8$, 88.9%) asked participants

to initiate self-report. These events to initiate self-reports included occurrence of behaviors [27-32,34], media exposure [33], or subjective experiences [26-28] during the monitoring period (Multimedia Appendix 3). Limited information about compliance was available from the 8 studies with event-based protocols. Among these studies, 6 reported compliance in the format of count of prompts that contained information about the behavior of interest. One (11.1%) study with an event-based protocol asked participants to respond to EMA prompts triggered by events (ie, use of the inhaler) sensed by a Bluetooth-enabled inhaler [34]. In this study, compliance was reported as the proportion of answered prompts triggered by the sensed use of the Bluetooth-enabled inhaler. The compliance rate reported in this study was 47.90%.

Reported Correlates of Participant Compliance

Among studies with clinical participants ($n=17$), 8 examined correlates of compliance and reported no significant association between prompt completion rate and day of the week [18,20], time of day [18,20], age [19,20,22,24,34], gender [18-20,22,34,38], disease status [19,20,22,24,34,38], and technical difficulties [17]. One study reported that prompt completion rate was positively associated with participants' intelligence quotient (IQ) [22]. Three studies reported declines in completion rates over time reported [17,19,20] and one did not observe such a difference [18]. One study documented reasons for missing assessments included technical issues ("did not hear the notification") [20] and the timing of the prompt [20].

Reported significant correlates of compliance among studies with nonclinical participants included gender [39-41], ethnicity [36,40], health condition [42], and baseline affect [41,43]. Weekday status [36,44] and participant age [36,40] were reportedly not associated with prompt completion rate. Mixed findings on the correlation of prompting time of the day and completion rate were reported in 4 studies [36,39,44,45]. Declines in participants' completion rate over the course of the study were reportedly tested in 4 studies [30,31,39,46] and 3 reported a significant completion rate decline [30,39,46]. It was found that 5 studies documented reasons for missed assessments. These reasons included participants' engagement in competing activities [42,47], device malfunction [36,46,47], not hearing the notification [42], and participant forgetfulness [30,46].

Average Rates and Moderators of Participant Compliance: Meta-Analysis Results

Average Compliance Rate

A total of 36 studies with a time-based EMA protocol were included in the meta-analysis portion of the study. After accounting for the cluster effect of momentary assessments within participants, the average compliance rate across the included studies was 78.26% (95% CI 75.49-80.78%), and the average compliance rate was not associated with the average age or gender proportion of the study participants. The average compliance rates were not statistically different between (1) studies with EMA data collected using one mobile platform (77.44%, 95% CI [73.59-80.88%]) compared with studies using a mobile platform with additional wearable devices ($n=5$, 73.00%, 95% CI [61.75-81.91%]; $z=-0.91$, $P=.36$); (2) studies using a fixed incentive strategy ($n=15$, 79.08%, 95% CI [69.08-86.48%]), an incremental incentive strategy ($n=10$, 72.95%, 95% CI [62.36-81.44%]), and did not report using an incentive strategy ($n=10$, 80.15%, 95% CI [73.00-85.77%]); and (3) studies with participants from clinical settings (76.92%, 95% CI [70.76-82.11%]) compared with studies with participants from nonclinical settings (79.15%, 95% CI [75.59-82.32%]; $z=-1.06$, $P=.29$). There was substantial between-study variation in compliance rates for both studies with clinical ($I^2=48.33\%$, $Q_{\text{total}}=27.09$, $df=14$, $P=.02$) and nonclinical participants ($I^2=66.93$, $Q_{\text{total}}=60.48$, $df=20$, $P<.001$). Thus, the examination of moderators of ESs was warranted.

Daily Prompting Frequency as Moderator of Compliance

Daily prompting frequency significantly moderated the compliance rates among clinical ($Q_{\text{between}}=9.78$, $df=2$, $P=.008$; $R^2=.74$) and nonclinical ($Q_{\text{between}}=15.13$, $df=2$, $P<.001$; $R^2=.44$) studies. Among studies with clinical participants, the compliance rates were significantly higher in studies that employed the *most* frequent prompts (6 or more times a day) compared with studies with less frequent sampling of 2-3 times per day ($z=-2.68$, $P=.007$) and 4-5 times per day ($z=-3.10$, $P=.002$). Conversely, among studies with only nonclinical participants, the compliance rates were significantly higher in studies that prompted participants *the least* frequently (2-3 times a day), compared with studies with prompting frequencies of 4-5 times ($z=-3.81$, $P<.001$), or 6 or more times per day ($z=-3.53$, $P<.001$; Table 1).

Table 1. Prompting frequency by intensity category.

Setting	Prompting frequency (# of prompts per day)	n	Compliance (95% CI)	R^2_{analog}	Q_{residual}^a	$I^2(\%)^b$
Clinical	2-3 times	11	73.47 (67.45-78.73) _d	0.74	14.97 (df=12)	19.82
	4-5 times	4	66.94 (53.50-78.09) _e			
	6+ times	2	89.28 (78.83-94.90) _c			
Nonclinical	2-3 times	6	91.73 (85.48-95.44) _g	0.44	38.27 (df=18) $P=.004$	52.96
	4-5 times	13	74.42 (59.37-85.29) _e			
	6+ times	5	75.00 (59.21-86.12) _f			

^a Q_{residual} : test for residual between-study variance (not explained by the moderator) against zero.

^c I^2 : percentage of the residual variability that is due to heterogeneity rather than sampling error.

^c $P=.007$, compared to study with a prompting frequency of 2-3 times.

^d $P=.007$, compared to study with a prompting frequency of 6+ times.

^d $P=.002$, compared to study with a prompting frequency of 6+ times.

^e $P<.001$, compared to study with a prompting frequency of 2-3 times.

^f $P<.001$, compared to study with a prompting frequency of 2-3 times.

^d $P<.001$, compared to study with a prompting frequency of 6+ times.

Table 2. Length of monitoring by week.

Settings	Length of EMA ^a monitoring (number of weeks)	n	Compliance (95% CI)	R^2_{analog}	Q_{residual}^b	$I^2(\%)^c$
Clinical	1	6	78.13 (64.37-87.61)	0	26.67 (df=12) $P=.009$	55.01
	2	5	73.46 (53.74-86.84)			
	3+	6	75.47 (56.86-87.78)			
Nonclinical	1	14	75.81 (70.39-80.52)	0.11	51.43 (df=18) $P<.001$	65
	2	5	76.77 (61.30-87.33)			
	3	5	83.95 (74.69-90.71)			

^aEMA: ecological momentary assessment.

^b Q_{residual} : test for residual between-study variance (not explained by the moderator) against zero.

^c I^2 : percentage of the residual variability that is due to heterogeneity rather than sampling error.

Study Length as Moderator of Compliance

There were no significant differences in reported compliance between studies that engaged participants in an EMA protocol for 2 and 3 or more weeks compared with studies that engaged participants for 1 week or less, among both studies with participants from clinical ($Q_{\text{between}}=0.33$, $df=2$, $P=.85$) and nonclinical ($Q_{\text{between}}=2.60$, $df=18$, $P=.27$) settings (Table 2).

Discussion

Principal Findings

The aim of this study was to provide an up-to-date review of evidence on youths' compliance to real-time EMA protocols operated on mobile platforms. Interest in using EMA with mobile technology in youth is growing rapidly, as documented by the sizable number of mobile-EMA studies conducted to capture various aspects of youth's life. In the reviewed studies, we estimated an average compliance rate of 78.3% across studies

using time-based prompting protocols. Although this rate is comparable with the rate of 71% (range 44-96%) observed by Liao et al [9], this study's estimate is lower than the EMA compliance rate reported in the adult populations [48] and just falls short of the 80% rate recommended by Stone and Shiffman [49]. Considering that close to 30% ($n=11$) of the reviewed studies reported a compliance rate that is lower than 70%, there is a need to identify factors that may impact youths' compliance to mobile-EMA protocols in order to facilitate more optimal compliance rates.

Study Design and Completion Rates in Time-Based Protocols

The study designs varied considerably both in terms of the overall length of EMA monitoring and in terms of the frequency with which youths were prompted to complete momentary assessments per day. This allowed us to examine whether these specific EMA study design factors moderate compliance rates. Our meta-analytic findings provided evidence that the compliance rates are significantly different among studies of

different daily frequency of assessments. Importantly, although being significant for both nonclinical and clinical samples, the effect was in opposite directions for clinic and nonclinic participants. Among the 17 studies with participants from *clinical* settings, the two studies with the *highest prompting frequency* had a significantly higher compliance rate of 89.3% compared with studies with a less intense prompting frequency (73.5% for studies with the lowest frequency; 66.9% for studies prompted participant 4-6 times a day). Conversely, among studies with *nonclinical* participants, we estimated an average compliance rate of 91.73% among studies with the *lowest prompting frequency* and was significantly higher than the rate in studies that prompted participants for 4-5 times or 6+ times per day (74.4% and 75.0%, respectively). These results suggest that the association between prompting frequency and compliance differ between participants from nonclinical and clinical settings.

We can only speculate on the potential reasons for this result. One possibility is that studies in nonclinical and clinical settings differ in the content of the questions and how meaningful they are to respondents. Clinical studies commonly tap into medical and disease-related aspects of daily life that may be intrinsically relevant to the young patients. On the other hand, the content of EMA prompts in nonclinical studies may appear less intrinsically relevant to respondents, which may decrease compliance when the assessments are more frequent.

On the other hand, the meta-analytic results indicate that the overall compliance rates were similarly moderate among studies with different lengths (number of weeks) of EMA monitoring in either setting. However, as several reviewed studies with clinical [17,19,20] and nonclinical [30,39,46] participants reported declines in compliance rates over the course of study period, these results suggest that young participants' compliance to EMA protocols might deteriorate over time. These results emphasize the need for developing a more nuanced understanding of the possible factors and strategies that sustain youths' motivation to complete EMA prompts over extended periods of time. Several strategies (eg, reward systems, rotating item administration) were utilized as mechanisms for maintaining youths' motivation to comply with longer EMA protocols. The post hoc analysis results further indicate that the average compliance rates did not differ between studies with a fixed incentive structure and an incremental incentive structure. Although this result suggests that these two incentive strategies may have similar effectiveness in motivating young participants to comply, further investigations on crucial aspects of how these strategies affect participant compliance, for example, the mechanism of which young participants are motivated, is necessary. However, to the authors' knowledge, there is no published evidence that systemically examined the effectiveness of reward systems or other promising strategies on compliance to EMA among young participants.

Mobile Technologies Used in Current EMA Protocols

Several studies with a time-based sampling protocol in this review incorporated wearable or deployable devices such as accelerometers [34-37], heart rate monitor [36], GPS trackers [36], and glucometers [21] in their field data collection efforts,

in addition to the EMA reporting platforms. Since these devices often collect certain behavioral or contextual data passively with minimal inputs required from study participants, incorporating additional wearable devices in EMA data collection may be acceptable to youth without impacting participant compliance. Although providing a detailed review of feasibility and utility of existing mobile technology in youth is beyond the scope of this study, we identified several existing reviews that examined the use of wearable and mobile technology in assessing particular behaviors (eg, physical activity [5,50,51], dietary behavior [5,50,52], and smoking [53]). Considering the suboptimal average compliance rate estimated among mobile-EMA studies using a time-based protocol, combining wearable technologies with EMA data collection may offer a viable alternative to collect some self-reported data to alleviate participant burden. In addition to reducing participant burden, recent developments in human-computer interaction reveal the possibilities of using data captured by mobile phone sensors to identify ideal timing, or the "opportune moments" [54-57], to send prompts in order to minimize interruption or participant's engagement in competing activities. Several studies with adult participants [55,57,58] have developed algorithms that predict users' availability to respond and receptivity to be intervened. Although the extent to which data collected using this approach is subject to selection bias and its utility in youth are yet to be determined, utilizing the data passively collected from mobile devices may offer researchers valuable opportunities to understand participant behaviors and improve compliance.

Among the event-based protocols reviewed, only a small proportion of studies ($n=1$, 12.5%) used a protocol that emits event-based prompts based on objectively measured behavior of interest using wearable devices [34]. Currently, the majority of the mobile-EMA studies operationalize event-based sampling using a participant self-initiated self-report approach. Nonetheless, as self-reported information obtained using this unsolicited approach may subject to systemic under- or over-reporting [59], quantifying participant compliance to this type of protocol using the current common practice (reporting number of event of interest recorded) can be misleading and overly optimistic. Mobile technologies are increasingly sophisticated in capturing objective measurements of various behaviors [5,50-53]. Therefore, incorporating wearable mobile devices in event-based sampling procedure, in parallel with participant self-report, may present researchers the opportunity to capture objectively measured data about participants' behavior without impacting participant compliance.

Limitations in the Current Compliance Reporting Practice

Another finding from this review is that there are areas where compliance-related results and procedures were inadequately or inconsistently reported. First, among the time-based protocols, participant compliance was considered to be synonymous with average prompt completion rate (ie, mean percentage of prompt answered). The distribution of compliance rates is often negatively skewed (with the mass of the distribution concentrated at the higher end). If this is the case, the arithmetic mean provides a conservative representation of overall

compliance in the sample, and robust measures of central tendency (median or geometric mean) should be reported. In addition, this relatively vague definition of compliance does not account for important information, such as response latency, that could allow for assessment of response patterns or approximation of item cognitive load. Response latency, or the time difference between a prompt and its corresponding response, is especially relevant when assessing experiences that are time-varying and context-dependent (eg, pain, emotion, and hunger). For example, past emotions or experiences like pain are prone to be distorted by events or experiences occurred during the active reconstruction process of recall [60]. As the time-stamp information is automatically collected in modern mobile technologies, incorporating response latency in defining and reporting compliance can provide fine-grain insights to young users' compliant behavior. Therefore, we recommend that future EMA studies incorporate the time-frame in which the EMA must be completed in order to be considered compliant and report the latency of prompt completion when applicable.

Second, although a number of studies examined correlates of compliance (ie, quantitative assessment of compliance) and participant-reported reasons for noncompliance (qualitative assessment of compliance), results of both quantitative and qualitative assessment of compliance were inconsistently reported across studies, in part, because the data may not have been collected in the study. Obtaining and reporting information about how individual (eg, age and gender), technological (eg, software malfunction, device power depletion, and network connectivity), or time-varying (eg, time of day, environmental factors, and activities) factors relate to compliance and to missing data is important for at least two reasons. For one, identifying these factors is necessary for improvement of compliance in future EMA data collection. For example, by understanding which participant groups need to be specifically targeted (and when they need to be targeted), improved retention strategies can be formulated to address the challenges unique to participants of specific demographic groups and to facilitate overall compliance rates. In addition, without this information one cannot determine if missing data in EMA studies is merely random "noise" or if it is systematically linked to individual or situational characteristics. Systematic noncompliance is clearly a major threat to the validity of conclusions and analytic steps can be taken to attenuate the bias if the attributing factors are known. Therefore, we encourage future EMA studies to report both quantitative and qualitative compliance results.

Third, several studies reported compliance rates only among those in the final analytic sample, after removing participants with low compliance. Many studies provided rationales for excluding low- or noncompliant participants. Nonetheless, the compliance rate reported with these subsamples can be viewed as inflated and would be likely to affect our ability to accurately estimate the average compliance across studies. Therefore, we recommend that future studies report compliance rates before and after removing the participants from analyses to enhance transparency of the analysis process.

Limitations of This Review

A major strength of this study was that we were able to quantitatively assess the compliance rates for mobile-EMA studies of various health-related behaviors and the association between the reported prompt completion rate and some design factors. Our findings, however, only pertain to two aspects of a real-time EMA protocol (ie, prompting frequency and sampling length) that may affect participants' compliance [2]. Examples of other important aspects include the quality and complexity of the prompts, effort required to complete each assessment, attractiveness of the interface (ie, developmental appropriateness, and esthetics), and stability of the reporting platforms. To date, however, information on EMA protocol designs is inconsistently reported; therefore, this study was not able to assess the effect of these factors on the reported compliance. In addition, the average compliance rate estimated in this study may be somewhat inflated, as some reviewed studies reportedly excluded low-compliant participants from the analytic datasets used for calculating and reporting compliance. The possibility of overestimating the average compliance rate reflects the methodological limitation of conducting meta-analysis using data extracted from text of published literature. Future research that analyzes data retrieved from each individual study will be able to provide a more precise estimate of participant compliance rate and to further understand how individual participant characteristics affect participant compliance.

Conclusions and Future Directions

Using mobile technologies as data collection platforms in EMA studies has demonstrated generally moderate, but suboptimal, compliance rates among the youth population. In this review, we have further identified that sampling intensity, a possible proximate of participant burden, might impact compliance of participants from different settings. The study results suggest that youth from nonclinical settings may comply better with mobile-EMA protocol with a lowered daily prompting frequency, whereas youth from clinical settings comply better otherwise. Nonetheless, the nonexperimental nature of this review limits our ability to make further recommendations and highlights the need for experimental studies to investigate the impact of these study design factors on participant compliance. Moreover, this review identified several areas of compliance-related results that are currently inconsistently or inadequately reported among the reviewed studies. This suggests the need for thorough reports of participant compliance, which would potentially advance the current understanding of participants' compliance to EMA protocols and to aid development of future EMA study designs. Therefore, we suggest that future studies use the proposed reporting guidelines by Liao et al ([Multimedia Appendix 2](#)) [9].

We further emphasize the importance for future studies to report results in several areas that have been most inconsistently reported. These areas include (1) the reporting of EMA design features that were used to reduce participant burden or potentially improve data quality (eg, minimizing item "over exposure" by administering items in rotated order); (2) the number of prompts delivered and actually received by the

participants, and whether nonresponse was due to technical issues or participant noncompliance; (3) response latency, or the amount of time from prompt signal to prompt answering; (4) distributional characteristics of noncompliance rates (ie, standard deviation and skewness of participant prompt completion rates), participant compliance results based on the full sample to improve the transparency and consistency in reporting prompt response rate; and (5) demographic and

time-varying correlates of EMA compliance. Furthermore, we suggest that future studies should incorporate the time-frame information when defining participant compliance. As one of the central promises of EMA the collection of data with a reduced recall bias, providing this information could aid future studies and meta-analytic reviews to determine the effect of latency on data collected, which may further improve the current understanding of participant compliance.

Acknowledgments

Author CKFW is supported by the University of Southern California Ph.D Fellowship and the National Institute of Health (NCCIH: R01AT008330). Author SS and AAS are supported by the National Institute of Health (NIAMS: R01AR066200 and NIA: R01AG042407). Author DSM is supported by the National Science Foundation CISE/SCH (1521722) and the National Institutes of Health (U24OD23176).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study characteristics.

[[PDF File \(Adobe PDF File\), 45KB](#) - [jmir_v19i4e132_app1.pdf](#)]

Multimedia Appendix 2

Ecological momentary assessment (EMA) study with time-based design.

[[PDF File \(Adobe PDF File\), 45KB](#) - [jmir_v19i4e132_app2.pdf](#)]

Multimedia Appendix 3

Ecological momentary assessment (EMA) study with event-based design.

[[PDF File \(Adobe PDF File\), 25KB](#) - [jmir_v19i4e132_app3.pdf](#)]

References

1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008;4:1-32. [Medline: [18509902](#)]
2. Hufford M. Special methodological challenges and opportunities in ecological momentary assessment. In: Stone AA, editor. *The Science of Real-Time Data Capture: Self-Reports in Health Research*. Oxford: Oxford University Press; 2007:54-75.
3. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials* 2003 Apr;24(2):182-199. [Medline: [12689739](#)]
4. Prensky M. Digital natives, digital immigrants part 1. *On the Horizon* 2001 Sep;9(5):1-6. [doi: [10.1108/10748120110424816](#)]
5. Turner T, Spruijt-Metz D, Wen CK, Hingle MD. Prevention and treatment of pediatric obesity using mobile and wireless technologies: a systematic review. *Pediatr Obes* 2015 Dec;10(6):403-409 [[FREE Full text](#)] [doi: [10.1111/ijpo.12002](#)] [Medline: [25641770](#)]
6. Lenhart A, Duggan M, Perrin A, Stepler R, Rainie L, Parker K. Pew Research Center. 2015 Apr 09. Teens, social media, and technology overview 2015 URL: <http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/> [accessed 2017-04-06] [[WebCite Cache ID 6pW390fzT](#)]
7. Cain AE, Depp CA, Jeste DV. Ecological momentary assessment in aging research: a critical review. *J Psychiatr Res* 2009 Jul;43(11):987-996 [[FREE Full text](#)] [doi: [10.1016/j.jpsychires.2009.01.014](#)] [Medline: [19272611](#)]
8. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE checklist for reporting EMA studies (CREMAS). *J Med Internet Res* 2016 Jun 21;18(6):e151 [[FREE Full text](#)] [doi: [10.2196/jmir.4954](#)] [Medline: [27328833](#)]
9. Liao Y, Skelton K, Dunton G, Bruening M. A Systematic Review of Methods and Procedures Used in Ecological Momentary Assessments of Diet and Physical Activity Research in Youth: An Adapted STROBE Checklist for Reporting EMA Studies (CREMAS). *J Med Internet Res* 2016 Jun 21;18(6):e151 [[FREE Full text](#)] [doi: [10.2196/jmir.4954](#)] [Medline: [27328833](#)]
10. Lipsey MW, Wilson DA. *Practical meta-analysis*. Thousand Oaks, CA: Sage; 2001.

11. Higgins JPT. In: Green S, editor. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. London: The Cochrane Collaboration; 2011.
12. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003 Sep 06;327(7414):557-560 [FREE Full text] [doi: [10.1136/bmj.327.7414.557](https://doi.org/10.1136/bmj.327.7414.557)] [Medline: [12958120](https://pubmed.ncbi.nlm.nih.gov/12958120/)]
13. Liberati A, Altman D, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009 Oct;62(10):e1-34 [FREE Full text] [doi: [10.1016/j.jclinepi.2009.06.006](https://doi.org/10.1016/j.jclinepi.2009.06.006)] [Medline: [19631507](https://pubmed.ncbi.nlm.nih.gov/19631507/)]
14. Whalen CK, Henker B, Ishikawa SS, Jamner LD, Floro JN, Johnston JA, et al. An electronic diary study of contextual triggers and ADHD: get ready, get set, get mad. *J Am Acad Child Adolesc Psychiatry* 2006 Feb;45(2):166-174. [doi: [10.1097/01.chi.0000189057.67902.10](https://doi.org/10.1097/01.chi.0000189057.67902.10)] [Medline: [16429087](https://pubmed.ncbi.nlm.nih.gov/16429087/)]
15. Rosen PJ, Epstein JN, Van OG. I know it when I quantify it: ecological momentary assessment and recurrence quantification analysis of emotion dysregulation in children with ADHD. *Atten Defic Hyperact Disord* 2013 Sep;5(3):283-294 [FREE Full text] [doi: [10.1007/s12402-013-0101-2](https://doi.org/10.1007/s12402-013-0101-2)] [Medline: [23338519](https://pubmed.ncbi.nlm.nih.gov/23338519/)]
16. Rosen PJ, Factor PI. Emotional impulsivity and emotional and behavioral difficulties among children with ADHD: an ecological momentary assessment study. *J Atten Disord* 2015 Sep;19(9):779-793. [doi: [10.1177/1087054712463064](https://doi.org/10.1177/1087054712463064)] [Medline: [23172248](https://pubmed.ncbi.nlm.nih.gov/23172248/)]
17. Stinson JN, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, Dupuis A, et al. Construct validity of a multidimensional electronic pain diary for adolescents with arthritis. *Pain* 2008 Jun;136(3):281-292. [doi: [10.1016/j.pain.2007.07.002](https://doi.org/10.1016/j.pain.2007.07.002)] [Medline: [17723279](https://pubmed.ncbi.nlm.nih.gov/17723279/)]
18. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. *J Med Internet Res* 2013 Mar 08;15(3):e51. [doi: [10.2196/jmir.2350](https://doi.org/10.2196/jmir.2350)] [Medline: [23475457](https://pubmed.ncbi.nlm.nih.gov/23475457/)]
19. Connelly M, Bromberg MH, Anthony KK, Gil KM, Franks L, Schanberg LE. Emotion regulation predicts pain and functioning in children with juvenile idiopathic arthritis: an electronic diary study. *J Pediatr Psychol* 2012;37(1):43-52 [FREE Full text] [doi: [10.1093/jpepsy/jsr088](https://doi.org/10.1093/jpepsy/jsr088)] [Medline: [22037006](https://pubmed.ncbi.nlm.nih.gov/22037006/)]
20. Borus JS, Blood E, Volkening LK, Laffel L, Shrier LA. Momentary assessment of social context and glucose monitoring adherence in adolescents with type 1 diabetes. *J Adolesc Health* 2013 May;52(5):578-583 [FREE Full text] [doi: [10.1016/j.jadohealth.2012.10.003](https://doi.org/10.1016/j.jadohealth.2012.10.003)] [Medline: [23298986](https://pubmed.ncbi.nlm.nih.gov/23298986/)]
21. Helgeson VS, Lopez LC, Kamarck T. Peer relationships and diabetes: retrospective and ecological momentary assessment approaches. *Health Psychol* 2009 May;28(3):273-282 [FREE Full text] [doi: [10.1037/a0013784](https://doi.org/10.1037/a0013784)] [Medline: [19450032](https://pubmed.ncbi.nlm.nih.gov/19450032/)]
22. Khor AS, Gray KM, Reid SC, Melvin GA. Feasibility and validity of ecological momentary assessment in adolescents with high-functioning autism and Asperger's disorder. *J Adolesc* 2014 Jan;37(1):37-46. [doi: [10.1016/j.adolescence.2013.10.005](https://doi.org/10.1016/j.adolescence.2013.10.005)] [Medline: [24331303](https://pubmed.ncbi.nlm.nih.gov/24331303/)]
23. Lewandowski L, Rieger B, Smyth J, Perry L, Gathje R. Measuring post-concussion symptoms in adolescents: feasibility of ecological momentary assessment. *Arch Clin Neuropsychol* 2009 Dec;24(8):791-796. [doi: [10.1093/arclin/acp087](https://doi.org/10.1093/arclin/acp087)] [Medline: [19892712](https://pubmed.ncbi.nlm.nih.gov/19892712/)]
24. Connelly M, Miller T, Gerry G, Bickel J. Electronic momentary assessment of weather changes as a trigger of headaches in children. *Headache* 2010 May;50(5):779-789. [doi: [10.1111/j.1526-4610.2009.01586.x](https://doi.org/10.1111/j.1526-4610.2009.01586.x)] [Medline: [20039960](https://pubmed.ncbi.nlm.nih.gov/20039960/)]
25. Dennis ML, Scott CK, Funk RR, Nicholson L. A pilot study to examine the feasibility and potential effectiveness of using smartphones to provide recovery support for adolescents. *Subst Abuse* 2015;36(4):486-492 [FREE Full text] [doi: [10.1080/08897077.2014.970323](https://doi.org/10.1080/08897077.2014.970323)] [Medline: [25310057](https://pubmed.ncbi.nlm.nih.gov/25310057/)]
26. Garcia C, Hardeman RR, Kwon G, Lando-King E, Zhang L, Genis T, et al. Teenagers and texting: use of a youth ecological momentary assessment system in trajectory health research with latina adolescents. *JMIR Mhealth Uhealth* 2014 Jan 24;2(1):e3 [FREE Full text] [doi: [10.2196/mhealth.2576](https://doi.org/10.2196/mhealth.2576)] [Medline: [25098355](https://pubmed.ncbi.nlm.nih.gov/25098355/)]
27. Kim J, Nakamura T, Kikuchi H, Sasaki T, Yamamoto Y. Co-variation of depressive mood and locomotor dynamics evaluated by ecological momentary assessment in healthy humans. *PLoS One* 2013;8(9):e74979 [FREE Full text] [doi: [10.1371/journal.pone.0074979](https://doi.org/10.1371/journal.pone.0074979)] [Medline: [24058642](https://pubmed.ncbi.nlm.nih.gov/24058642/)]
28. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: a real-time ecological assessment study among adolescents and young adults. *J Abnorm Psychol* 2009 Nov;118(4):816-827 [FREE Full text] [doi: [10.1037/a0016948](https://doi.org/10.1037/a0016948)] [Medline: [19899851](https://pubmed.ncbi.nlm.nih.gov/19899851/)]
29. Miranda R, Monti PM, Ray L, Treloar HR, Reynolds EK, Ramirez J, et al. Characterizing subjective responses to alcohol among adolescent problem drinkers. *J Abnorm Psychol* 2014 Feb;123(1):117-129 [FREE Full text] [doi: [10.1037/a0035328](https://doi.org/10.1037/a0035328)] [Medline: [24661164](https://pubmed.ncbi.nlm.nih.gov/24661164/)]
30. Hoepfner BB, Kahler CW, Gwaltney CJ. Relationship between momentary affect states and self-efficacy in adolescent smokers. *Health Psychol* 2014;33(12):1507-1517. [doi: [10.1037/hea0000075](https://doi.org/10.1037/hea0000075)] [Medline: [25020151](https://pubmed.ncbi.nlm.nih.gov/25020151/)]
31. Gwaltney CJ, Bartolomei R, Colby SM, Kahler CW. Ecological momentary assessment of adolescent smoking cessation: a feasibility study. *Nicotine Tob Res* 2008 Jul;10(7):1185-1190 [FREE Full text] [doi: [10.1080/14622200802163118](https://doi.org/10.1080/14622200802163118)] [Medline: [18629728](https://pubmed.ncbi.nlm.nih.gov/18629728/)]

32. Grenard JL, Stacy AW, Shiffman S, Baraldi AN, MacKinnon DP, Lockhart G, et al. Sweetened drink and snacking cues in adolescents: a study using ecological momentary assessment. *Appetite* 2013 Aug;67:61-73 [FREE Full text] [doi: [10.1016/j.appet.2013.03.016](https://doi.org/10.1016/j.appet.2013.03.016)] [Medline: [23583312](https://pubmed.ncbi.nlm.nih.gov/23583312/)]
33. Scharf DM, Martino SC, Setodji CM, Staplefoote BL, Shadel WG. Middle and high school students' exposure to alcohol- and smoking-related media: a pilot study using ecological momentary assessment. *Psychol Addict Behav* 2013 Dec;27(4):1201-1206 [FREE Full text] [doi: [10.1037/a0032555](https://doi.org/10.1037/a0032555)] [Medline: [23772763](https://pubmed.ncbi.nlm.nih.gov/23772763/)]
34. Dunton G, Dzibur E, Li M, Huh J, Intille S, McConnell R. Momentary assessment of psychosocial stressors, context, and asthma symptoms in hispanic adolescents. *Behav Modif* 2016 Jan;40(1-2):257-280. [doi: [10.1177/0145445515608145](https://doi.org/10.1177/0145445515608145)] [Medline: [26438664](https://pubmed.ncbi.nlm.nih.gov/26438664/)]
35. Dunton GF, Whalen CK, Jamner LD, Henker B, Floro JN. Using ecologic momentary assessment to measure physical activity during adolescence. *Am J Prev Med* 2005 Nov;29(4):281-287. [doi: [10.1016/j.amepre.2005.07.020](https://doi.org/10.1016/j.amepre.2005.07.020)] [Medline: [16242590](https://pubmed.ncbi.nlm.nih.gov/16242590/)]
36. Dunton GF, Liao Y, Intille SS, Spruijt-Metz D, Pentz M. Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones. *Obesity (Silver Spring)* 2011 Jun;19(6):1205-1212 [FREE Full text] [doi: [10.1038/oby.2010.302](https://doi.org/10.1038/oby.2010.302)] [Medline: [21164502](https://pubmed.ncbi.nlm.nih.gov/21164502/)]
37. Dunton GF, Liao Y, Dzibur E, Leventhal AM, Huh J, Gruenewald T, et al. Investigating within-day and longitudinal effects of maternal stress on children's physical activity, dietary intake, and body composition: Protocol for the MATCH study. *Contemp Clin Trials* 2015 Jul;43:142-154 [FREE Full text] [doi: [10.1016/j.cct.2015.05.007](https://doi.org/10.1016/j.cct.2015.05.007)] [Medline: [25987483](https://pubmed.ncbi.nlm.nih.gov/25987483/)]
38. Bromberg MH, Connelly M, Anthony KK, Gil KM, Schanberg LE. Self-reported pain and disease symptoms persist in juvenile idiopathic arthritis despite treatment advances: an electronic diary study. *Arthritis Rheumatol* 2014 Feb;66(2):462-469 [FREE Full text] [doi: [10.1002/art.38223](https://doi.org/10.1002/art.38223)] [Medline: [24504820](https://pubmed.ncbi.nlm.nih.gov/24504820/)]
39. Reid SC, Kauer SD, Dudgeon P, Sanci LA, Shrier LA, Patton GC. A mobile phone program to track young people's experiences of mood, stress and coping. Development and testing of the mobiletype program. *Soc Psychiatry Psychiatr Epidemiol* 2009 Jun;44(6):501-507. [doi: [10.1007/s00127-008-0455-5](https://doi.org/10.1007/s00127-008-0455-5)] [Medline: [19011723](https://pubmed.ncbi.nlm.nih.gov/19011723/)]
40. Dunton GF, Whalen CK, Jamner LD, Floro JN. Mapping the social and physical contexts of physical activity across adolescence using ecological momentary assessment. *Ann Behav Med* 2007 Oct;34(2):144-153. [doi: [10.1080/08836610701566803](https://doi.org/10.1080/08836610701566803)] [Medline: [17927553](https://pubmed.ncbi.nlm.nih.gov/17927553/)]
41. Sokolovsky AW, Mermelstein RJ, Hedeker D. Factors predicting compliance to ecological momentary assessment among adolescent smokers. *Nicotine Tob Res* 2014 Mar;16(3):351-358 [FREE Full text] [doi: [10.1093/ntr/ntt154](https://doi.org/10.1093/ntr/ntt154)] [Medline: [24097816](https://pubmed.ncbi.nlm.nih.gov/24097816/)]
42. Björling EA. The momentary relationship between stress and headaches in adolescent girls. *Headache* 2009 Sep;49(8):1186-1197. [doi: [10.1111/j.1526-4610.2009.01406.x](https://doi.org/10.1111/j.1526-4610.2009.01406.x)] [Medline: [19389139](https://pubmed.ncbi.nlm.nih.gov/19389139/)]
43. van Roekel E, Goossens L, Verhagen M, Wouters S, Engels R, Scholte R. Loneliness, affect, and adolescents' appraisals of company: an experience sampling method study. *J Res Adolesc* 2013 Jun 21;24(2):350-363. [doi: [10.1111/jora.12061](https://doi.org/10.1111/jora.12061)]
44. Ranzenhofer LM, Engel SG, Crosby RD, Anderson M, Vannucci A, Cohen LA, et al. Using ecological momentary assessment to examine interpersonal and affective predictors of loss of control eating in adolescent girls. *Int J Eat Disord* 2014 Nov;47(7):748-757 [FREE Full text] [doi: [10.1002/eat.22333](https://doi.org/10.1002/eat.22333)] [Medline: [25046850](https://pubmed.ncbi.nlm.nih.gov/25046850/)]
45. Croke AH, Reid SC, Kauer SD, McKenzie DP, Hearps SJ, Khor AS, et al. Temporal mood changes associated with different levels of adolescent drinking: using mobile phones and experience sampling methods to explore motivations for adolescent alcohol use. *Drug Alcohol Rev* 2013 May;32(3):262-268. [doi: [10.1111/dar.12034](https://doi.org/10.1111/dar.12034)] [Medline: [23432520](https://pubmed.ncbi.nlm.nih.gov/23432520/)]
46. Suveg C, Payne M, Thomassin K, Jacob M. Electronic diaries: a feasible method of assessing emotional experiences in youth? *J Psychopathol Behav Assess* 2009 Sep 16;32(1):57-67. [doi: [10.1007/s10862-009-9162-0](https://doi.org/10.1007/s10862-009-9162-0)]
47. Rusby JC, Westling E, Crowley R, Light JM. Concurrent and predictive associations between early adolescent perceptions of peer affiliates and mood states collected in real time via ecological momentary assessment methodology. *Psychol Assess* 2013 Mar;25(1):47-60 [FREE Full text] [doi: [10.1037/a0030393](https://doi.org/10.1037/a0030393)] [Medline: [23088201](https://pubmed.ncbi.nlm.nih.gov/23088201/)]
48. Morren M, van DS, Ouwerkerk J, Bensing J. Compliance with momentary pain measurement using electronic diaries: a systematic review. *Eur J Pain* 2009 Apr;13(4):354-365. [doi: [10.1016/j.ejpain.2008.05.010](https://doi.org/10.1016/j.ejpain.2008.05.010)] [Medline: [18603458](https://pubmed.ncbi.nlm.nih.gov/18603458/)]
49. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. *Ann Behav Med* 2002;24(3):236-243. [Medline: [12173681](https://pubmed.ncbi.nlm.nih.gov/12173681/)]
50. Wen CK, Bell B, Spruijt-Metz D. Use of mobile health strategies for the prevention and treatment of childhood obesity. In: Goran MI, editor. *Childhood Obesity: Causes, Consequences Intervention Approaches*. United States: Taylor & Francis Group; Oct 26, 2016:431-441.
51. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. *Sports Med* 2014 May;44(5):671-686. [doi: [10.1007/s40279-014-0142-5](https://doi.org/10.1007/s40279-014-0142-5)] [Medline: [24497157](https://pubmed.ncbi.nlm.nih.gov/24497157/)]
52. Steele R. An overview of the state of the art of automated capture of dietary intake information. *Crit Rev Food Sci Nutr* 2015;55(13):1929-1938. [doi: [10.1080/10408398.2013.765828](https://doi.org/10.1080/10408398.2013.765828)] [Medline: [24950017](https://pubmed.ncbi.nlm.nih.gov/24950017/)]
53. Schüz N, Cianchi J, Shiffman S, Ferguson SG. Novel technologies to study smoking behavior: current developments in ecological momentary assessment. *Curr Addict Rep* 2015 Jan 18;2(1):8-14. [doi: [10.1007/s40429-015-0039-x](https://doi.org/10.1007/s40429-015-0039-x)]

54. Mehrotra A, Pejovic V, Vermeulen J, Hendley R, Musolesi M. My phone and me: understanding people's receptivity to mobile notifications. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 May 07 Presented at: 2016 CHI Conference on Human Factors in Computing Systems; May 5-12, 2016; Santa Clara p. 1021-1032. [doi: [10.1145/2858036.2858566](https://doi.org/10.1145/2858036.2858566)]
55. Poppinga B, Heuten W, Boll S. Sensor-based identification of opportune moments for triggering notifications. IEEE Pervasive Comput 2014 Jan;13(1):22-29. [doi: [10.1109/MPRV.2014.15](https://doi.org/10.1109/MPRV.2014.15)]
56. Ko M, Choi S, Yatani K, Lee U. Lock n'LoL: group-based limiting assistance app to mitigate smartphone distractions in group activities. 2016 May 07 Presented at: The 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; Santa Clara p. 998-1010. [doi: [10.1145/2858036.2858568](https://doi.org/10.1145/2858036.2858568)]
57. Sarker H, Sharmin M, Ali AA, Rahman M, Bari R, Hossain SM, et al. Assessing the availability of users to engage in just-in-time intervention in the natural environment. 2014 Sep 13 Presented at: The 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13-17, 2014; Seattle p. 909-920. [doi: [10.1145/2632048.2636082](https://doi.org/10.1145/2632048.2636082)]
58. Hayes TL, Cobbinah K, Dishongh T, Kaye JA, Kimel J, Labhard M, et al. A study of medication-taking and unobtrusive, intelligent reminding. Telemed J E Health 2009 Oct;15(8):770-776 [FREE Full text] [doi: [10.1089/tmj.2009.0033](https://doi.org/10.1089/tmj.2009.0033)] [Medline: [19780692](https://pubmed.ncbi.nlm.nih.gov/19780692/)]
59. Spruijt-Metz D, Belcher BR, Hsu Y, McClain AD, Chou C, Nguyen-Rodriguez S, et al. Temporal relationship between insulin sensitivity and the pubertal decline in physical activity in peripubertal hispanic and African American females. Diabetes Care 2013 Jul 11;36(11):3739-3745. [doi: [10.2337/dc13-0083](https://doi.org/10.2337/dc13-0083)] [Medline: [23846812](https://pubmed.ncbi.nlm.nih.gov/23846812/)]
60. Gorin A, Stone A. Recall biases and cognitive errors in retrospective self-reports: a call for momentary assessments. Handbook of Health Psychology 2001 Jan 01;23:405-413.

Abbreviations

ANOVA: analysis of variance

EMA: ecological momentary assessment

ES: effect size

ESM: experience sampling methods

GPS: Global Positioning System

HFASD: high function autism and Asperger's syndrome

ICC: intraclass correlation coefficient

IQ: intelligence quotient

JIA: juvenile idiopathic arthritis

PDA: personal digital assistant

PRISMA: preferred reporting items for systemic review and meta-analysis

T1D: type 1 diabetes

Edited by G Eysenbach; submitted 30.09.16; peer-reviewed by M Businelle, F Lamers, Y Liao; comments to author 03.11.16; revised version received 15.12.16; accepted 08.02.17; published 26.04.17.

Please cite as:

Wen CKF, Schneider S, Stone AA, Spruijt-Metz D

Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis

J Med Internet Res 2017;19(4):e132

URL: <http://www.jmir.org/2017/4/e132/>

doi: [10.2196/jmir.6641](https://doi.org/10.2196/jmir.6641)

PMID: [28446418](https://pubmed.ncbi.nlm.nih.gov/28446418/)

©Cheng K Fred Wen, Stefan Schneider, Arthur A Stone, Donna Spruijt-Metz. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

#WhyWeTweetMH: Understanding Why People Use Twitter to Discuss Mental Health Problems

Natalie Berry^{1,2}, BSc (Hons), MSc; Fiona Lobban³, BA (Hons), DClinPsy, PhD; Maksim Belousov⁴, BSc (Hons); Richard Emsley⁵, BSc (Hons), PhD; Goran Nenadic^{1,4}, BSc (Hons), MSc, PhD; Sandra Bucci^{2,6}, BSc (Hons), ClinPsyD

¹Health eResearch Centre, University of Manchester, Manchester, United Kingdom

²Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

³Spectrum Centre for Mental Health Research, School of Health and Medicine, Lancaster University, Lancaster, United Kingdom

⁴School of Computer Science, University of Manchester, Manchester, United Kingdom

⁵Centre for Biostatistics, Division of Population Health, Health Services Research & Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

⁶Manchester Academic Health Science Centre, Greater Manchester Mental Health NHS Foundation Trust, Manchester, United Kingdom

Corresponding Author:

Natalie Berry, BSc (Hons), MSc
Division of Psychology and Mental Health
School of Health Sciences, Faculty of Biology, Medicine and Health
University of Manchester
Room S42, Zochonis Building
Brunswick Street
Manchester, M13 9PL
United Kingdom
Phone: 44 1613060428
Fax: 44 1613060406
Email: natalie.berry@manchester.ac.uk

Abstract

Background: Use of the social media website Twitter is highly prevalent and has led to a plethora of Web-based social and health-related data available for use by researchers. As such, researchers are increasingly using data from social media to retrieve and analyze mental health-related content. However, there is limited evidence regarding why people use this emerging platform to discuss mental health problems in the first place.

Objectives: The aim of this study was to explore the reasons why individuals discuss mental health on the social media website Twitter. The study was the first of its kind to implement a study-specific hashtag for research; therefore, we also examined how feasible it was to circulate and analyze a study-specific hashtag for mental health research.

Methods: Text mining methods using the Twitter Streaming Application Programming Interface (API) and Twitter Search API were used to collect and organize tweets from the hashtag #WhyWeTweetMH, circulated between September 2015 and November 2015. Tweets were analyzed thematically to understand the key reasons for discussing mental health using the Twitter platform.

Results: Four overarching themes were derived from the 132 tweets collected: (1) sense of community; (2) raising awareness and combatting stigma; (3) safe space for expression; and (4) coping and empowerment. In addition, 11 associated subthemes were also identified.

Conclusions: The themes derived from the content of the tweets highlight the perceived therapeutic benefits of Twitter through the provision of support and information and the potential for self-management strategies. The ability to use Twitter to combat stigma and raise awareness of mental health problems indicates the societal benefits that can be facilitated via the platform. The number of tweets and themes identified demonstrates the feasibility of implementing study-specific hashtags to explore research questions in the field of mental health and can be used as a basis for other health-related research.

(*J Med Internet Res* 2017;19(4):e107) doi:[10.2196/jmir.6173](https://doi.org/10.2196/jmir.6173)

KEYWORDS

mental health; Twitter; social media

Introduction

Background

Use of social media websites such as Facebook and Twitter is commonplace, with around 65% of American adults [1] and 66% of British adults [2] reporting ownership of at least one active social media account. High rates of social media use are also evident by individuals who experience mental health problems [3,4]. Research in the field of social media and mental health has largely focused on the potential harm of social media engagement. For example, researchers have observed or empirically evidenced associations between social media use and the occurrence and exacerbation of experiences associated with psychosis [5-7], mood disorders [8-10] personality disorders [10], eating disorders [11,12], and obsessive compulsive disorder [13]. However, others have reported that there are no associations between mental health problems and social media use and, in some cases, significant improvements in social functioning have been observed following social media engagement [14-16]. Mixed and correlational findings in the field and limitations in the methodological design of studies highlight the infancy of our understanding of the relationship between social media use and mental health [17-19]. In addition, much of the current research has focused on the use of Facebook, rather than Twitter, but the nature of the two sites and users differ extensively. For example, a recent comparative analysis of user behavior of individuals with Facebook and Twitter accounts demonstrated no significant overlap between Facebook “friends” and Twitter “followers,” and reported that Facebook was often used as the main platform for communication, whereas Twitter was used as a secondary platform [20]. In addition, user preference for the two different social media platforms has been found to differ based on user personality traits [21]. Therefore, caution needs to be taken when applying findings relating to Facebook to Twitter use.

Twitter (www.Twitter.com) is a popular microblogging site, with 313 million monthly users [22]. Twitter users are able to post 140-character limit posts or “tweets,” which others can respond to via retweeting, replying, or liking posts [23]. Such posts are often publicly accessible and, therefore, available for collection and analysis by researchers. As such, recent studies have collected tweets that included hashtags such as #depression, #schizophrenia, and #dearmentalhealthprofessionals to analyze

mental health-related attitudes and experiences [24-26]. A recent editorial argued that the use of mental health-related hashtags facilitates connections, enables sharing without barriers, and provides the opportunity to voice opinions [27]. Furthermore, a mental health ambassador and educator with lived experience described the “helping hands” of Twitter that can guide people to safety [28]. However, there is little empirical research examining reasons why individuals use Twitter to discuss mental health problems.

Aims of the Study

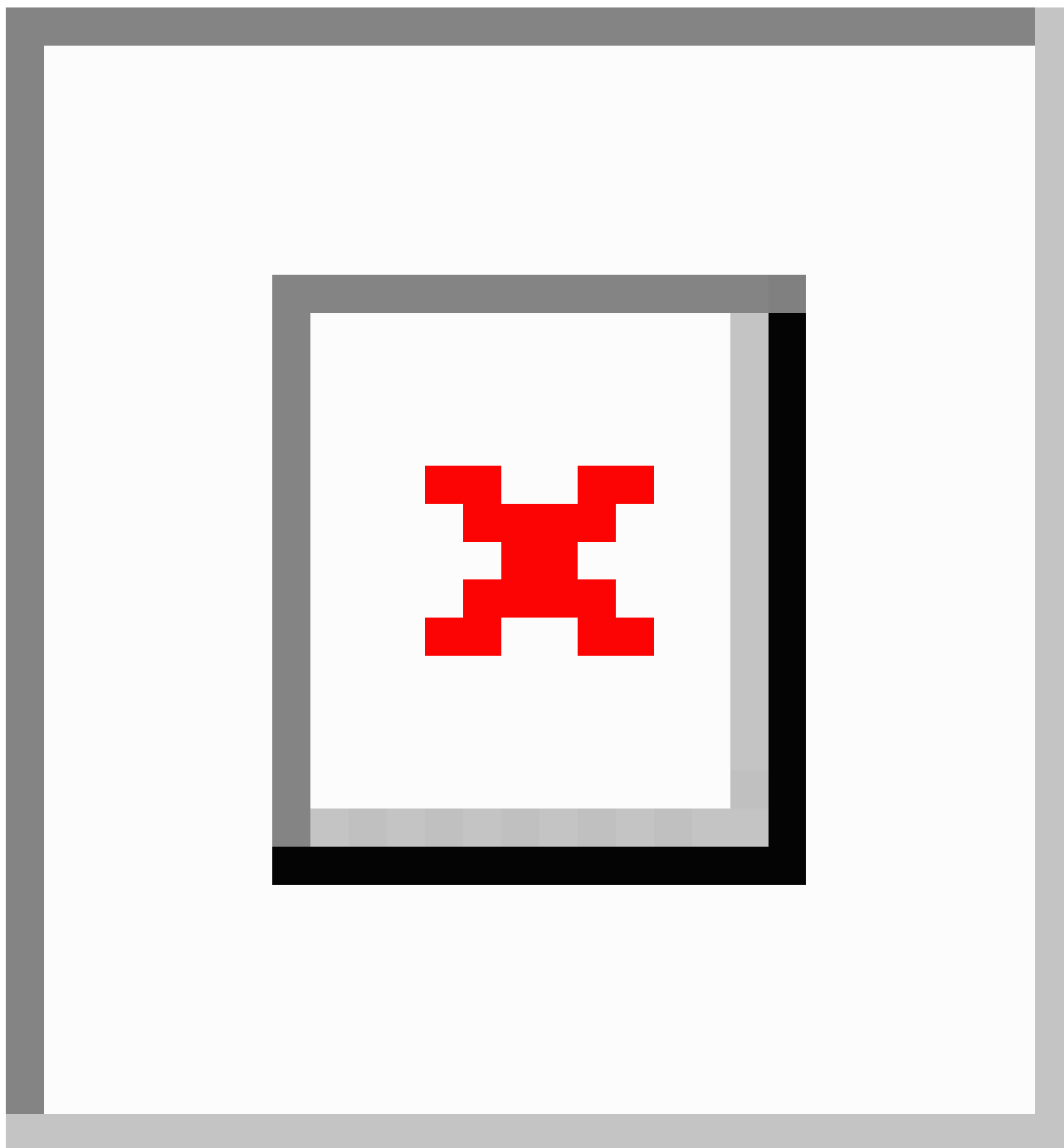
This study implemented the hashtag #WhyWeTweetMH to (1) examine why people use Twitter to discuss mental health problems and (2) investigate whether it is feasible for researchers to directly implement a Twitter hashtag that yields meaningful data for analysis. Twitter was specifically chosen as the social media platform of interest due to the prevalence and popularity of discourse surrounding mental health that is evident on the website.

Methods

Hashtag Development

Twitter allows users to post any information that they wish to share in the form of a 140-character tweet. Tweets posted by users can be “retweeted” so that any tweet an individual wishes to share can be posted on their Twitter profile for their followers to see. Twitter also affords users the opportunity to include hashtags within tweets, which can facilitate communication about, and efficient search for, a specific topic. To this end, the hashtag #WhyWeTweetMH was selected to be circulated on Twitter by the research team (see Figure 1). The decision to use the hashtag #WhyWeTweetMH was based on a number of discussions within the research team. The initial hashtag #WhyWeTweet was developed due to the small number of characters that would be used within response tweets and the alliterative and, therefore, memorable phrasing used. Additionally, popular mental health-related hashtags such as #MHAwareness, #MHcare, and #MHservices use the acronym “MH” to refer mental health on the platform. Therefore, the letters MH were added during the development of the hashtag to ensure that the users were aware that the study was seeking reasons for discussing mental health specifically on Twitter.

Figure 1. Example circulation tweet on the social media website Twitter using the #WhyWeTweetMH hashtag, research study disclaimer, and link for additional information.



Data Collection

The first author (NB) posted the circulation tweet on the researcher's own Twitter page; this was retweeted by other members of the research team. The researcher then individually contacted various mental health charities, campaigners, and advocates asking them to retweet information about the study. Initially, we were only seeking to collect responses from people with current or past experiences of mental health problems; however, some responses were written from other perspectives; for example, academics, clinicians, and charities. Therefore, it was decided that any tweet including the hashtag #WhyWeTweetMH would be analyzed. Frequent attempts were made by the research team to circulate the hashtag until no new

tweets were posted including #WhyWeTweetMH. Collection of tweets using the hashtag occurred between September 2015 and November 2015.

Tweets were automatically collected and stored in a password-protected database. We used both the Twitter Streaming Application Programming Interface (API), for real-time data [29] collection, and the Search API for daily data collection [30] to minimize the risk of missing data due to any network connection failures. This approach ensured that if network errors resulted in a loss of real-time data, past data could still be obtained through the Search API. The hashtag #WhyWeTweetMH was used as the search and streaming keyword.

Data Analysis

Once data collection was completed, all tweets including the hashtag #WhyWeTweetMH were imported to a password-protected Excel (Microsoft) file for qualitative thematic analysis. The Twitter handles (usernames) of users were removed to protect anonymity. During this process, retweets and any tweets posted to circulate the hashtag were removed. In addition, user geolocation was also collected and stored in a password-protected file.

Tweets containing #WhyWeTweetMH were visually inspected several times for common terms. Thematic analysis was used to identify the key reasons that users gave for discussing mental health problems on Twitter. The research team conducting the analysis consisted of a researcher with limited clinical experience and two clinical academics with extensive experience working with people with mental health problems. To ensure transparency and reliability, all tweets were read and analyzed by two members of the research team (NB and SB), who developed an emergent coding scheme to arrange the data. A hierarchical structure of descriptive headings and subheadings was produced and compared across all tweets. This structure was independently reviewed by FL and, as recommended by Turpin and colleagues [31], these categories were discussed again and refined with all members of the research team.

Ethical Considerations

Ethical issues surrounding research using social media websites are complex and some individuals may perceive researchers “lurking” on Internet communities as intrusive [32]. However, as Twitter is considered a public platform, content posted on Twitter is publicly available to be used for research purposes [26]. Throughout the development and implementation of this study, several guidelines for Internet research were consulted and adhered to, specifically, the Association of Internet Researchers [33], the British Psychological Society [34], and INVOLVE [35]. In addition, ethical approval was granted by the University of Manchester Research Ethics Committee (ref: 15347). The use of these guidelines and consultation with the local ethics committee during the development process enabled the formulation of several methodological considerations to protect the safety and privacy of Twitter users.

As this was the first study of its kind to implement a mental health hashtag for research purposes, rather than collecting data from an already trending hashtag, new methods were employed to ensure that the study was ethically sound. First, the tweet circulating the hashtag #WhyWeTweetMH explicitly stated that the hashtag was being used for research purposes. The tweet circulating the hashtag also contained a link to an information sheet, which detailed a list of helplines that individuals would be able to contact should they require further support. The

hashtag was also monitored several times a day to ensure that any potentially offensive or bullying comments to individuals who tweeted using the hashtag could be reported to Twitter. However, it is of note that none of the tweets identified contained offensive or bullying responses. Individual Twitter handles (usernames) were removed from the tweets to maintain confidentiality and, after thematic analysis, all tweets for presentation and publication purposes were paraphrased to ensure anonymity. Tweets were paraphrased by NB and reviewed by SB to confirm that the paraphrased tweets accurately reflected the content of the original tweets. Each paraphrased tweet was inputted into search engines and the Twitter search function to ensure that users’ profiles were not identified in the search results. In line with recommendations for the reporting of research conducted via Twitter [36], a full list of paraphrased Tweets is available as a [Multimedia Appendix 1](#).

Results

Tweet Features

After the removal of retweets, a total of 132 original tweets posted by 90 different users contained #WhyWeTweetMH. The participant information sheet from the study was viewed 145 times during the study period. Respondents were located in the United Kingdom (n=44), the United States (n=22), Canada (n=4), South Africa (n=1), and Australia (n=1). The remaining users either listed a fictional location or did not have their location available (n=18). Respondents’ tweets were analyzed to determine whether experiences of using Twitter to discuss mental health problems were from personal or professional perspectives. The majority of the Twitter users who responded to the hashtag were identified from their responses as having personal experiences of mental health problems (n=50) and others were identified as working in the field of mental health (n=8). Inferences about user experience could not be made for the remaining respondents (n=32). We identified 4 themes and 11 associated subthemes. Some tweets presented several reasons for tweeting about mental health and are, therefore, applicable to multiple themes and subthemes. The frequency of themes and subthemes derived from the data, words used within subthemes, and the numbers of retweets and “likes” for each subtheme are presented in [Table 1](#).

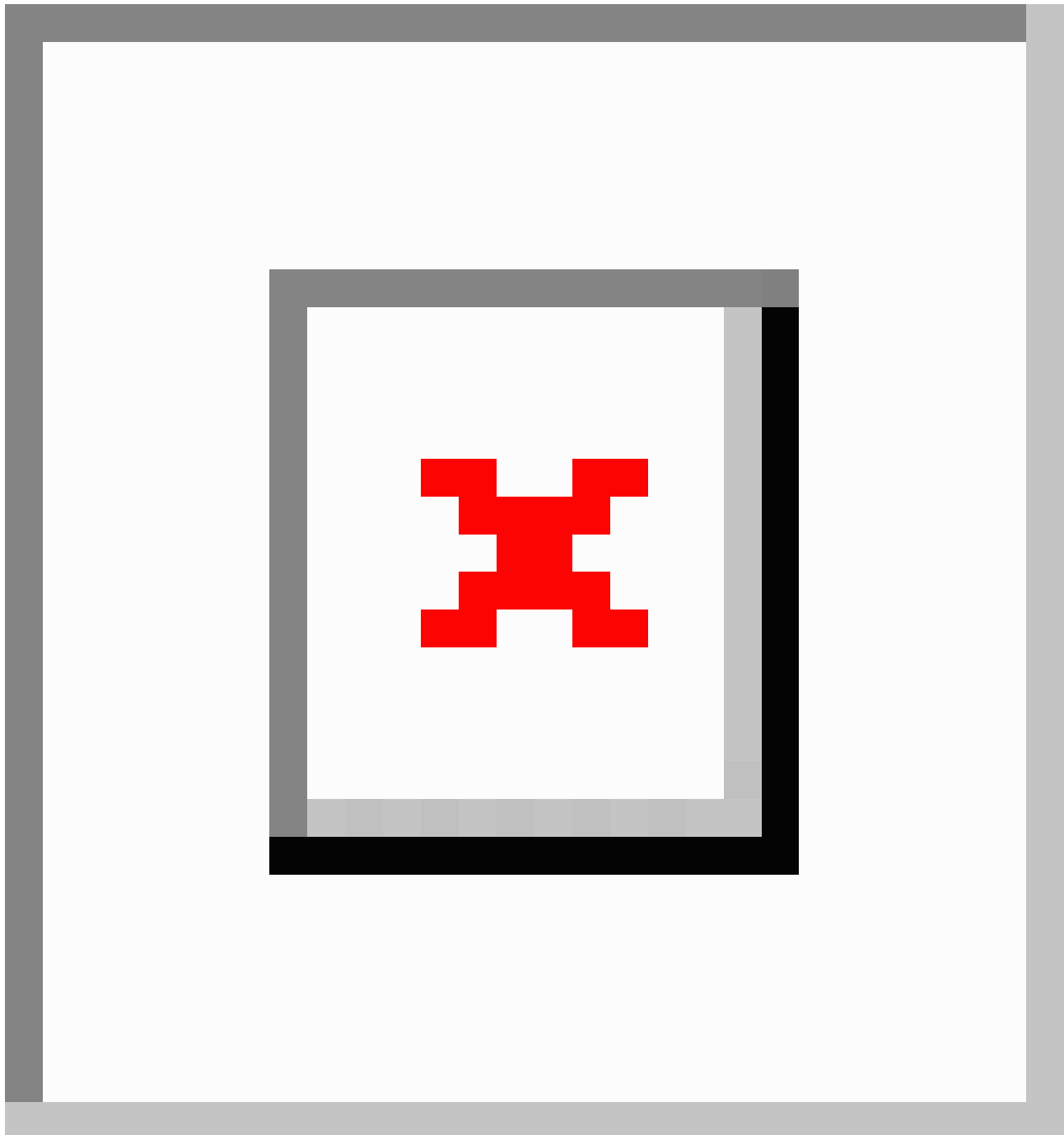
Information regarding the frequency of common words in the tweets collected was recorded through splitting the text into single words. The words most frequently mentioned in the tweets were (1) stigma; (2) support or supporting; (3) alone; (4) connect; (5) awareness; (6) others; and (7) share or sharing appear in [Figure 2](#), which was created using QSR International’s NVivo 11 software. These terms reflect some of the key themes and subthemes resulting from the tweets.

Table 1. Themes and subthemes derived from the #WhyWeTweetMHhashtag and the associated frequencies of retweets, “likes,” and popular words for each of the subthemes.

Theme and subthemes	Tweet frequency	Number of retweets	Proportion of tweets retweeted n (%)	Number of “likes”	Proportion of tweets “liked” n (%)	Commonly used words within subthemes	Word frequency
Sense of community							
To connect or socialize and reduce isolation	47	55	25 (53)	61	27 (57)	Alone	13
						Connect	11
						Others	9
						People or ppl	8
						Friends	4
						Isolation	4
						Community	4
To send and receive messages of hope and support	35	44	20 (57)	49	18 (51)	Support or supporting	14
						Hope	5
						Help or helps	4
						Hugs	3
To share and receive information	21	20	9 (43)	15	9 (43)	Share or sharing	9
						Information or info	7
						Resources	5
						Learn	5
Stigma and awareness							
To combat stigma	23	40	15 (65)	29	14 (61)	Stigma	19
						Combat	3
						Eradicate	2
To raise awareness	22	42	13 (59)	27	13 (59)	Awareness	11
						Raise	7
						Educate	4
						Understanding	3
To fight and campaign	11	29	9 (82)	15	7 (64)	Services	3
						Advocate	3
						Improve	3
Safe space for expression							
To share honest experiences without feeling judged	32	23	15 (47)	44	16 (50)	Experiences	8
						Share or sharing	8
						Honest	5
						Judge	4
						Safe	3
To vent, give people a voice, and feel heard	21	25	12 (57)	29	9 (43)	Vent	6
						Express	3
						Frustration	2
Perceived benefit over Twitter and other social media platforms	7	6	4 (57)	17	5 (71)	Facebook	2
						Networking or media	2
Coping and empowerment							
To escape	4	0	0 (0)	2	2 (50)	Escape	2
						Distract	1
						Suspend	1

Theme and subthemes	Tweet frequency	Number of retweets	Proportion of tweets retweeted n (%)	Number of "likes"	Proportion of tweets "liked" n (%)	Commonly used words within subthemes	Word frequency
Empowering form of self-monitoring and management	11	7	4 (36)	10	7 (64)	Empower	1
						Resilience	1
						Monitor	1
						Manage	1

Figure 2. Word cloud reflecting the frequency of common words identified in tweets including the hashtag #WhyWeTweetMH.



Theme 1: Tweeting About Mental Health Provides a Sense of Community

The overall sense of a “Twitter community” was evident through the explicit use of the word “community” in some of the tweets.

The terms “Twitter friends” and “virtual hugs” were also prevalent, which implies a reciprocal relationship within the Twitter mental health community. In total, 51% (42/83) of the tweets included within this theme were retweeted and 53% (44/83) received “likes” from other users.

Tweeting to Connect, Socialize, and Reduce Isolation

Some users expressed that Twitter is the only setting where they are able to connect and socialize with others. The use of Twitter for some people as the sole avenue for communication may be due to the accessible nature of websites:

*Because it is...the one space I can speak with people.
Because I am with friends even when I am unable to go out.*

Users also commented that tweeting about mental health provided them with their only opportunity to connect with others with shared understanding:

I am able to communicate with other people with the same experiences...

Additionally, tweeting about mental health was perceived by users as a way to reduce feelings of isolation and loneliness and allowed them to show others and themselves that they are not alone:

*...so I do not feel that I am the only person with MH concerns.
If it helps even one person recognize they aren't alone in their pain.*

Tweeting to Send and Receive Messages of Support and Hope

Support was detailed in many tweets, with users expressing that they often tweet about mental health to provide and receive messages of hope and support:

*I enjoy supporting people and receiving support from them.
Suicide might be complicated but extending a hand to someone is simple and it may save their life...
I tweet humour to show people that there is light at the end of the dark tunnel...*

Some users also expressed that by sharing their experiences on Twitter, they could help people who were facing similar challenges:

So that, perhaps, my tweets and experiences may help others. Even if it's only one person.

Twitter was perceived as an accessible avenue for support due to the instantaneous nature of the responses:

I am able to get fast, insightful and appreciated support in a way that's meaningful for me...

Tweeting to Share and Receive Information

Some users also reported that tweeting about mental health provided them with the opportunity to ask questions, learn more about mental health, and to seek and signpost useful resources:

*To advise, support, and to ask questions...
Tweeting about mental health helps people to obtain helpful info they would not normally hear about.*

Theme 2: Tweeting About Mental Health to Combat Stigma and Raise Awareness

Tweets that contained information about using Twitter to raise awareness of mental health problems, combat stigma, and fight and campaign received the largest proportion of retweets (65%, 31/48) and “likes” (58%, 28/48). Additionally, 82% (9/11) of the tweets in the subtheme tweeting to fight and campaign received retweets and 64% (7/11) were “liked.” The high proportion of responses to such tweets may be due to other users sharing these tweets in an attempt to further campaign for people experiencing mental health problems and the high number of followers that campaigners or advocates may have on the site. Tweeting to raise awareness, combat stigma, and fight and campaign were often detailed by users as an attempt to achieve a final outcome; for example, developing empathy and compassion, to show people that others care and to provide hope for the future:

*To raise awareness, stop stigma, create networks, & build empathy & compassion is #WhyWeTweetMH.
To bring buried, misjudged, and shameful disorders out of the darkness. To relieve the struggle of those still to come.*

Tweeting to Combat Stigma

Many of the antistigma tweets contained particularly strong and emotive language such as “combat,” “demolish,” and “fight” to describe the concept of using Twitter to address stigmatizing attitudes. In addition, some Twitter users embedded the already popular hashtag #endthestigma into their #WhyWeTweetMH responses:

*We do not only need to challenge stigma we have to eliminate stigma...
...to attempt to battle stigma...*

Tweeting to Raise Awareness

Some users reported that Twitter was a common starting point for important conversations about mental health problems:

*...begin speaking about what's actually important...
To begin the conversation and open the barriers...*

Additionally, some tweets also included already trending mental health awareness hashtags; for example, #mentalhealthawareness and #everyonesbusiness.

Tweeting to Fight and Campaign

Some people saw Twitter as an avenue for campaigning about mental health, which allowed them to represent others experiencing mental health problems:

*To inform, empower, and inspire. We must advocate for and show others how to advocate for themselves.
...An advocate told me that my voice was required on here to confront the “Master Narratives” about mental health, trauma, and suicide.*

Theme 3: Tweeting About Mental Health Because Twitter Is a Safe Space for Expression

Twitter was perceived as a safe setting in which users could discuss mental health honestly and openly without feeling judged by others. Perceptions of safety in comparison to other social media platforms were also evident in some tweets. On average, just under half of the tweets assigned to this theme were retweeted (48%, 22/46) and over half were “liked” by other users (54%, 25/46).

Tweeting to Share Honest Experiences Without Feeling Judged

Several respondents noted that the perceived anonymity of Twitter allowed them to feel safe and, therefore, felt comfortable in being open and honest about their experiences of mental health problems:

I tweet because I am able to be anonymous so honest...

Some users also stated that Twitter allowed them to share thoughts and feelings relating to mental health on Twitter without feeling judged by others. Although efforts have been made to reduce stigma and judgmental attitudes toward mental health problems, these issues are still prevalent in society [37,38]. However, perceptions of safety and accepting attitudes reported by users suggest that Twitter may provide a protective platform for communication and expression that is, perhaps, not available in everyday life:

...because I'm never dismissed by my Twitter friends as being over sensitive, needing attention, or not making enough of an effort.

Tweeting to Vent, Have a Voice and Feel Heard

Users expressed that tweeting about mental health was a release and provided them with an outlet to voice any worries or concerns they were experiencing:

...When I tweet about mental health it's a release...I also want the world to see how rubbish I feel...

Respondents also reported that they tweeted to share their experiences of the mental health system and service availability:

I can voice my infuriation with the professional support systems or lack of.

I like to tweet when I am angry at mental health services, so that even if nothing is resolved, my complaint is still public.

Twitter was perceived to be a platform on which to vent because some users felt unable to share thoughts and feelings in face-to-face settings with people who they personally knew.

Mainly I use Twitter as a soapbox so I am able to avoid burdening my friends...

Benefit of Twitter Over Other Social Media Websites

Some users reported that discussing mental health on Twitter was more appropriate than other platforms because they did not feel judged by others on Twitter and could avoid the heavily embellished version of peoples' lives evident on Facebook:

I tweet about mental health problems, information, and feelings because no one judges me on Twitter, unlike other social networking sites...

Facebook is the sparkly sunny version of people, Twitter is the authentic version...

Theme 4: Tweeting About Mental Health Is an Empowering Coping Mechanism

A smaller number of users revealed tweeting about mental health as a self-directed coping mechanism, which enabled them to escape from challenges faced in daily life; recognize and reflect on thoughts, feelings, and experiences; and facilitate feelings of empowerment. There were comparatively far fewer retweets of responses included in this theme than the other themes noted (27%, 4/15), although the proportion of “liked” tweets was similar (60%, 9/15).

Tweeting About Mental Health Provides an Opportunity to Escape

Some users reported that using Twitter provides them with the opportunity to escape from the “real-world” and distract themselves from difficult thoughts or feelings:

To distract myself from my mental health. I enjoy being able to laugh and joke on Twitter—that's the part of me that I like...

...interrupt my irrational and obsessive thoughts—it does work.

Tweeting About Mental Health as an Empowering Self-Management Strategy

The concept of using Twitter as a mood monitor was evident in several tweets, as it allowed respondents to express themselves on Twitter and reflect back on the tweets to recognize their thoughts and feelings over time:

I began tweeting so that I will someday be able to look back at how bad things have been, as blogging was too much for me...

My Twitter timeline performs as a sort of mood monitor for myself and those who personally know me...

The potential strength of Twitter as a coping mechanism was evident in some tweets, which stated that Twitter was “cheaper than therapy,” “Twitter saves lives,” and the inclusion of the hashtag #lifehack.

Some users also commented that they felt empowered by tweeting about their mental health, which suggests that tweeting about mental health can be an empowering experience:

Tweeting's empowering...

Discussion

Principal Findings

The aims of this study were to (1) explore reasons why people use Twitter to discuss mental health problems and (2) examine whether study-specific Twitter hashtags can be implemented by researchers as a method for data collection. The collective

experiences noted are indicative of the positive role that Twitter can provide in mental health discussions and the number of tweets collected suggests that the circulation of study-specific hashtags on Twitter is a feasible avenue for investigating mental health-related phenomena.

The content expressed in collected tweets conveyed the notion of a “Twitter community” that allowed communication to flourish, awareness to be raised, stigma to be fought, and support that could be both offered and received. These perceived functions of Twitter support previous assertions that the platform provides a space for mental health-related discussions [26,27] and self-disclosures [39] and the wider literature regarding the social ties, sense of community, and support mechanisms that can be developed when communicating about health and experiences on the Internet [40–42]. Sense of belongingness and integration within a community can benefit an individual’s mental health and may be a protective factor in the development and exacerbation of symptoms associated with mental health problems [43–45]. In addition, social disconnectedness is often associated with higher rates of relapse [46], increased mortality [47], and poorer physical and mental health [48]. Therefore, being a member of a large Twitter mental health community may act as a protective factor by facilitating communication and support. Furthermore, combatting stigma and raising awareness were key reasons identified for tweeting about mental health, which may help foster the sense of community that was evident in the tweets.

The positive evaluations of the Twitter mental health community for the provision of support may, in part, be due to the value of a shared understanding on Twitter. Some users noted that Twitter allows them to communicate and receive support from others with similar experiences. There has been a growing movement in psychological practice toward the inclusion of peer support approaches, whereby individuals with experience of mental health problems provide support for people with similar experiences [49]. The potential value of peer support has been widely discussed in the literature and is associated with improved functioning, empowerment, and confidence [49], reductions in hospital admissions [50], and increased social networks and wellbeing [51]. The notion that social media could provide an accessible avenue for peer support is not new necessarily. A recent commentary regarding social media usage in severe mental health problems reported that social media could facilitate help-seeking behaviors, reciprocal support, and antistigma campaigns [52]. In addition, Naslund and colleagues [53] analyzed comments on videos created by individuals experiencing severe mental health problems on the video sharing platform YouTube (www.youtube.com). The authors reported that there was evidence of naturally occurring peer support within the comments, which provided supportive messages and coping strategies and reduced isolation. As such, the findings from this study support the view that social media websites, in this case, Twitter, could be a valuable tool for people who experience mental health problems.

Many users also noted that they were able to access resources and information on Twitter that they ordinarily would not be able to retrieve. The availability and subsequent use of Web-based material may help to facilitate self-directed

psychoeducation, which is a potentially effective psychological intervention [54]. Therefore, resources on Twitter could be employed by users as a self-directed psychoeducation intervention. In addition, access to Web-based health-related information is reportedly beneficial for improving health behaviors, awareness and care of conditions, and could facilitate help-seeking [55–57]. Individuals experiencing mental health problems, clinicians, and academics could also use Web-based resources shared by other Twitter users to remain informed about recent advances in clinical practice and current research in the field. Some of the tweets that included #WhyWeTweetMH also contained other trending hashtags; for example, #everyonesbusiness and #mentalhealthawareness. The inclusion of such hashtags illustrates the popularity of incorporating mental health hashtags within tweets and supports the notion that hashtags can be an effective method to facilitate communication about specific topics.

The use of Twitter to share experiences of mental health services was also evident in some of the tweets and supports previous conclusions that mental health services could use Twitter to receive feedback on the care that they provide [26]. Users also reported that Twitter allowed them to be open and honest about their experiences. Providing the individual consents, mental health professionals may have the opportunity to review clinically relevant information disclosed by users on Twitter accounts that they may ordinarily feel uncomfortable sharing in a formal clinical setting. However, further research assessing Twitter user and health care professional views toward the collection of clinically relevant information via Twitter is warranted. Additionally, the use of Twitter as a coping mechanism, which is evident in some tweets, suggests that social networking tools may be popular as a component for psychological interventions.

Feasibility of Circulating a Study-Specific Hashtag on Twitter for Research Purposes

Previous research using Twitter hashtags for data collection has relied on the analysis of already trending hashtags [24–26]. Therefore, the second aim of this study was to assess the feasibility of circulating a study-specific hashtag for research purposes. As there was no precedence for what constituted a “sufficient” number of tweets for research purposes, we gathered a sufficient number for qualitative analysis, demonstrating that it is feasible to employ research hashtags on Twitter. Indeed, the ethical integrity of providing a study-specific hashtag that users are aware is being used for research, rather than collecting preexisting data, may negate the potential disadvantages of reduced data. However, when considering the implementation of study-specific hashtags in mental health research, researchers should remain mindful about the ethical considerations associated with asking people to tweet about their mental health and moral issues surrounding the duty of care toward users who choose to share their views. Additionally, further research should seek to identify Twitter users’ views about the collection of their data using Twitter hashtags to determine whether or not they find this approach acceptable or potentially intrusive. Researchers seeking to use study-specific hashtags in future work may also wish to consider creating a specific Twitter account for research studies. The circulation of a research

hashtag on a study-specific Twitter account may allow users to feel more comfortable tweeting about their experiences due to the anonymous nature of a study account; therefore, potentially increasing responses.

Study Strengths

There were some strengths and limitations to the novel methodology employed. First, the ethical integrity of the study was a considerable strength. Specifically, a research disclaimer was included in the tweets circulating #WhyWeTweetMH, usernames were removed from all tweets before analysis, and tweets were paraphrased after analysis for presentation and publication purposes. Additionally, the investigation of tweeting behavior took place in the setting in which the behavior directly occurred, which ensured that respondents to the hashtag were active Twitter users. The use of both the streaming API and search API to collect tweets reduced the likelihood of missing data. The truly interdisciplinary nature of the research team (ie, computer scientists and psychologists) ensured that the approach to collect tweets was technologically and methodologically sound, and the research question, analysis of tweets, and implications for clinical practice were appropriate. Importantly, conducting this research on Twitter allowed people to provide views for a study without the constraints of traditional research such as location, time pressures, and effort required.

Study Limitations

Some users may have chosen not to tweet using #WhyWeTweetMH due to the disclaimer that data would be used for research purposes, which may have led to missing data. Additionally, the use of Twitter to obtain reasons for tweeting about mental health may have led to biased responses of positive experiences, as people who do not use Twitter due to negative experiences will not have been able to detail reasons for not tweeting. As such, future research should seek to explore both the potential positive and negative experiences people have encountered when using Twitter to discuss mental health. The publicly accessible nature of Twitter may have also resulted in some users observing others responses and adapting their reasons for tweeting accordingly. The study also relied on the assumption that respondents actively tweeted about mental health and the 140-character limit of tweets may have prevented users from being able to give an in-depth insight about their reasons for tweeting. Therefore, the amount of material available for an in-depth exploration of tweet content was limited and

could be utilized further as a method to identify a broad sample and purposively select participants from this sample for further questioning. It is generally seen as good practice to ask participants whether they agree that the analyzed data and paraphrased quotes accurately captured discussions during qualitative interviews. However, this is not feasible via Twitter due to the ethical issues surrounding directly contacting individuals tweeting with the hashtag. Additionally, tweets were limited to English-speakers, which may impact on the generalizability of the findings. To prevent the collection, analysis, and dissemination of potentially identifiable information, only user location (city and country) and the content of tweets containing #WhyWeTweetMH were collected. Although the majority of respondents indicated in the tweets that they experienced mental health problems, for some, interpretations regarding personal or professional experience could not be made. Finally, information such as diagnosis, age, and gender could not be collected due to the ethical considerations surrounding the collection of identifiable information without specific user consent.

Conclusions

The number of tweets collected in the study and the thematic analysis applied demonstrates the feasibility of researchers directly implementing a hashtag for mental health research. Furthermore, the unique methodology employed resulted in the development and identification of several ethical considerations to ensure the safety and anonymity of Twitter users. The findings from #WhyWeTweetMH tweets suggest that individuals may actively use Twitter to discuss mental health as way of developing a sense of belonging within a community, accessing support, challenging stigma and raising awareness, sharing experiences, and as an empowering coping mechanism. Future research is planned to explore whether Twitter users are open to their data being used for research purposes and the acceptability of using Twitter as an avenue for evidence-based psychological interventions. In addition, further work regarding clinician views about Twitter use in mental health and how Twitter could help or hinder clinical practice should be considered. Although potential drawbacks of Twitter use must be considered, for example, cyberbullying and Web-based predators, the strong expressions within the tweets suggest that respondents to #WhyWeTweetMH have experienced Twitter as welcoming and supportive and a useful forum for an open and honest dialog about mental health.

Acknowledgments

The authors would like to thank the Twitter users who retweeted and responded to the #WhyWeTweetMH hashtag. NB is a PhD student funded by the MRC Health eResearch Centre (HeRC), Farr Institute, United Kingdom (MR/K006665/1). MB is a PhD student funded by the Centre for Doctoral Training in Computer Science (United Kingdom, EPSRC).

Authors' Contributions

All authors contributed to the development of the research question, study design, and final manuscript. NB completed the first draft of the manuscript, with SB, FL, RE, MB, and GN providing critical feedback. All authors approved and contributed toward the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Paraphrased versions of tweets that included the #WhyWeTweetMH hashtag.

[[PDF File \(Adobe PDF File\), 79KB - jmir_v19i4e107_app1.pdf](#)]

References

1. Pew Research Center. Pewinternet. 2015 Oct 08. Social media usage: 2005-2015 URL: <http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/> [accessed 2015-11-01] [WebCite Cache ID 6i74Dsd3e]
2. Deloitte. 2014. Media consumer 2014: the digital divide URL: <http://www.deloitte.co.uk/mediaconsumer2014/assets/downloads/Deloitte-Media-Consumer-Survey-2014.pdf> [accessed 2016-01-08] [WebCite Cache ID 6i73mRGij]
3. Gowen K, Deschaine M, Gruttadara D, Markey D. Young adults with mental health conditions and social networking websites: seeking tools to build community. *Psychiatr Rehabil J* 2012;35(3):245-250. [doi: [10.2975/35.3.2012.245.250](https://doi.org/10.2975/35.3.2012.245.250)] [Medline: [22246123](https://pubmed.ncbi.nlm.nih.gov/22246123/)]
4. Birnbaum ML, Rizvi AF, Correll CU, Kane JM. Role of social media and the Internet in pathways to care for adolescents and young adults with psychotic disorders and non-psychotic mood disorders. *Early Interv Psychiatry* 2015 Mar 23; Epub ahead of print. [doi: [10.1111/eip.12237](https://doi.org/10.1111/eip.12237)] [Medline: [25808317](https://pubmed.ncbi.nlm.nih.gov/25808317/)]
5. Kalbitzer J, Mell T, Bermpohl F, Rapp MA, Heinz A. Twitter psychosis: a rare variation or a distinct syndrome? *J Nerv Ment Dis* 2014 Aug;202(8):623. [doi: [10.1097/NMD.0000000000000173](https://doi.org/10.1097/NMD.0000000000000173)] [Medline: [25075647](https://pubmed.ncbi.nlm.nih.gov/25075647/)]
6. Krishna N, Fischer BA, Miller M, Register-Brown K, Patchan K, Hackman A. The role of social media networks in psychotic disorders: a case report. *Gen Hosp Psychiatry* 2013;35(5):576.e1-576.e2. [doi: [10.1016/j.genhosppsych.2012.10.006](https://doi.org/10.1016/j.genhosppsych.2012.10.006)] [Medline: [23174460](https://pubmed.ncbi.nlm.nih.gov/23174460/)]
7. Nitzan U, Shoshan E, Lev-Ran S, Fennig S. Internet-related psychosis—a sign of the times. *Isr J Psychiatry Relat Sci* 2011;48(3):207-211 [FREE Full text] [Medline: [22141146](https://pubmed.ncbi.nlm.nih.gov/22141146/)]
8. Lin LY, Sidani JE, Shensa A, Radovic A, Miller E, Colditz JB, et al. Association between social media use and depression among U.S. young adults. *Depress Anxiety* 2016 Apr;33(4):323-331. [doi: [10.1002/da.22466](https://doi.org/10.1002/da.22466)] [Medline: [26783723](https://pubmed.ncbi.nlm.nih.gov/26783723/)]
9. Pantic I, Damjanovic A, Todorovic J, Topalovic D, Bojovic-Jovic D, Ristic S, et al. Association between online social networking and depression in high school students: behavioral physiology viewpoint. *Psychiatr Danub* 2012 Mar;24(1):90-93 [FREE Full text] [Medline: [22447092](https://pubmed.ncbi.nlm.nih.gov/22447092/)]
10. Rosen L, Whaling K, Rab S, Carrier L, Cheever N. Is Facebook creating “iDisorders”? The link between clinical symptoms of psychiatric disorders and technology use, attitudes and anxiety. *Comput Human Behav* 2013 May;29(3):1243-1254. [doi: [10.1016/j.chb.2012.11.012](https://doi.org/10.1016/j.chb.2012.11.012)]
11. Mabe AG, Forney KJ, Keel PK. Do you “like” my photo? Facebook use maintains eating disorder risk. *Int J Eat Disord* 2014 Jul;47(5):516-523. [Medline: [25035882](https://pubmed.ncbi.nlm.nih.gov/25035882/)]
12. Smith AR, Hames JL, Joiner Jr TE. Status update: maladaptive Facebook usage predicts increases in body dissatisfaction and bulimic symptoms. *J Affect Disord* 2013 Jul;149(1-3):235-240. [doi: [10.1016/j.jad.2013.01.032](https://doi.org/10.1016/j.jad.2013.01.032)] [Medline: [23453676](https://pubmed.ncbi.nlm.nih.gov/23453676/)]
13. Lee SL, Park MS, Tam CL. The relationship between Facebook attachment and obsessive-compulsive disorder severity. *Cyberpsychology* 2015 Aug;9(2). [doi: [10.5817/CP2015-2-6](https://doi.org/10.5817/CP2015-2-6)]
14. Jelenchick LA, Eickhoff JC, Moreno MA. “Facebook depression?” social networking site use and depression in older adolescents. *J Adolesc Health* 2013 Jan;52(1):128-130. [doi: [10.1016/j.jadohealth.2012.05.008](https://doi.org/10.1016/j.jadohealth.2012.05.008)] [Medline: [23260846](https://pubmed.ncbi.nlm.nih.gov/23260846/)]
15. Veretilo P, Billick SB. Psychiatric illness and facebook: a case report. *Psychiatr Q* 2012 Sep;83(3):385-389. [doi: [10.1007/s11126-012-9207-5](https://doi.org/10.1007/s11126-012-9207-5)] [Medline: [22274630](https://pubmed.ncbi.nlm.nih.gov/22274630/)]
16. Simoncic TE, Kuhlman KR, Vargas I, Houchins S, Lopez-Duran NL. Facebook use and depressive symptomatology: investigating the role of neuroticism and extraversion in youth. *Comput Human Behav* 2014 Nov 1;40:1-5 [FREE Full text] [Medline: [25861155](https://pubmed.ncbi.nlm.nih.gov/25861155/)]
17. Toseeb U, Inkster B. Online social networking sites and mental health research. *Front Psychiatry* 2015;6:36 [FREE Full text] [doi: [10.3389/fpsy.2015.00036](https://doi.org/10.3389/fpsy.2015.00036)] [Medline: [25814958](https://pubmed.ncbi.nlm.nih.gov/25814958/)]
18. Best P, Manktelow R, Taylor B. Online communication, social media and adolescent wellbeing: a systematic narrative review. *Child Youth Serv Rev* 2014 Jun;41:27-36. [doi: [10.1016/j.childyouth.2014.03.001](https://doi.org/10.1016/j.childyouth.2014.03.001)]
19. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. *JMIR Ment Health* 2016 Nov 23;3(4):e50 [FREE Full text] [doi: [10.2196/mental.5842](https://doi.org/10.2196/mental.5842)] [Medline: [27881357](https://pubmed.ncbi.nlm.nih.gov/27881357/)]
20. Buccafurri F, Lax G, Nicolazzo S, Nocera A. Comparing Twitter and Facebook user behavior: privacy and other aspects. *Comput Human Behav* 2015 Nov;52:87-95. [doi: [10.1016/j.chb.2015.05.045](https://doi.org/10.1016/j.chb.2015.05.045)]
21. Hughes DJ, Rowe M, Batey M, Lee A. A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Comput Human Behav* 2012 Mar;28(2):561-569. [doi: [10.1016/j.chb.2011.11.001](https://doi.org/10.1016/j.chb.2011.11.001)]
22. Twitter. 2016. URL: <https://about.twitter.com/company> [accessed 2016-11-22] [WebCite Cache ID 6mDWcRS4k]

23. Twitter. support.twitter. 2016. Help Center: Using Twitter URL: <https://support.twitter.com/categories/281> [accessed 2017-03-26] [WebCite Cache ID 6pFkr9XFo]
24. Joseph AJ, Tandon N, Yang LH, Duckworth K, Torous J, Seidman LJ, et al. #Schizophrenia: use and misuse on Twitter. *Schizophr Res* 2015 Jul;165(2-3):111-115. [doi: [10.1016/j.schres.2015.04.009](https://doi.org/10.1016/j.schres.2015.04.009)] [Medline: [25937459](https://pubmed.ncbi.nlm.nih.gov/25937459/)]
25. Reavley NJ, Pilkington PD. Use of Twitter to monitor attitudes toward depression and schizophrenia: an exploratory study. *Peer J* 2014;2:e647 [FREE Full text] [doi: [10.7717/peerj.647](https://doi.org/10.7717/peerj.647)] [Medline: [25374786](https://pubmed.ncbi.nlm.nih.gov/25374786/)]
26. Shepherd A, Sanders C, Doyle M, Shaw J. Using social media for support and feedback by mental health service users: thematic analysis of a twitter conversation. *BMC Psychiatry* 2015;15:29. [doi: [10.1186/s12888-015-0408-y](https://doi.org/10.1186/s12888-015-0408-y)]
27. Betton V, Borschmann R, Docherty M, Coleman S, Brown M, Henderson C. The role of social media in reducing stigma and discrimination. *Br J Psychiatry* 2015 Jun;206(6):443-444 [FREE Full text] [doi: [10.1192/bjp.bp.114.152835](https://doi.org/10.1192/bjp.bp.114.152835)] [Medline: [26034176](https://pubmed.ncbi.nlm.nih.gov/26034176/)]
28. Knightsmith P. Huffingtonpost. 2016. Twitter Saved My Life - Why Social Media Isn't All Bad URL: http://www.huffingtonpost.co.uk/dr-pooky-knightsmith/social-media-mental-health_b_8943202.html [accessed 2017-03-26] [WebCite Cache ID 6pFIJvqPC]
29. Twitter. Dev.twitter. The streaming APIs URL: <https://dev.twitter.com/streaming/overview> [accessed 2016-06-08] [WebCite Cache ID 6i74Vxde5]
30. Twitter. Dev.twitter. The Search API URL: <https://dev.twitter.com/rest/public/search> [accessed 2016-06-08] [WebCite Cache ID 6i74b0n7g]
31. Turpin G, Bartley V, Beail N, Seaire J, Slade P, Smith JA. Standards for research projects and theses involving qualitative methods suggested guidelines for trainees and course. *Clinical Psychology Forum* 1997;108:3-7.
32. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. *Br Med J* 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: [11701577](https://pubmed.ncbi.nlm.nih.gov/11701577/)]
33. Association of Internet Researchers (AoIR). AoIR. Ethical decision-making and internet research 2.0 URL: <http://aoir.org/ethics/> [accessed 2016-06-08] [WebCite Cache ID 6i74nSjoZ]
34. British Psychological Society (BPS). BPS. Ethics guidelines for internet-mediated research URL: <http://www.bps.org.uk/news/guidelines-internet-mediated-research> [accessed 2016-06-08] [WebCite Cache ID 6i7508Jua]
35. INVOLVE. Invo. Guidance on the use of social media to actively involve people in research URL: <http://www.invo.org.uk/posttypepublication/guidance-on-the-use-of-social-media/> [accessed 2016-06-08] [WebCite Cache ID 6i755p08M]
36. Rivers CM, Lewis BL. Ethical research standards in a world of big data. *F1000Res* 2014 Aug 21;3:38. [doi: [10.12688/f1000research.3-38.v2](https://doi.org/10.12688/f1000research.3-38.v2)]
37. Pescosolido BA. The public stigma of mental illness: what do we think; what do we know; what can we prove? *J Health Soc Behav* 2013 Mar;54(1):1-21 [FREE Full text] [doi: [10.1177/0022146512471197](https://doi.org/10.1177/0022146512471197)] [Medline: [23325423](https://pubmed.ncbi.nlm.nih.gov/23325423/)]
38. Corker E, Hamilton S, Henderson C, Weeks C, Pinfold V, Rose D, et al. Experiences of discrimination among people using mental health services in England 2008-2011. *Br J Psychiatry Suppl* 2013 Apr;55:s58-s63 [FREE Full text] [doi: [10.1192/bjp.bp.112.112912](https://doi.org/10.1192/bjp.bp.112.112912)] [Medline: [23553696](https://pubmed.ncbi.nlm.nih.gov/23553696/)]
39. Walton S, Rice R. Mediated disclosure on Twitter: the roles of gender and identity in boundary impermeability, valence, disclosure, and stage. *Comput Human Behav* 2013 Jul 2013 Jul;29(4):1465-1474. [doi: [10.1016/j.chb.2013.01.033](https://doi.org/10.1016/j.chb.2013.01.033)]
40. Mikal JP, Rice RE, Abeyta A, DeVilbiss J. Transition, stress and computer-mediated social support. *Computers in Human Behavior* 2013 Sep;29(5):A40-A53. [doi: [10.1016/j.chb.2012.12.012](https://doi.org/10.1016/j.chb.2012.12.012)]
41. 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 Apr 23;15(4):e85 [FREE Full text] [doi: [10.2196/jmir.1933](https://doi.org/10.2196/jmir.1933)] [Medline: [23615206](https://pubmed.ncbi.nlm.nih.gov/23615206/)]
42. Rains SA, Peterson EB, Wright KB. Communicating social support in computer-mediated contexts: a meta-analytic review of content analyses examining support messages shared online among individuals coping with illness. *Commun Monogr* 2015 Mar 17;82(4):403-430. [doi: [10.1080/03637751.2015.1019530](https://doi.org/10.1080/03637751.2015.1019530)]
43. Fowler K, Wareham-Fowler S, Barnes C. Social context and depression severity and duration in Canadian men and women: exploring the influence of social support and sense of community belongingness. *J Appl Soc Psychol* 2013 May 03;43:E85-E96. [doi: [10.1111/jasp.12050](https://doi.org/10.1111/jasp.12050)]
44. Kitchen P, Williams A, Chowhan J. Sense of community belonging and health in Canada: a regional analysis. *Soc Indic Res* 2011 Mar 26;107(1):103-126. [doi: [10.1007/s11205-011-9830-9](https://doi.org/10.1007/s11205-011-9830-9)]
45. Tew J, Ramon S, Slade M, Bird V, Melton J, Le Boutillier C. Social factors and recovery from mental health difficulties: a review of the evidence. *Br J Soc Work* 2011 Jun 15;42(3):443-460. [doi: [10.1093/bjsw/bcr076](https://doi.org/10.1093/bjsw/bcr076)]
46. Hultman CM, Wieselgren IM, Ohman A. Relationships between social support, social coping and life events in the relapse of schizophrenic patients. *Scand J Psychol* 1997 Mar;38(1):3-13. [Medline: [9104101](https://pubmed.ncbi.nlm.nih.gov/9104101/)]
47. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. *PLoS Med* 2010 Jul;7(7):e1000316 [FREE Full text] [doi: [10.1371/journal.pmed.1000316](https://doi.org/10.1371/journal.pmed.1000316)] [Medline: [20668659](https://pubmed.ncbi.nlm.nih.gov/20668659/)]
48. Cornwell EY, Waite LJ. Social disconnectedness, perceived isolation, and health among older adults. *J Health Soc Behav* 2009 Mar;50(1):31-48 [FREE Full text] [Medline: [19413133](https://pubmed.ncbi.nlm.nih.gov/19413133/)]

49. Resnick SG, Rosenheck RA. Integrating peer-provided services: a quasi-experimental study of recovery orientation, confidence, and empowerment. *Psychiatr Serv* 2008 Nov;59(11):1307-1314. [doi: [10.1176/appi.ps.59.11.1307](https://doi.org/10.1176/appi.ps.59.11.1307)] [Medline: [18971407](https://pubmed.ncbi.nlm.nih.gov/18971407/)]
50. Sledge WH, Lawless M, Sells D, Wieland M, O'Connell MJ, Davidson L. Effectiveness of peer support in reducing readmissions of persons with multiple psychiatric hospitalizations. *Psychiatr Serv* 2011 May;62(5):541-544. [doi: [10.1176/ps.62.5.pss6205_0541](https://doi.org/10.1176/ps.62.5.pss6205_0541)] [Medline: [21532082](https://pubmed.ncbi.nlm.nih.gov/21532082/)]
51. Walker G, Bryant W. Peer support in adult mental health services: a metasynthesis of qualitative findings. *Psychiatr Rehabil J* 2013 Mar;36(1):28-34. [doi: [10.1037/h0094744](https://doi.org/10.1037/h0094744)] [Medline: [23477647](https://pubmed.ncbi.nlm.nih.gov/23477647/)]
52. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. *Epidemiol Psychiatr Sci* 2016 Apr;25(2):113-122 [FREE Full text] [doi: [10.1017/S2045796015001067](https://doi.org/10.1017/S2045796015001067)] [Medline: [26744309](https://pubmed.ncbi.nlm.nih.gov/26744309/)]
53. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. *PLoS One* 2014;9(10):e110171 [FREE Full text] [doi: [10.1371/journal.pone.0110171](https://doi.org/10.1371/journal.pone.0110171)] [Medline: [25333470](https://pubmed.ncbi.nlm.nih.gov/25333470/)]
54. Bond K, Anderson IM. Psychoeducation for relapse prevention in bipolar disorder: a systematic review of efficacy in randomized controlled trials. *Bipolar Disord* 2015 Jun;17(4):349-362. [doi: [10.1111/bdi.12287](https://doi.org/10.1111/bdi.12287)] [Medline: [25594775](https://pubmed.ncbi.nlm.nih.gov/25594775/)]
55. 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](https://doi.org/10.1016/j.ijmedinf.2005.07.032)] [Medline: [16125453](https://pubmed.ncbi.nlm.nih.gov/16125453/)]
56. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. *J Med Internet Res* 2014;16(3):e66 [FREE Full text] [doi: [10.2196/jmir.3103](https://doi.org/10.2196/jmir.3103)] [Medline: [24594922](https://pubmed.ncbi.nlm.nih.gov/24594922/)]
57. Win KT, Hassan NM, Bonney A, Iverson D. Benefits of online health education: perception from consumers and health professionals. *J Med Syst* 2015 Mar;39(3):27. [doi: [10.1007/s10916-015-0224-4](https://doi.org/10.1007/s10916-015-0224-4)] [Medline: [25666928](https://pubmed.ncbi.nlm.nih.gov/25666928/)]

Abbreviations

API: Application Programming Interface

Edited by G Eysenbach; submitted 08.06.16; peer-reviewed by J Torous, M Lahti, R Rice; comments to author 26.10.16; revised version received 28.11.16; accepted 22.01.17; published 05.04.17.

Please cite as:

Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S

#WhyWeTweetMH: Understanding Why People Use Twitter to Discuss Mental Health Problems

J Med Internet Res 2017;19(4):e107

URL: <http://www.jmir.org/2017/4/e107/>

doi: [10.2196/jmir.6173](https://doi.org/10.2196/jmir.6173)

PMID: [28381392](https://pubmed.ncbi.nlm.nih.gov/28381392/)

©Natalie Berry, Fiona Lobban, Maksim Belousov, Richard Emsley, Goran Nenadic, Sandra Bucci. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 05.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Understanding Health Care Social Media Use From Different Stakeholder Perspectives: A Content Analysis of an Online Health Community

Yingjie Lu¹, PhD; Yang Wu¹, MBA; Jingfang Liu², PhD; Jia Li³, PhD; Pengzhu Zhang⁴, PhD

¹School of Economics and Management, Beijing University of Chemical Technology, Beijing, China

²School of Management, Shanghai University, Shanghai, China

³School of Business, East China University of Science and Technology, Shanghai, China

⁴Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China

Corresponding Author:

Yingjie Lu, PhD

School of Economics and Management

Beijing University of Chemical Technology

North Third Ring Road 15, Chaoyang District

Beijing, 100029

China

Phone: 86 1064448681

Fax: 86 1064448681

Email: lyyj@mail.buct.edu.cn

Abstract

Background: Health care social media used for health information exchange and emotional communication involves different types of users, including patients, caregivers, and health professionals. However, it is difficult to identify different stakeholders because user identification data are lacking due to privacy protection and proprietary interests. Therefore, identifying the concerns of different stakeholders and how they use health care social media when confronted with huge amounts of health-related messages posted by users is a critical problem.

Objective: We aimed to develop a new content analysis method using text mining techniques applied in health care social media to (1) identify different health care stakeholders, (2) determine hot topics of concern, and (3) measure sentiment expression by different stakeholders.

Methods: We collected 138,161 messages posted by 39,606 members in lung cancer, diabetes, and breast cancer forums in the online community MedHelp.org over 10 years (January 2007 to October 2016) as experimental data. We used text mining techniques to process text data to identify different stakeholders and determine health-related hot topics, and then analyzed sentiment expression.

Results: We identified 3 significantly different stakeholder groups using expectation maximization clustering (3 performance metrics: Rand=0.802, Jaccard=0.393, Fowlkes-Mallows=0.537; $P<.001$), in which patients (24,429/39,606, 61.68%) and caregivers (12,232/39,606, 30.88%) represented the majority of the population, in contrast to specialists (2945/39,606, 7.43%). We identified 5 significantly different health-related topics: symptom, examination, drug, procedure, and complication (Rand=0.783, Jaccard=0.369, Fowlkes-Mallows=0.495; $P<.001$). Patients were concerned most about symptom topics related to lung cancer (536/1657, 32.34%), drug topics related to diabetes (1883/5904, 31.89%), and examination topics related to breast cancer (8728/23,934, 36.47%). By comparison, caregivers were more concerned about drug topics related to lung cancer (300/2721, 11.03% vs 109/1657, 6.58%), procedure topics related to breast cancer (3952/13,954, 28.32% vs 5822/23,934, 24.33%), and complication topics (4449/25,701, 17.31% vs 4070/31,495, 12.92%). In addition, patients (9040/36,081, 25.05%) were more likely than caregivers (2659/18,470, 14.39%) and specialists (17,943/83,610, 21.46%) to express their emotions. However, patients' sentiment intensity score (2.46) was lower than those of caregivers (4.66) and specialists (5.14). In particular, for caregivers, negative sentiment scores were higher than positive scores (2.56 vs 2.18), with the opposite among specialists (2.62 vs 2.46). Overall, the proportion of negative messages was greater than that of positive messages related to symptom, complication, and examination. The pattern was opposite for drug and procedure topics. A trend analysis showed that patients and caregivers gradually changed their emotional state in a positive direction.

Conclusions: The hot topics of interest and sentiment expression differed significantly among different stakeholders in different disease forums. These findings could help improve social media services to facilitate diverse stakeholder engagement for health information sharing and social interaction more effectively.

(*J Med Internet Res* 2017;19(4):e109) doi:[10.2196/jmir.7087](https://doi.org/10.2196/jmir.7087)

KEYWORDS

health care social media; stakeholder analysis; topic analysis; sentiment analysis

Introduction

People are increasingly turning to the Internet for health information. Some social media services, such as Wikipedia, Facebook, online forums, and message boards, are increasingly being used to obtain health information and share health care experiences [1,2]. A Pew survey showed that 61% of American adults looked online for health information and that 41% of this group had read about someone else's medical experiences posted in social media. More than half thought the information online affected a decision about how to treat an illness or condition [3].

There are some reasons why patients and their caregivers turn to social media for health care information. First, they feel that doctors are too busy to answer their questions [4]; many doctors only tell their patients how to take a treatment but do not take time to fully explain the detailed reasons [5]. Second, health care social media thus allows users to learn more about their health problems through the use of social support and make decisions about treatment options [6], especially for those with chronic diseases [7]. Third, convenience and anonymity [4] are also important reasons why patients and their caregivers turn to social media, where they feel less embarrassed about asking online experts or communicating with other online members about their health conditions [8].

In addition to patients and their caregivers, other health care stakeholders are involved in health care social media, including doctors, physicians, nurses, aides, social workers, spiritual caregivers, counselors, therapists, and volunteers [9-17]. Some medical experts like to communicate with social media users and offer them professional medical knowledge. In addition, some fellow patients are willing to share their experiences with treatment, which are considered valuable for patients with newly diagnosed conditions. These individuals devote their time to helping other patients get answers in their time of need. These information providers are recognized as health specialists and make great contributions to health care social media.

Health care social media is not only used to exchange health information [18-20], but also provides a platform for emotional communication [21-23]. Some emotional expressions, such as venting, sending positive energy, and showing compassion or empathy, can help patients with serious diseases to cope. It is important especially for patients with cancer and chronic illnesses to receive emotional support and encouragement from fellow patients [24]. A study indicated that patients with low-survival-rate diseases have more emotional communication than those with high-survival-rate diseases [25]. Another study found that an estimated 75% to 85% of patients engaged in

health care social media changed their sentiment in a positive direction through online interactions with other users.

Thus, different types of members discuss topics they are concerned about and share their feelings in health care social media. Understanding the concerns of different stakeholders and how they use health care social media is very meaningful work. Such information can enable users, especially newcomers, to obtain a sense of what social media is, help them quickly find the issues they are concerned about, increase their involvement more easily, and obtain the resources they need for health self-education, self-caring, and self-management. For health care social media sites, a better understanding of user online behaviors could help webmasters provide humanized supporting functions to facilitate diverse stakeholder engagement. For researchers in the field of health care, such studies could clearly summarize the behavior characteristics of different stakeholders and provide advice and guidance for health care social media.

However, given the tremendous amounts of health-related messages posted by social media users, it is difficult and time consuming to employ traditional statistical approaches to find valuable information. Moreover, social media users are usually reluctant to offer personal information due to privacy concerns, and user identification data are usually proprietary and hence not easily accessible. To address these problems, text mining techniques have recently been applied to health social media in some studies [26] to shed light on the demographics of health-related social media users by extracting keywords and other key attributes. However, clearly identifying different health care stakeholders is still a problem that needs to be solved. We therefore proposed a novel framework using text mining techniques to comprehensively analyze the content of an online health community from the perspectives of diverse stakeholders. We used a computational social science approach to process the large amount of text data for stakeholder analysis, topic analysis, and sentiment analysis. This approach allowed us to determine the characteristics of different stakeholders involved in health care social media to facilitate an understanding of health care social media use.

Methods

Data Collection

We chose the online health community MedHelp.org (MedHelp International, San Francisco, CA, USA) as our data source. This site is one of the most popular health care social media platforms. It consists of more than 170 discussion boards about different diseases. Since the site opened in 1994, millions of threads have been posted in the community. It also attracts over one million visitors every month.

We selected representative disease forums. Lung cancer and breast cancer are common cancers with high mortality, and some studies have shown that both cancers are among the most common cancers that Internet users seek information about [27]. Diabetes, as a chronic disease, is also among the most frequently discussed diseases in health care social media. Therefore, we chose lung cancer, breast cancer, and diabetes as our research subjects. We obtained all the webpages from these 3 disease

forums by using webcrawler software Offline Explorer 6.8 (MetaProducts Systems), and then we parsed the pages to extract available messages and stored the messages in a database. Next, we filtered out noisy and unreliable data by text preprocessing. We finally collected 138,161 messages posted by 39,606 members in these 3 disease forums in the 10 years from January 2007 to October 2016 as our experimental data. Table 1 shows the statistical results.

Table 1. Data collection statistics (January 2007 to October 2016) from 3 disease forums on MedHelp.org.

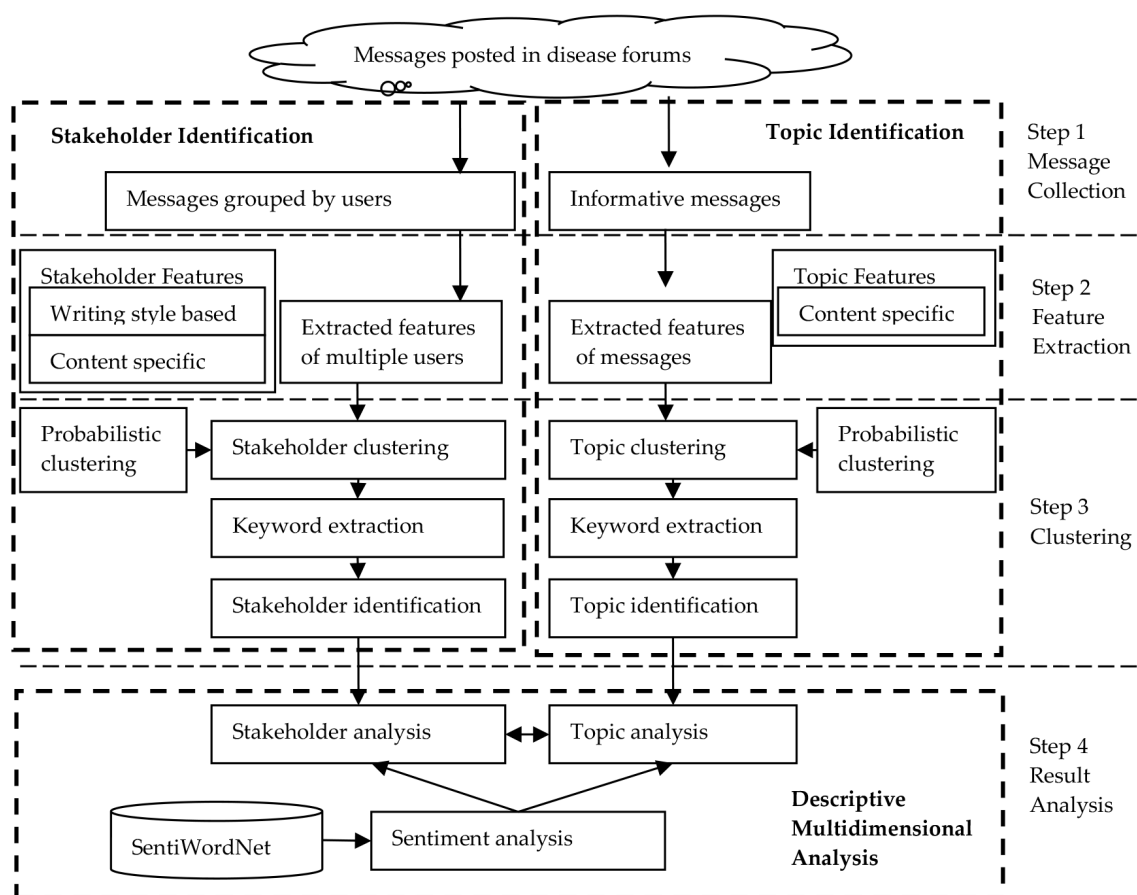
Disease type	No. of messages	No. of members	Messages per member
Lung cancer	5317	2416	2.20
Diabetes	35,193	11,571	3.04
Breast cancer	97,651	25,619	3.81

We only used messages that were open to the public. We never used any user identification data. Personal information such as name, age, and other demographics was not used or reported as part of the results of the study. Therefore, this study did not raise any ethical or legal concerns.

Experiment Design

In this paper, we propose a novel framework using text mining techniques to comprehensively analyze the content of the messages posted in 3 disease forums. The framework consists of 3 parts: stakeholder identification, topic identification, and sentiment analysis, as Figure 1 shows.

First, to distinguish different stakeholders engaged in the online health community, we based stakeholder analysis on text mining techniques to identify different user groups. Users were automatically partitioned into different groups based on the similarity of their posts. An unsupervised approach such as a clustering technique was applied to stakeholder identification. Then, to better understand the concerns of different stakeholders, we based topic analysis on text mining techniques to identify health-related hot topics. Next, we analyzed sentiment to assess the valence and intensity of the messages posted by different stakeholders, as the emotional support and encouragement offered by community members are also an important component of online communication.

Figure 1. Research framework.

Stakeholder Identification

Stakeholder identification included the following 4 steps: (1) message collection, (2) feature extraction, (3) probabilistic clustering, and (4) keyword extraction and stakeholder identification.

For message collection, we first collected all messages posted in the 3 disease forums, and then merged all the messages posted by the same user into one.

For feature extraction, we used a comprehensive textual feature representation encompassing writing style-based features and content-specific features for authorship analysis [28,29]. Writing style-based features were the number of messages, average number of sentences per message, and frequency of words per message. Content-specific features were word *n*-grams, medical domain-specific terminologies, and kinship terminologies. The *n*-grams referred to word unigrams, bigrams, and trigrams and represented important keywords and phrases in the messages. Medical domain-specific terminologies were helpful in distinguishing diverse stakeholders, as some studies have shown that patients and their caregivers used more lay concepts and fewer professional terminologies than health professionals do [30-33]. We selected Medical Subject Headings (MeSH), a controlled thesaurus, to extract health-related terminologies. Family members of patients serving as primary caregivers

usually used kinship terms to refer to the patients they cared for in their posts, and therefore we selected kinship terminologies as part of the content-specific features to distinguish caregivers from other stakeholders. Kinship terminologies in this study were mainly identified by one Unified Medical Language System (UMLS) semantic type, family group, which has been used in previous studies to identify family member concepts [34].

For probabilistic clustering, without a priori knowledge of the number of user groups and their specific characteristics, we chose a probabilistic clustering that leveraged the expectation maximization algorithm among various clustering techniques for stakeholder identification [35]. We found the optimal number of clusters by cross-validating different numbers of clusters. We then used 3 metrics to evaluate the clustering results: the Rand index, Jaccard similarity coefficient, and the Fowlkes-Mallows (FM) index, which we used in our previous study [36]. The clustering result was a probabilistic distribution of instances of belonging to each stakeholder cluster. We labeled each user with a group number according to the cluster with the highest probability assignment.

For keyword extraction, the keywords were terms that best distinguished the corresponding user group from other groups. We extracted these keywords from the messages posted by each user group according to their scores. Supposing all resultant

clusters were C_1, C_2, \dots, C_N , then for each n -gram (unigram, bigram, and trigram) term w in a cluster C_i , its score was calculated as $f(w, C_i) \times \log(N / |\{C_j | f(w, C_j) \geq f(w, C_i), j=1, 2, \dots, N\}|)$ (equation 1), where $f(w, C_i)$ is the frequency of w in cluster C_i , and $|\{C_j | f(w, C_j) \geq f(w, C_i), j=1, 2, \dots, N\}|$ is the total number of clusters with a frequency of term w greater than or equal to the term w frequency of the cluster evaluated. We ranked keywords with high scores, from which we could infer the stakeholder identification.

Topic Identification

We identified topics in a manner similar to identifying stakeholders, by the following steps: message collection, feature extraction, probabilistic clustering, keyword extraction, and topic identification.

For message collection, we only chose informational messages involving health conditions, online reviews of particular drugs or medical treatments, and so on. We filtered out messages containing no informational support but only offering emotional support and other spam messages.

For feature extraction, the features used in topic identification included 2 parts: word n -grams and domain-specific terminologies. Domain-specific terminologies were from the UMLS Metathesaurus. The UMLS Metathesaurus as the world's largest repository of biomedical concepts, consisting of 1.7 million biomedical concepts, where each concept is assigned to at least one of the 134 semantic types. We only chose the terminologies mapped to health-related semantic types that we used in our previous study [36]. To obtain these medical terminologies automatically, we used MetaMap 2014 (US National Library of Medicine), a highly configurable program that maps biomedical text to concepts in the UMLS Metathesaurus. Using a Java application programming interface offered by MetaMap, we could parse the messages to get the health-related terminologies.

We performed clustering and keyword extraction for topic identification in the same manner as stakeholder identification.

Sentiment Analysis

We analyzed sentiment to assess the valence and intensity of the messages by community members. Lexicon-based approaches are widely used in sentiment analysis, and some well-known sentiment lexicons, such as SentiWordNet [37], have been successfully applied in sentiment analysis [38,39]. The SentiWordNet lexicon provides positive- and negative-intensity scores for each sentiment term. In this study, we used the SentiWordNet lexicon to extract sentiment terms from the messages and calculate their sentiment scores. We used the following 3 sentiment measures to evaluate sentiment expression: PositiveScores, NegativeScores, and SubjectiveScores [40,41], where PositiveScores is positive-polarity scores divided by the number of messages; NegativeScores is negative-polarity scores divided by the number of messages; and SubjectiveScores is subjective-polarity scores divided by the number of messages.

Results

Stakeholder Analysis

We used the expectation maximization clustering algorithm for stakeholder identification. Expectation maximization clustering can evaluate and determine the optimal number of clusters by cross-validating different numbers of clusters. We identified 3 stakeholder groups in 3 disease forums. We then performed the 2-sample t test to evaluate whether there was a significant difference in the 3 stakeholder groups. The null hypothesis was that the difference between 2 groups was 0. The conclusion was to reject the null hypothesis ($P < .001$) and that the 3 stakeholder groups were significantly different. Then, we extracted some keywords from each group according to equation (1) after filtering out meaningless or invalid phrases. Table 2 shows the results.

Table 2. Stakeholder analysis in the 3 disease forums.

Cluster	Keywords	Authorship
Lung cancer		
1	in my chest, on my chest, my lungs, my left lung, my right lung, please help me, of my chest, I was diagnosed, my ct scan, doctor tell me, my xray result, my question, i was wondering, I have cancer, be greatly appreciated	Patients
2	my husband, thank you, my mother, my dad, my mom, my sister, my question is, my father, do you think, i am worried, on his lung, thanks in advance, thanks so much, father in law	Caregivers
3	good luck, all the best, hope this helps, stay positive, you should, god bless, your mother, your father, your husband, with your doctor, you need to, let us know, sorry to hear, you could consider, see your doctor, your symptoms, with your physician, best of luck	Specialists
Diabetes		
1	my sugar, help me, I was diagnosed, my sugar levels, my blood, I was wondering, thank you, my body, my question, thanks for your, my blood sugar, I need to, my question is, I have diabetes, be greatly appreciated, type 1 diabetic	Patients
2	my husband, was diagnosed, he was diagnosed, her blood sugar, my son, my daughter, his blood sugar, low blood sugar, want to know, he has been, she has been, thanks so much, my son is, daughter was diagnosed	Caregivers
3	your blood, you need, your doctor, you need to, good luck, you should, your glucose, your blood sugar, I would suggest, your blood sugars, with your doctor, let us know, hope this helps, I hope you, your glucose levels,	Specialists
Breast cancer		
1	my breast, my nipple, my breasts, thank you, my question, should I, my left breast, in my right, on my left, in my breast, my right breast, found a lump, thanks for your, be greatly appreciated, of my breast	Patients
2	my mom, she had, my mother, her breast, she was diagnosed, family history, had breast cancer, my question is, worry about, her left breast, her right breast, my sister	Caregivers
3	I hope, you can, your doctor, you should, best wishes, your breast, good luck, you need to, your oncologist, let us know, with your doctor, hope this helps, all the best, a second opinion, second opinion	Specialists

Then, we could infer stakeholder identification according to the extracted keywords. Taking clustering results in the lung cancer forums as an example, in the cluster 1 group, members talked about their own conditions (eg, my chest, my lungs, my left lung, my right lung, of my chest, in my chest, on my chest) and procedures they underwent (eg, I was diagnosed, my xray result, my ct scan). They hoped to get help (eg, please help me, my question, i was wondering) and then gave thanks (eg, be greatly appreciated). We therefore assigned cluster 1 as the patient group. In the cluster 2 group, members talked more about their family members (eg, my husband, my mother, my father, my dad, my mom, my sister, father in law) and expressed their concern about their family members (eg, i am worried). They also raised some questions of concern (eg, my question is) and gave thanks (eg, thank you, thanks in advance, thanks so much). We assigned cluster 2 as the caregiver group. In the cluster 3 group, members were more likely to give other members advice and suggestions (eg, you should, you need to, you could consider) and offer help (eg, hope this helps) or advise them to

undergo further consultation and procedures with their doctors (eg, with your doctor, see your doctor, with your physician). Members in the cluster 3 group often expressed their compassion and encouragement to other members (eg, stay positive, good luck, all the best, sorry to hear, best of luck). We thus assigned cluster 3 as the health specialist group.

Then we used the 3 evaluation metrics to test whether the textual features proposed in this study could significantly distinguish different stakeholders. By incorporating writing style-based features (F1), word n -grams (F2), medical domain-specific terminologies (F3), and kinship terminologies (F4) into the feature set in turn, we evaluated the results of clustering based on different feature sets, as Table 3 shows. The results showed that the 3 measure values increased as more features were incorporated into the feature set, and reached a maximum (eg, Rand=0.802, Jaccard=0.393, FM=0.537 for breast cancer) when all features were incorporated, indicating that these textual features could improve the performance of stakeholder identification significantly.

Table 3. Performance measures for distinguishing stakeholders using different textual feature sets^a.

Disease	Feature set	Performance measure		
		Rand index	Jaccard similarity coefficient	Fowlkes-Mallows index
Lung cancer	F1	0.712	0.261	0.395
	F1+F2	0.731	0.321	0.441
	F1+F2+F3	0.757	0.349	0.473
	F1+F2+F3+F4	0.785	0.371	0.501
Diabetes	F1	0.717	0.273	0.401
	F1+F2	0.742	0.335	0.456
	F1+F2+F3	0.780	0.367	0.489
	F1+F2+F3+F4	0.792	0.381	0.523
Breast cancer	F1	0.725	0.297	0.421
	F1+F2	0.779	0.356	0.481
	F1+F2+F3	0.793	0.385	0.529
	F1+F2+F3+F4	0.802	0.393	0.537

^aFeature set components: style-based features (F1), word *n*-grams (F2), medical domain-specific terminologies (F3), and kinship terminologies (F4).

In terms of the distributions of different stakeholders in the 3 disease forums, as [Figure 2](#) shows, patients (24,429/39,606, 61.68%) and caregivers (12,232/39,606, 30.88%) were the majority of the population, in contrast to health specialists (2945/39,606, 7.43%). The proportions of patients (1202/2416, 49.75% vs 5738/11,571, 49.59%) and caregivers (1053/2416, 43.58% vs 4836/11,571, 41.79%) were similar in the lung cancer and diabetes forums. By contrast, in the breast cancer forum, the proportion of patients (17,489/25,619, 68.27%) was significantly greater.

The reason for the different distributions may be that women are the prominent group with breast cancer and are also more likely than men to seek online information, as indicated by a previous study [42].

[Table 4](#) shows descriptive statistics of postings by different stakeholders. Both patients and caregivers published fewer than 2 messages (1.48 and 1.51) on average. By contrast, health specialists, although constituting a minority of the population in the online community compared with patients and caregivers, published many more postings (28.4) on average than the other 2 groups.

We speculated about the reason for different distributions of posting. Generally, most patients and caregivers aimed to seek health information they were concerned about. They would leave the site after receiving satisfactory answers. Another possibility was that they simply became frustrated with the inability to find answers and left the site soon. Most patients and caregivers thus were short-term participants.

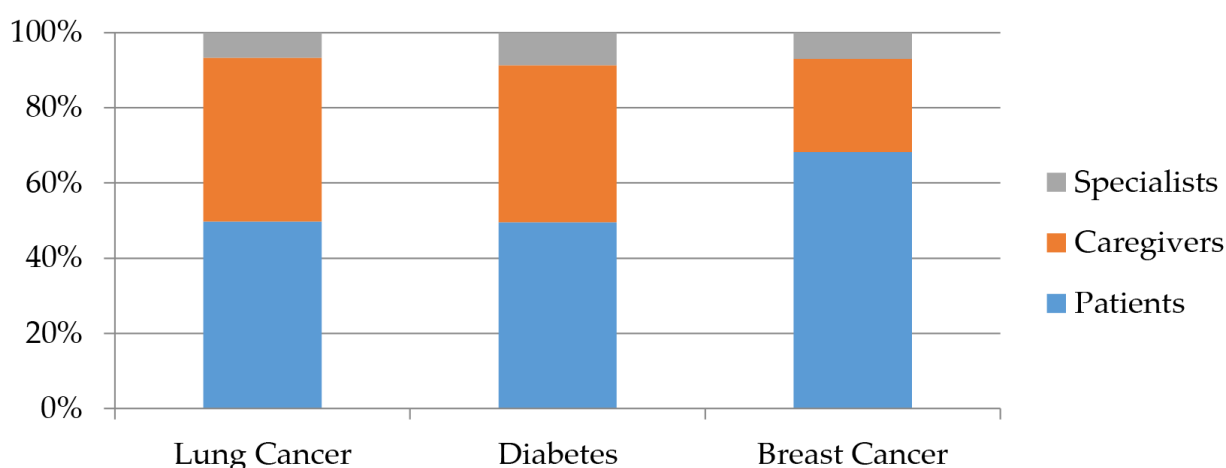
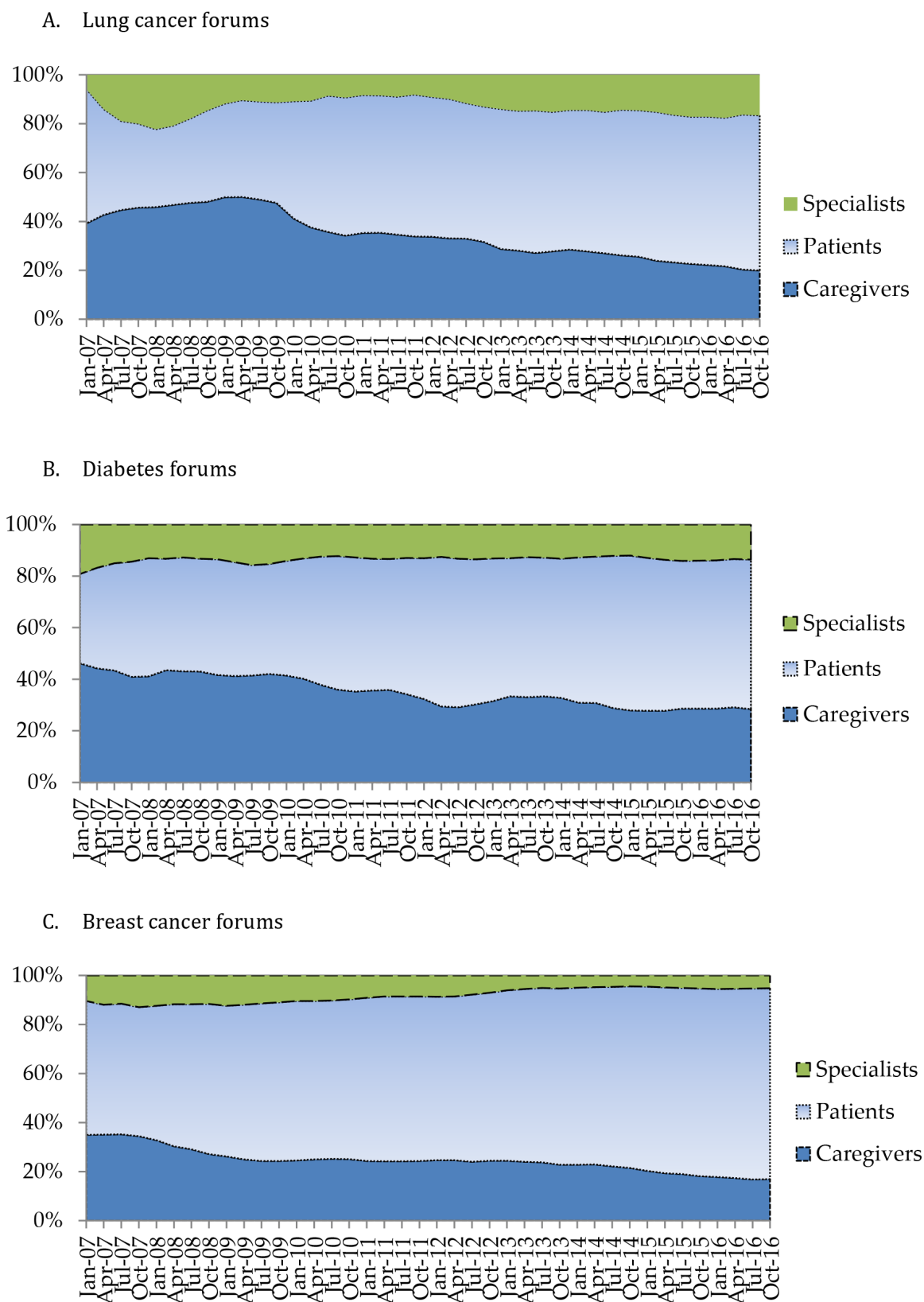
Figure 2. Distributions of the 3 stakeholder groups.

Table 4. Clustering results in the stakeholder analysis.

Disease	Cluster	No. of members	No. of messages	Messages per member	Authorship
Lung cancer	1	1202	1378	1.15	Patients
	2	1053	1607	1.53	Caregivers
	3	161	2332	14.48	Specialists
Diabetes	1	5738	7691	1.34	Patients
	2	4836	7136	1.48	Caregivers
	3	997	20,366	20.43	Specialists
Breast cancer	1	17,489	27,012	1.55	Patients
	2	6343	9727	1.53	Caregivers
	3	1787	60,912	34.09	Specialists

Figure 3. Changes in the proportions of stakeholder groups in (A) lung cancer, (B) diabetes, and (C) breast cancer forums, January 2007 to October 2016.



In terms of changes in the proportions of the 3 stakeholder groups (see Figure 3), the proportion of the specialist group tended to decrease and then stabilize. By contrast, the proportion of caregivers decreased gradually, and the proportion of patients increased.

One possible explanation is that most health specialists gradually formed a stable relationship with the online community and only a few left after short-term engagement. By contrast, as health care social media becomes more widespread and widely accepted, a higher proportion of patients than of caregivers

prefer to seek online health information. After all, patients themselves are most concerned about their own health conditions.

Topic Analysis

Clustering and keyword extraction for topic identification were performed in the same manner as stakeholder identification.

We identified 5 significantly different health hot topics ($P<.001$), as Table 5 shows. They were named based on the extracted keywords with similar UMLS semantic types: symptom, examination, procedure, drug, and complication.

Table 5. Topic analysis in the 3 disease forums.

Cluster	Topics	Keywords	UMLS ^a semantic types
Lung cancer			
1	Symptom	cough, pain, breathless, symptoms, chest pain, painful, shortness of breath, wheezing, short of breath, coughing up blood, nausea	sosy
2	Complication	infection, bronchitis, pneumonia, tuberculosis, asthma, pleural effusion, copd, emphysema, collapsed lung, atelectasis	dsyn, patf
3	Examination	scans, x-ray, cat scan, mri, biopsy, pet scan, chest x-ray, imaging, biopsy needle, bronchoscopy	diap
4	Procedure	chemo, operation, surgery, radiation, therapy, chemotherapy, removal, radiation therapy, wedge resection, lobectomy	topp
5	Drug	silicas, morphine, advil, tarceva, chantix, carboplatin, alimta, dilaudid, taxol, coumadin	phsu
Diabetes			
1	Drug	insulin, lantus, januvia, metformin, glucophage, actos, avandia, amaryl, marijuana, glipizide	phsu
2	Complication	infection, hypoglycemia, low blood sugar, dka, obesity, pcos, kidney disease, coma, diabetic neuropathy, bgs	dsyn, patf
3	Symptom	pain, tired, thirsty, nausea, fatigue, frequent urination, hungry, dizzy, itchy, sore, tingling	sosy
4	Examination	blood test, glucose test, fasting test, fasting blood sugar, cat scan, hemoglobin a1c test, gtts, glucose tolerance test, mri	lbpr, diap
5	Procedure	infusion, therapy, injection, transplant, dialysis, rx, ect, insulin injection, cde, amputation	topp
Breast cancer			
1	Examination	biopsy, mri, ultrasound, mammogram, screening, bi-rads, core biopsy, cat scan, imaging, biopsy needle	diap, lbpr
2	Procedure	chemo, operation, chemotherapy, radiation, radiotherapy, mastectomy, lumpectomy, implant, removal, surgical	topp
3	Symptom	sore, pain, painful, breast pain, nipple discharge, itching, tingling, hot flashes, nausea, itchy	sosy
4	Drug	tamoxifen, arimidex, taxol, femara, taxotere, carboplatin, effexor, docetaxel, valium, raloxifene	phsu
5	Complication	infection, rash, lymph edema, fibrocystic breast, mastitis, idc, eczema, complex cyst, complex cysts, neuropathy, fibrocystic breast disease, fibrocystic disease	dsyn

^aUMLS: Unified Medical Language System.

We then used the 3 evaluation metrics to test whether the textual features proposed in this study could be used to distinguish different hot topics significantly. By incorporating word n -grams (F1) and medical domain-specific terminologies (F2) into the feature set, we evaluated the results of clustering based on a

2-feature set, as s Table 6 shows. The results showed that feature set F1+F2 outperformed F1 significantly (eg, Rand=0.783, Jaccard=0.369, FM=0.495 for breast cancer), indicating that the 2 types of textual features improved the performance of topic identification significantly.

Table 6. Performance measures for distinguishing hot topics using different textual feature sets^a.

Disease	Feature set	Rand	Jaccard	FM ^b
Lung cancer	F1	0.703	0.242	0.382
	F1+F2	0.761	0.352	0.478
Diabetes	F1	0.718	0.275	0.411
	F1+F2	0.774	0.351	0.478
Breast cancer	F1	0.722	0.285	0.417
	F1+F2	0.783	0.369	0.495

^aFeature set components: word *n*-grams (F1) and medical domain-specific terminologies (F2).

^bFM: Fowlkes-Mallows index.

To further explore the concerns of different stakeholders, we examined the distributions of the 3 stakeholder groups in the 5 hot topics (Figure 4). Among the different stakeholders in the different types of diseases, their concerns about various health topics were fairly different.

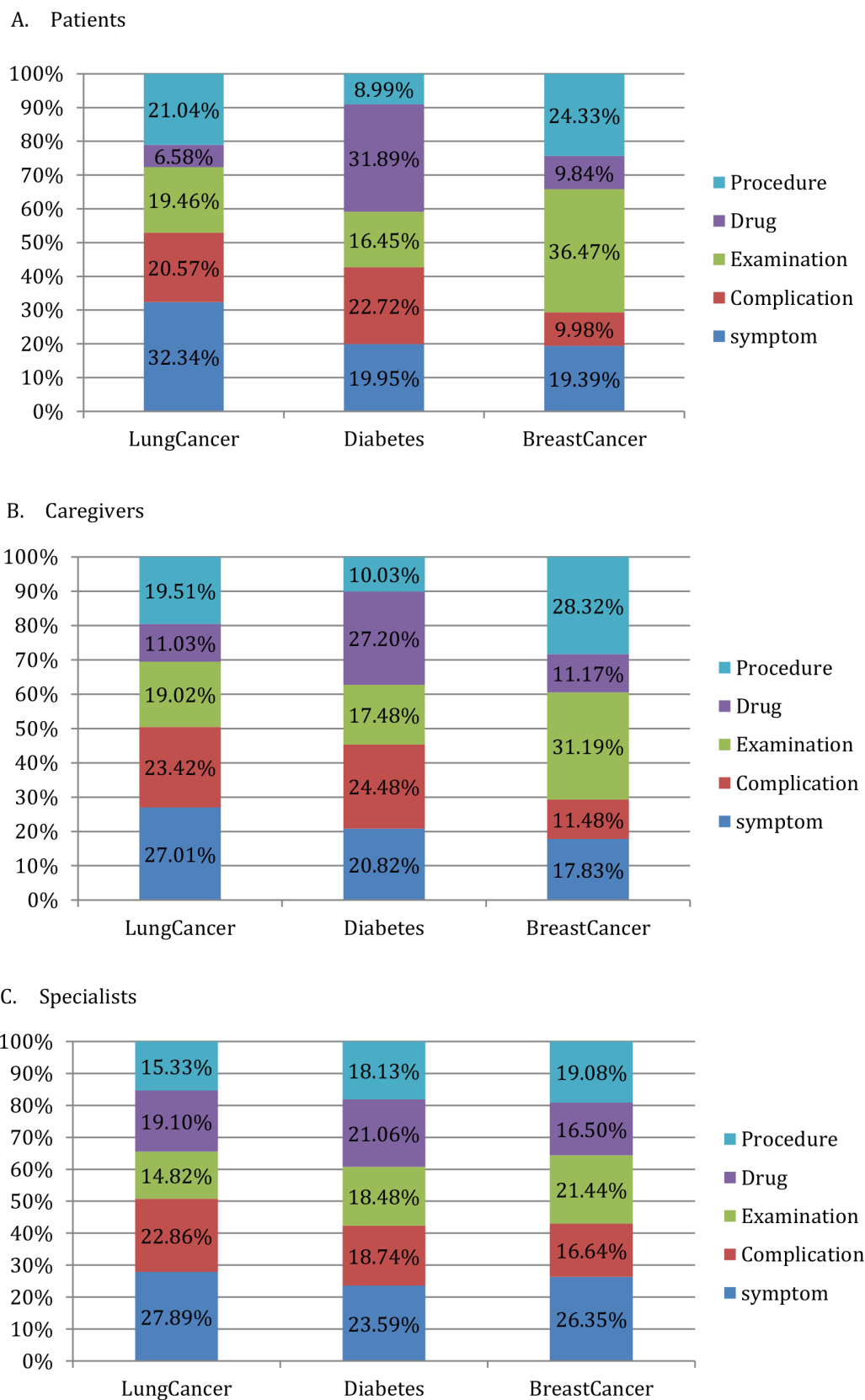
A significantly greater proportion of patients with lung cancer (536/1657, 32.34%) were involved in the symptom topics. Patients with diabetes were more interested in drug topics (1883/5904, 31.89%). Patients with breast cancer were more likely to mention their examination or tests online (8728/23,934, 36.47%).

The reasons may be that some early symptoms of lung cancer, such as coughing or wheezing, are very similar to ailments such as fevers and bronchitis. Patients with such ailments were not sure whether their conditions were signs of lung cancer and thus hoped to obtain valuable information by seeking the experiences of others or consulting about their conditions on health care social media. By contrast, patients who found an abnormality of their breast tissues usually underwent further examination. Because of a lack of other obvious symptoms, they mainly discussed their examination or tests for the diagnosis of breast cancer. Diabetes is a common chronic disease that is difficult to cure. Patients with diabetes use regular long-term medication and thus were naturally concerned about topics related to antidiabetic drugs to learn about their medical effects and possible side effects.

The distribution of hot topics among caregivers was similar to that of patients. Both groups appeared to be mainly concerned with the hottest health issues related to a specific disease, as patients and caregivers were both information requesters. However, the proportional distribution of topics was distinguishable between patients and caregivers. For example, a significantly greater proportion of caregivers than of patients were involved in drug topics (300/2721, 11.03% for caregivers vs 109/1657, 6.58% for patients) in lung cancer forums and in procedure topics (3952/13,954, 28.32% for caregivers vs 5822/23,934, 24.33% for patients) in breast cancer forums.

These results indicate that caregivers were more concerned about disease treatment topics than disease diagnosis because drug and procedure topics are both related to disease treatment. Complication was another health topic of interest to caregivers, possibly because it may greatly affect their caregiving responsibilities. As Figure 4 shows, the proportion of complication topics was significantly greater among caregivers than among patients in the 3 disease forums (4449/25,701, 17.31% for caregivers vs 4070/31,495, 12.92% for patients).

As Figure 4 shows, health specialists were equally distributed among the 5 health topics, in contrast to patients and caregivers. This result indicates that specialists, as information providers, were more likely to focus on various health topics and to share their experiences and knowledge. Broad knowledge is indispensable for health specialists to offer help to patients in their time of need.

Figure 4. Distributions of (A) patients, (B) caregivers, and (C) specialists in the 5 hot topics.

Sentiment Analysis

Emotional messages could be distinguished from informative messages by setting a threshold of sentiment measures in SubjectiveScores. As Figure 5 shows, informative messages

(108,519/138,161, 78.54%) were more prevalent than emotional messages (29,642/138,161, 21.45%).

This result suggests that most users make use of health care social media more to exchange health information than to provide a platform for emotional communication. This result

is also in agreement with the above argument that most patients and caregivers were short-term participants and that many only sought health information of interest.

The distribution of informative and emotional messages among different stakeholders indicated that patients were more likely to express their emotions than caregivers and specialists. The proportion of emotional messages was 25.05% (9040/36,081)

for patients, 14.39% (2659/18,470) for caregivers, and 21.46% (17,943/83,610) for specialists. Particularly among patients in the diabetes forums, patients usually encouraged each other to overcome their diseases by pursuing long-term treatment. By contrast, caregivers preferred to share informative messages and were the least likely to exchange emotions among the 3 stakeholders.

Figure 5. Distributions of informative messages and emotional messages by (A) disease and by stakeholder group in (B) lung cancer, (C) diabetes, and (D) breast cancer forums.

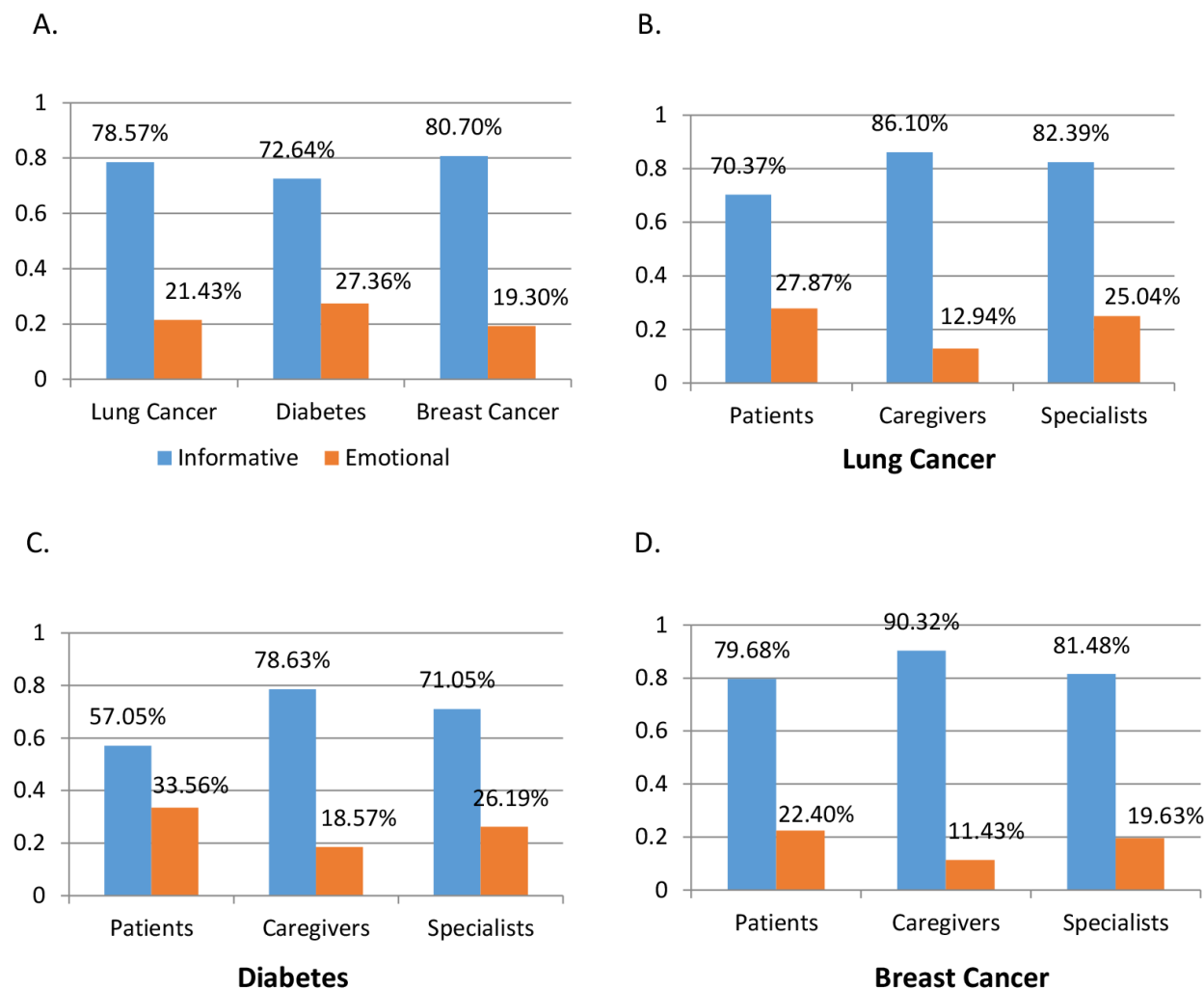
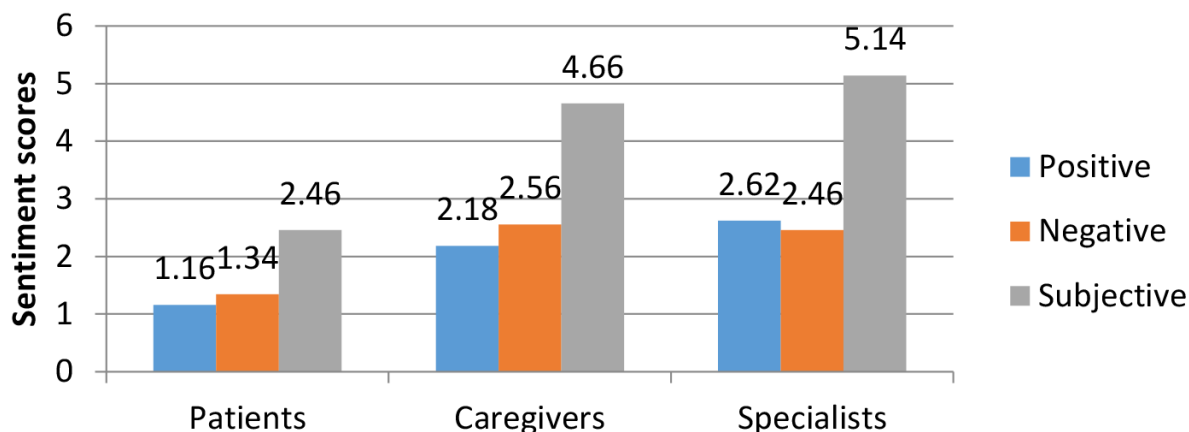


Figure 6. Sentiment measures of the 3 stakeholder groups.



One possibility is that the caregiving responsibilities of caregivers do not leave sufficient time for online communication. Especially for the caregivers involved in the 2 cancer forums, most only sought diagnosis and treatment information of interest online to help their patients cure their high-survival-rate diseases.

We calculated the 3 sentiment measures of PositiveScores, NegativeScores, and SubjectiveScores to evaluate sentiment expression, as Figure 6 shows. The subjective sentiment scores of patients (2.46) were lower than those of caregivers (4.66) and specialists (5.14), indicating that, although patients were more likely to express their emotions, as Figure 5 shows, the emotional intensity of patients was lower than that of specialists and caregivers. The negative sentiment scores of caregivers (2.56) were higher than their positive scores (2.18). Among health specialists, positive sentiment scores (2.62) were higher than their negative sentiment scores (2.46).

We speculated about the reasons for the differences in sentiment expressions. Patients appear to prefer to regard health care social media as an emotion-exchange platform rather than a place to vent their negative feelings. Conversely, caregivers undertaking long-term caretaking activities usually expressed strong emotions about their patients. They preferred to use negative words with high emotional intensity to vent their emotional stress and burden [43,44], as Figure 6 shows that the negative sentiment scores of caregivers were higher than their positive scores. Many caregivers appeared to regard health care social media as a good platform to vent and relieve stress. For health specialists, higher positive sentiment scores implied that this group provided emotional support by encouraging patients and their caregivers to confront their diseases rather than expressing sympathy and pity.

Further analysis identified differences in sentiment expression by different stakeholders in different hot topics, as Figure 7 shows. Overall, for patients and caregivers, the proportions of negative emotional messages were greater than those of positive emotional messages among the topics related to symptom, complication, and examination; for example, for patients, the proportion was 30.83% (562/1823) positive versus 69.17% (1261/1823) negative in symptom topics, 36.43% (1055/2896) versus 63.57% (1841/2896) in examination topics, and 37.85% (450/1189) versus 62.15% (739/1189) in complication topics.

This result indicates that patients and caregivers who discussed complications and examination topics feared illness, depression, hopelessness, anxiety, and other negative emotions and thus were more likely to express negative emotions. In particular, among caregivers, negative emotional intensity was significantly higher for the complication topic than for the other topics, indicating that complications were a considerable burden for caregivers performing long-term caretaking activities, and this group expressed strong negative feelings of disappointment and even despair. By contrast, drug and procedure topics are usually more closely related to the treatment of disease, and thus all stakeholders who discussed these 2 topics were more likely to

have a positive attitude; for example, for patients, the proportion was 47.94% (580/1210) positive versus 52.06% (630/1210) negative in drug topics and 59.11% (1136/1922) versus 40.89% (786/1922) in procedure topics. In their discussion of the 2 topics, patients were happy to be getting better, and the caregivers expressed their satisfaction with the treatment; in addition, specialists conveyed their congratulations and best wishes to the patients.

Finally, we analyzed trends in emotional changes among the different stakeholders to illustrate whether members who engaged in health care social media changed their sentiment expression based on online interactions with other users. We incorporated the messages posted by different members every half month and calculated the changes in the 3 sentiment measures of PositiveScores, NegativeScores, and SubjectiveScores over time in the first year of their involvement. Figure 8 shows the results.

As Figure 8 shows, the subjective sentiment scores of the messages from patients tended to increase in the early weeks and then stabilized. Patients preferred to express concern about their health topics of interest during their early involvement and then gradually began to express their feelings for social support. From the perspective of sentiment polarity, we found that positive sentiment scores increased gradually while negative sentiment scores decreased, implying that the patients engaged in health care social media changed their emotional state in a positive direction through online interactions with other users. Some possible reasons for this change are that the patients eliminated fears and anxiety after their problems were resolved or that professionals and fellow patients encouraged them to confront their diseases.

The subjective sentiment scores of the messages from caregivers remained high. Caregivers apparently preferred to use subjective words with high emotional intensity in their posts compared with patients. From the perspective of sentiment polarity, positive sentiment scores increased gradually while negative sentiment scores decreased, similar to the trend changes in patients, possibly for the same reasons discussed above. However, in contrast to the sentiment expression of patients, for caregivers the negative sentiment scores remained higher than the positive sentiment scores, implying that caregivers appeared to regard health care social media as a platform to share their negative feelings. Caregivers may complain about their long-term, heavy caretaking work online to relieve their stress at any time.

All 3 sentiment scores of the messages from specialists remained stable over time. This result indicates that specialists were involved long term in the online community to provide emotional support and to help patients and caregivers in various ways. They expressed their sympathy and pity for the unfortunate users by using some negative words while encouraging them to fight their illness using some positive words.

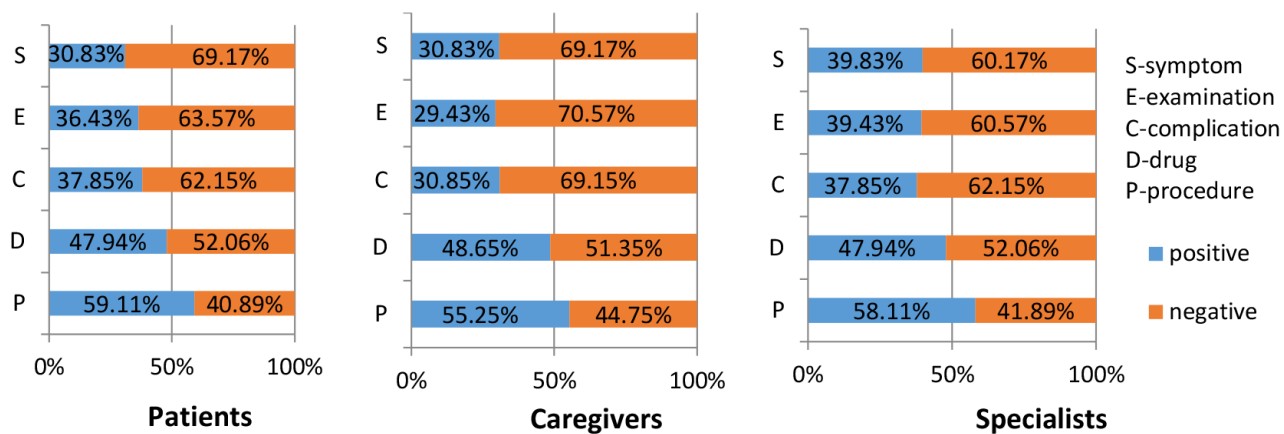
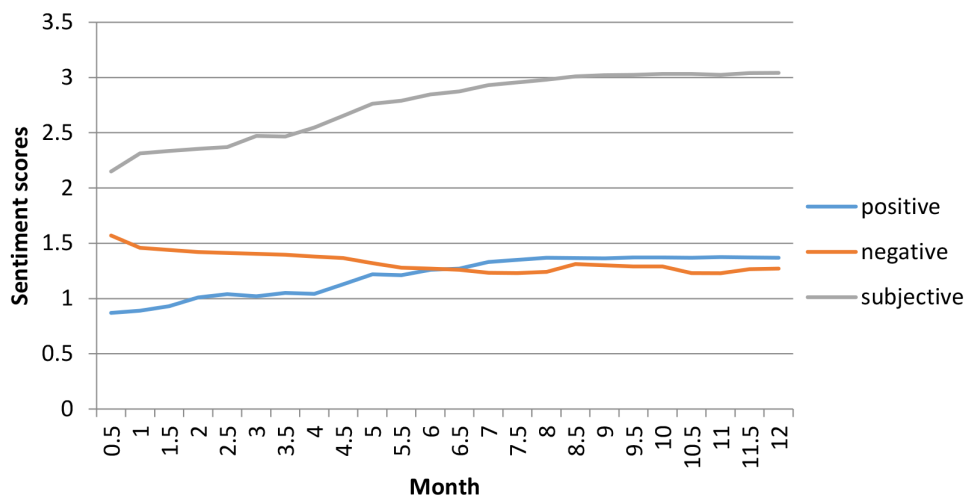
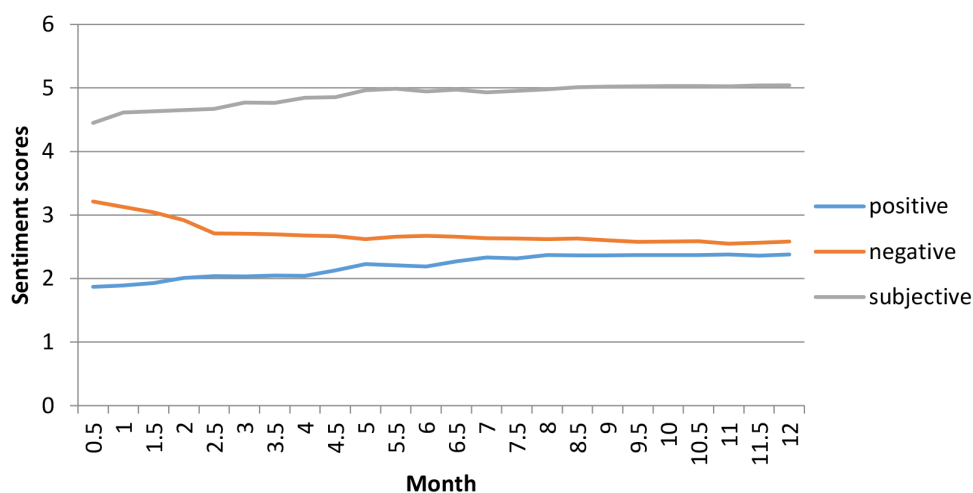
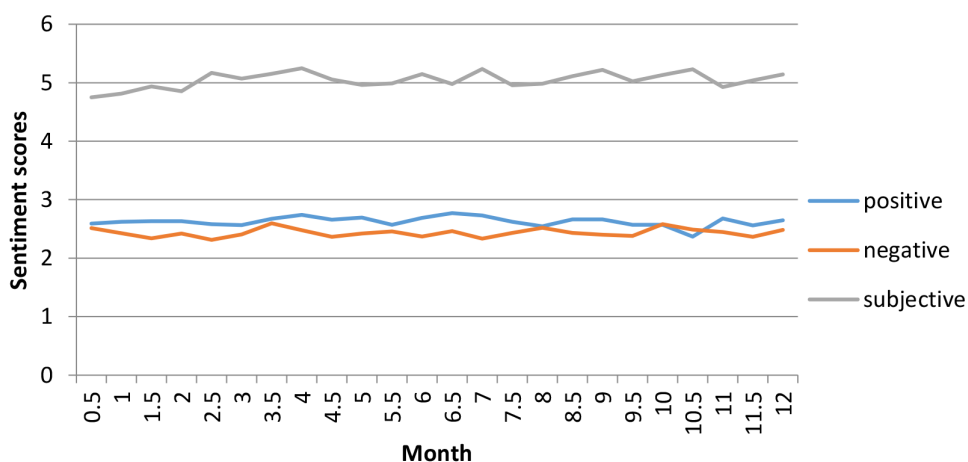
Figure 7. Distributions of positive and negative messages posted by the 3 stakeholder groups in the 5 hot topics.

Figure 8. Changes in the sentiment expression of (A) patients, (B) caregivers, and (C) specialists over the first year of each user's involvement.**A. Patients****B. Caregivers****C. Specialists**

Discussion

Understanding the concerns of different stakeholders and how they use social media is very meaningful work. In this paper, we proposed a novel framework using text mining techniques

to perform a comprehensive content analysis of an online health community from the perspectives of diverse stakeholders. We used a computational social science approach to process the large amount of text data for stakeholder analysis, topic analysis, and sentiment analysis. We identified significant differences in

hot topics of interest and sentiment expression among different stakeholders involved in different types of disease forums. These valuable conclusions provide a better understanding of health care social media use by different stakeholders that may aid improvements in social media services to facilitate diverse stakeholder engagement for more effective health information sharing and social interaction.

This study also has some limitations that must be considered further. Further research should examine how to describe and

measure the impacts of health care social media use on health self-management. Users involved in social media maintained good communication and developed online social networks, and thus future studies should include social network analysis of different stakeholder groups to determine the impacts of these relationships on their engagement. In addition, some deeper stakeholder analysis by using text mining techniques, such as how to distinguish knowledgeable patients and professional doctors in specialist groups, would be considered in our further study.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 71401005 and no. 71371005) and the Major Research Plan of the National Natural Science Foundation of China (grant no. 91646205).

Conflicts of Interest

None declared.

References

1. Lasker JN, Sogolow ED, Sharim RR. The role of an online community for people with a rare disease: content analysis of messages posted on a primary biliary cirrhosis mailinglist. *J Med Internet Res* 2005 Mar 31;7(1):e10 [FREE Full text] [doi: [10.2196/jmir.7.1.e10](https://doi.org/10.2196/jmir.7.1.e10)] [Medline: [15829472](#)]
2. Adams S, de Bont A, Berg M. Looking for answers, constructing reliability: an exploration into how Dutch patients check web-based medical information. *Int J Med Inform* 2006 Jan;75(1):66-72. [doi: [10.1016/j.ijmedinf.2005.07.036](https://doi.org/10.1016/j.ijmedinf.2005.07.036)] [Medline: [16137921](#)]
3. Fox S, Jones S. The social life of health information. Washington, DC: Pew Internet & American Life Project; 2009. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2009/PIP_Health_2009.pdf [accessed 2017-03-29] [WebCite Cache ID 6pKPvrr9]
4. Umefjord G, Petersson G, Hamberg K. Reasons for consulting a doctor on the Internet: Web survey of users of an Ask the Doctor service. *J Med Internet Res* 2003 Oct 22;5(4):e26 [FREE Full text] [doi: [10.2196/jmir.5.4.e26](https://doi.org/10.2196/jmir.5.4.e26)] [Medline: [14713654](#)]
5. Dickerson SS, Boehmke M, Ogle C, Brown JK. Seeking and managing hope: patients' experiences using the Internet for cancer care. *Oncol Nurs Forum* 2006 Jan 01;33(1):E8-17. [doi: [10.1188/06.ONF.E8-E17](https://doi.org/10.1188/06.ONF.E8-E17)] [Medline: [16470231](#)]
6. Gerber BS, Eiser AR. The patient physician relationship in the Internet age: future prospects and the research agenda. *J Med Internet Res* 2001;3(2):E15 [FREE Full text] [doi: [10.2196/jmir.3.2.e15](https://doi.org/10.2196/jmir.3.2.e15)] [Medline: [11720957](#)]
7. Bansil P, Keenan NL, Zlot AI, Gilliland JC. Health-related information on the Web: results from the HealthStyles Survey, 2002-2003. *Prev Chronic Dis* 2006 Apr;3(2):A36 [FREE Full text] [Medline: [16539777](#)]
8. 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](#)]
9. Culver JD, Gerr F, Frumkin H. Medical information on the Internet: a study of an electronic bulletin board. *J Gen Intern Med* 1997 Aug;12(8):466-470 [FREE Full text] [Medline: [9276651](#)]
10. Varlamis I, Apostolakis I. Medical informatics in the Web 2.0 era. In: Tsihrintzis GA, Virvou M, Howlett RJ, Jain LC, editors. *New Directions in Intelligent Interactive Multimedia*. Berlin, Germany: Springer; 2008:513-522.
11. Kernisan LP, Sudore RL, Knight SJ. Information-seeking at a caregiving website: a qualitative analysis. *J Med Internet Res* 2010;12(3):e31 [FREE Full text] [doi: [10.2196/jmir.1548](https://doi.org/10.2196/jmir.1548)] [Medline: [20675292](#)]
12. Ginossar T. Online participation: a content analysis of differences in utilization of two online cancer communities by men and women, patients and family members. *Health Commun* 2008;23(1):1-12. [doi: [10.1080/10410230701697100](https://doi.org/10.1080/10410230701697100)] [Medline: [18443988](#)]
13. Colineau N, Paris C. Talking about your health to strangers: understanding the use of online social networks by patients. *New Rev Hypermedia Multimedia* 2010 Apr;16(1-2):141-160. [doi: [10.1080/13614568.2010.496131](https://doi.org/10.1080/13614568.2010.496131)]
14. Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. *Int J Med Inform* 2009 Oct;78(10):645-655. [doi: [10.1016/j.ijmedinf.2009.04.008](https://doi.org/10.1016/j.ijmedinf.2009.04.008)] [Medline: [19501017](#)]
15. Domingo MC. Managing healthcare through social networks. *Computer* 2010 Jul;43(7):20-25. [doi: [10.1109/MC.2010.92](https://doi.org/10.1109/MC.2010.92)]
16. Durant KT, McCray AT, Safran C. Identifying temporal changes and topics that promote growth within online communities: a prospective study of six online cancer forums. *Int J Comput Models Algorithms Med* 2011 Apr;2(2):1-22 [FREE Full text] [doi: [10.4018/jcmam.2011040101](https://doi.org/10.4018/jcmam.2011040101)] [Medline: [22022636](#)]
17. Cho J, Noh H, Ha MH, Kang SN, Choi J, Chang YJ. What kind of cancer information do Internet users need? *Support Care Cancer* 2011 Sep;19(9):1465-1469. [doi: [10.1007/s00520-010-1057-9](https://doi.org/10.1007/s00520-010-1057-9)] [Medline: [21717273](#)]

18. Macias W, Lewis LS, Smith TL. Health-related message boards/chat rooms on the Web: discussion content and implications for pharmaceutical sponsorships. *J Health Commun* 2005 Apr;10(3):209-223. [doi: [10.1080/10810730590934235](https://doi.org/10.1080/10810730590934235)] [Medline: [16036729](https://pubmed.ncbi.nlm.nih.gov/16036729/)]
19. Schultz PN, Stava C, Beck ML, Vassilopoulou-Sellin R. Internet message board use by patients with cancer and their families. *Clin J Oncol Nurs* 2003;7(6):663-667. [doi: [10.1188/03.CJON.663-667](https://doi.org/10.1188/03.CJON.663-667)] [Medline: [14705483](https://pubmed.ncbi.nlm.nih.gov/14705483/)]
20. Attard A, Coulson NS. A thematic analysis of patient communication in Parkinson's disease online support group discussion forums. *Comput Hum Behav* 2012 Mar;28(2):500-506. [doi: [10.1016/j.chb.2011.10.022](https://doi.org/10.1016/j.chb.2011.10.022)]
21. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *J Med Internet Res* 2013;15(11):e239 [FREE Full text] [doi: [10.2196/jmir.2721](https://doi.org/10.2196/jmir.2721)] [Medline: [24184993](https://pubmed.ncbi.nlm.nih.gov/24184993/)]
22. Mohd Roffeei SH, Abdullah N, Basar SKR. Seeking social support on Facebook for children with autism spectrum disorders (ASDs). *Int J Med Inform* 2015 May;84(5):375-385. [doi: [10.1016/j.ijmedinf.2015.01.015](https://doi.org/10.1016/j.ijmedinf.2015.01.015)] [Medline: [25701266](https://pubmed.ncbi.nlm.nih.gov/25701266/)]
23. Jin J, Yan X, Li Y, Li Y. How users adopt healthcare information: an empirical study of an online Q&A community. *Int J Med Inform* 2016 Feb;86:91-103. [doi: [10.1016/j.ijmedinf.2015.11.002](https://doi.org/10.1016/j.ijmedinf.2015.11.002)] [Medline: [26616406](https://pubmed.ncbi.nlm.nih.gov/26616406/)]
24. Klemm P, Wheeler E. Cancer caregivers online: hope, emotional roller coaster, and physical/emotional/psychological responses. *Comput Inform Nurs* 2005;23(1):38-45. [Medline: [15681993](https://pubmed.ncbi.nlm.nih.gov/15681993/)]
25. Buis LR, Whitten P. Comparison of social support content within online communities for high- and low-survival-rate cancers. *Comput Inform Nurs* 2011 Aug;29(8):461-467. [doi: [10.1097/NCN.0b013e318214093b](https://doi.org/10.1097/NCN.0b013e318214093b)] [Medline: [21876394](https://pubmed.ncbi.nlm.nih.gov/21876394/)]
26. Sadah SA, Shahbazi M, Wiley MT, Hristidis V. A study of the demographics of Web-based health-related social media users. *J Med Internet Res* 2015;17(8):e194 [FREE Full text] [doi: [10.2196/jmir.4308](https://doi.org/10.2196/jmir.4308)] [Medline: [26250986](https://pubmed.ncbi.nlm.nih.gov/26250986/)]
27. 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](https://doi.org/10.1007/s00520-010-0935-5)] [Medline: [20556435](https://pubmed.ncbi.nlm.nih.gov/20556435/)]
28. Zheng R, Li J, Chen H, Huang Z. A framework for authorship identification of online messages: writing-style features and classification techniques. *J Am Soc Inf Sci* 2006 Feb 01;57(3):378-393. [doi: [10.1002/asi.20316](https://doi.org/10.1002/asi.20316)]
29. Abbasi A, Chen H. Writeprints: a stylometric approach to identity-level identification and similarity detection in cyberspace. *ACM Trans Inf Syst* 2008 Mar 01;26(2):1-29. [doi: [10.1145/1344411.1344413](https://doi.org/10.1145/1344411.1344413)]
30. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. *J Am Med Inform Assoc* 2006;13(1):24-29 [FREE Full text] [doi: [10.1197/jamia.M1761](https://doi.org/10.1197/jamia.M1761)] [Medline: [16221948](https://pubmed.ncbi.nlm.nih.gov/16221948/)]
31. Zeng QT, Tse T, Divita G, Keselman A, Crowell J, Browne AC, et al. Term identification methods for consumer health vocabulary development. *J Med Internet Res* 2007 Feb 28;9(1):e4 [FREE Full text] [doi: [10.2196/jmir.9.1.e4](https://doi.org/10.2196/jmir.9.1.e4)] [Medline: [17478413](https://pubmed.ncbi.nlm.nih.gov/17478413/)]
32. Keselman A, Smith CA, Divita G, Kim H, Browne AC, Leroy G, et al. Consumer health concepts that do not map to the UMLS: where do they fit? *J Am Med Inform Assoc* 2008;15(4):496-505 [FREE Full text] [doi: [10.1197/jamia.M2599](https://doi.org/10.1197/jamia.M2599)] [Medline: [18436906](https://pubmed.ncbi.nlm.nih.gov/18436906/)]
33. Doing-Harris KM, Zeng-Treitler Q. Computer-assisted update of a consumer health vocabulary through mining of social network data. *J Med Internet Res* 2011 May 17;13(2):e37 [FREE Full text] [doi: [10.2196/jmir.1636](https://doi.org/10.2196/jmir.1636)] [Medline: [21586386](https://pubmed.ncbi.nlm.nih.gov/21586386/)]
34. Goryachev S, Kim H, Zeng-Treitler Q. Identification and extraction of family history information from clinical reports. *AMIA Annu Symp Proc* 2008 Nov 06:247-251 [FREE Full text] [Medline: [18999129](https://pubmed.ncbi.nlm.nih.gov/18999129/)]
35. Witten I, Frank E. *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco, CA: Morgan Kaufmann Publishers; 2011.
36. Lu Y, Zhang P, Deng S. Exploring health-related topics in online health community using cluster analysis. 2013 Presented at: 46th Hawaii International Conference on System Sciences; January 7-10, 2013; Wailea, HI, USA p. 802-811. [doi: [10.1109/HICSS.2013.216](https://doi.org/10.1109/HICSS.2013.216)]
37. Esuli A, Sebastiani F. SENTIWORDNET: a publicly available lexical resource for opinion mining. 2006 Presented at: 5th Conference on Language Resources and Evaluation; May 24-26, 2006; Genoa, Italy p. 417-422.
38. Tetlock PC. Giving content to investor sentiment: the role of media in the stock market. *J Finance* 2007;62(3):1139-1168. [doi: [10.1111/j.1540-6261.2007.01232.x](https://doi.org/10.1111/j.1540-6261.2007.01232.x)]
39. Tetlock PC, Saar-Tsechansky M, Macskassy S. More than words: quantifying language to measure firms' fundamentals. *J Finance* 2008 May 09;63(3):1437-1467. [doi: [10.1111/j.1540-6261.2008.01362.x](https://doi.org/10.1111/j.1540-6261.2008.01362.x)]
40. Qiu B, Zhao K, Mitra P, Wu D, Caragea C, Yen J, et al. Get online support, feel better: sentiment analysis dynamics in an online cancer survivor community. 2011 Presented at: 3rd IEEE International Conference on Social Computing; October 9-11, 2011; Boston, MA, USA.
41. Denecke K, Taytsarau M, Palpanas T, Brosowski M. Topic-related sentiment analysis for discovering contradicting opinions in weblogs. Technical report #DISI-09-037. Trento, Italy: Dipartimento di Ingegneria e Scienze Dell'Informazione; 2009. URL: <http://eprints.biblio.unitn.it/1644/1/037.pdf> [accessed 2017-03-29] [WebCite Cache ID 6pKWH8hAq]
42. Blank TO, Adams-Blodnieks M. The who and the what of usage of two cancer online communities. *Comput Hum Behav* 2007 May;23(3):1249-1257. [doi: [10.1016/j.chb.2004.12.003](https://doi.org/10.1016/j.chb.2004.12.003)]

43. Grbich C, Parker D, Maddocks I. The emotions and coping strategies of caregivers of family members with a terminal cancer. *J Palliat Care* 2001;17(1):30-36. [Medline: [11324182](#)]
44. Flaskerud JH, Carter PA, Lee P. Distressing emotions in female caregivers of people with AIDS, age-related dementias, and advanced-stage cancers. *Perspect Psychiatr Care* 2000;36(4):121-130. [Medline: [12035204](#)]

Abbreviations

FM: Fowlkes-Mallows.

MeSH: Medical Subject Headings

UMLS: Unified Medical Language System

Edited by G Eysenbach; submitted 02.12.16; peer-reviewed by T Vaughan, B Bie; comments to author 10.02.17; revised version received 27.02.17; accepted 05.03.17; published 07.04.17.

Please cite as:

Lu Y, Wu Y, Liu J, Li J, Zhang P

Understanding Health Care Social Media Use From Different Stakeholder Perspectives: A Content Analysis of an Online Health Community

J Med Internet Res 2017;19(4):e109

URL: <http://www.jmir.org/2017/4/e109/>

doi: [10.2196/jmir.7087](https://doi.org/10.2196/jmir.7087)

PMID: [28389418](https://pubmed.ncbi.nlm.nih.gov/28389418/)

©Yingjie Lu, Yang Wu, Jingfang Liu, Jia Li, Pengzhu Zhang. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 07.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Pregnant Women's Views on the Feasibility and Acceptability of Web-Based Mental Health E-Screening Versus Paper-Based Screening: A Randomized Controlled Trial

Dawn Kingston¹, RN, MSc, PhD; Marie-Paule Austin², FRANZCP, MD; Sander Veldhuyzen van Zanten³, MSc, MPH, MD, PhD; Paula Harvalik¹, RN; Rebecca Giallo⁴, PhD; Sarah D McDonald⁵, MSc, FRCSC, MD; Glenda MacQueen⁶, PhD; Lydia Vermeyden¹, MSc; Gerri Lasiuk⁷, RN, RPN, PhD; Wendy Sword⁸, RN, PhD; Anne Biringer⁹, CCFP, FCFP, MD

¹Faculty of Nursing, University of Calgary, Calgary, AB, Canada

²University of New South Wales, Sydney, Australia

³Division of Gastroenterology, University of Alberta, Edmonton, AB, Canada

⁴Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Australia

⁵McMaster University, Hamilton, ON, Canada

⁶University of Calgary, Calgary, AB, Canada

⁷College of Nursing, University of Saskatchewan, Regina, SK, Canada

⁸University of Ottawa, Ottawa, ON, Canada

⁹Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Dawn Kingston, RN, MSc, PhD

Faculty of Nursing

University of Calgary

2500 University Ave NW

Calgary, AB, T2N 1N4

Canada

Phone: 1 4032202634

Fax: 1 4032844803

Email: dawn.kingston@ucalgary.ca

Abstract

Background: Major international guidelines recommend mental health screening during the perinatal period. However, substantial barriers to screening have been reported by pregnant and postpartum women and perinatal care providers. E-screening offers benefits that may address implementation challenges.

Objective: The primary objective of this randomized controlled trial was to evaluate the feasibility and acceptability of Web-based mental health e-screening compared with paper-based screening among pregnant women. A secondary objective was to identify factors associated with women's preferences for e-screening and disclosure of mental health concerns.

Methods: Pregnant women recruited from community and hospital-based antenatal clinics and hospital-based prenatal classes were computer-randomized to a fully automated Web-based e-screening intervention group or a paper-based control group. Women were eligible if they spoke or read English, were willing to be randomized to e-screening, and were willing to participate in a follow-up diagnostic interview. The intervention group completed the Antenatal Psychosocial Health Assessment and the Edinburgh Postnatal Depression Scale on a tablet computer, while controls completed them on paper. All women completed self-report baseline questions and were telephoned 1 week after randomization by a blinded research assistant for a MINI International Neuropsychiatric Interview. Renker and Tonkin's tool of feasibility and acceptability of computerized screening was used to assess the feasibility and acceptability of e-screening compared with paper-based screening. Intention-to-treat analysis was used. To identify factors associated with preference for e-screening and disclosure, variables associated with each outcome at $P < .20$ were simultaneously entered into final multivariable models to estimate adjusted odds ratios (AORs) and 95% CIs.

Results: Of the 675 eligible women approached, 636 agreed to participate (participation rate 94.2%) and were randomized to the intervention ($n=305$) or control ($n=331$) groups. There were no significant baseline differences between groups. More women

in the e-screening group strongly or somewhat agreed that they would like to use a tablet for answering questions on emotional health (57.9%, 175/302 vs 37.2%, 121/325) and would prefer using a tablet to paper (46.0%, 139/302 vs 29.2%, 95/325), compared with women in the paper-based screening group. There were no differences between groups in women's disclosure of emotional health concerns (94.1%, 284/302 vs 90.2%, 293/325). Women in the e-screening group consistently reported the features of e-screening more favorably than controls (more private or confidential, less impersonal, less time-consuming). In the multivariable models, being in the e-screening group was significantly associated with preferring e-screening (AOR 2.29, 95% CI 1.66-3.17), while no factors were significantly associated with disclosure.

Conclusions: The findings suggest that mental health e-screening is feasible and acceptable to pregnant women.

Trial Registration: Clinicaltrials.gov NCT01899534; <https://clinicaltrials.gov/ct2/show/NCT01899534> (Archived by WebCite at <http://www.webcitation.org/6ntWg1yWb>)

(*J Med Internet Res* 2017;19(4):e88) doi:[10.2196/jmir.6866](https://doi.org/10.2196/jmir.6866)

KEYWORDS

pregnancy; mental health; screening; prenatal care; computers

Introduction

The Need for Mental Health Screening

Depression and anxiety are among the most common morbidities in pregnancy, with prevalence rates of 13%-29% [1-3], and are the leading causes of maternal mortality in Western countries [4]. Without early screening and treatment, 50% to 70% of women with prenatal anxiety or depression symptoms [5] will experience persistent symptoms through their child's early years [6,7]. Recent findings from longitudinal birth cohorts also reveal that chronicity of depression starting in pregnancy, whether severe or subclinical, is emerging as a major risk factor in child development and mental health [8-11].

In the absence of routine screening, mental health problems are severely underdetected and undertreated in perinatal settings [12,13]. This is, in part, due to the fact that the majority of pregnant and postpartum women do not volunteer information about their mental health without being prompted by a perinatal care provider [14-18]. In our previous studies, we reported that while 67% of pregnant women did not raise concerns about their mental health with their physician or obstetrician, nurse, or midwife, 97% indicated they were comfortable with provider-initiated screening [19]. Our studies have identified other common barriers that deter women from self-identifying mental health problems, including "false" reassurance that they have received from friends or family, not knowing whether their symptoms are "normal" or not during pregnancy, and stigma-induced concerns such as not wanting their care provider to see them as depressed or anxious and not wanting to be seen as a bad mother [19]. However, 79% of pregnant women reported they would disclose mental health concerns if asked as part of their routine prenatal care [19,20]. Clearly, routine, provider-initiated mental health screening plays an important role in women's willingness to disclose prenatal mental health problems.

Studies also show that underdetection of mental health problems in the perinatal population is due to the lack of standardized screening, where up to 80% of cases remain unidentified by perinatal care providers who used unstandardized approaches for detection of mental health disorders (eg, without a validated tool) [21,22]. However, recent studies reveal that routine

prenatal and postnatal screening enhances detection and increases the likelihood that women with a positive screening result for depression will link to mental health services [23,24]. Thus, antenatal mental health screening holds benefits for enhanced detection, increased linkage to services, and improved clinical outcomes.

Although current evidence and international guidelines from the United Kingdom [25], Australia [26], and the United States [27,28] support the need for antenatal screening as a key intervention for interrupting the cycle of perinatal mental disorders and their negative impact on maternal and child well-being, serious challenges exist for its implementation. Despite high acceptance by women [29-31] and providers [32-35], only 20% of North American perinatal care providers conduct proactive screening as part of prenatal care [36] and less than 15% of pregnant or postpartum women receive the help they need [37]. A systematic review conducted by our team (manuscript in preparation) found that substantial personal and system-based barriers to routine screening exist for health care providers, including lack of time to screen, lack of accurate assessment tools and knowing how to interpret them, lack of defined referral processes, the absence of connections with mental health services, and frustration with the lack of availability of timely services for their pregnant patients [38-40].

Taken together, this body of research underscores the need to develop screening processes that are feasible for and acceptable to both women and service providers, are designed to overcome barriers to implementation, and are cost-effective and clinically useful. Indeed, the most effective perinatal mental health screening and management programs are those characterized by screening processes that are incorporated into routine care with designated systems of referral and treatment that are initiated immediately after screening [26,41,42].

The Potential Impact of E-Screening

E-screening has potential to be an effective, low-resource screening approach that can be feasibly embedded in a variety of perinatal settings [43,44]. Importantly, it has potential to overcome the personal and system-based barriers to screening identified by pregnant women and health care providers, and it can screen for sensitive issues such as intimate partner violence [45-47] and postpartum depression [48]. It can increase

efficiency of mental health care by reallocating scarce human resources to where they are most needed—in-depth follow-up assessment, referral, and treatment. It is well-suited for busy clinical settings, can be personalized to patient needs, offers audio or video options for low literacy, provides real-time data [43,44], achieves similar or greater rates of disclosure compared with interviews, and is preferred by patients because of its anonymity [44,45,49,50]. However, to date no studies have evaluated the feasibility or acceptability of e-screening as an approach to routine screening in pregnant women.

Objectives

The objective of this study was to compare the views of pregnant women randomized to an intervention group and a paper-based screening control group on the feasibility (eg, disclosure of

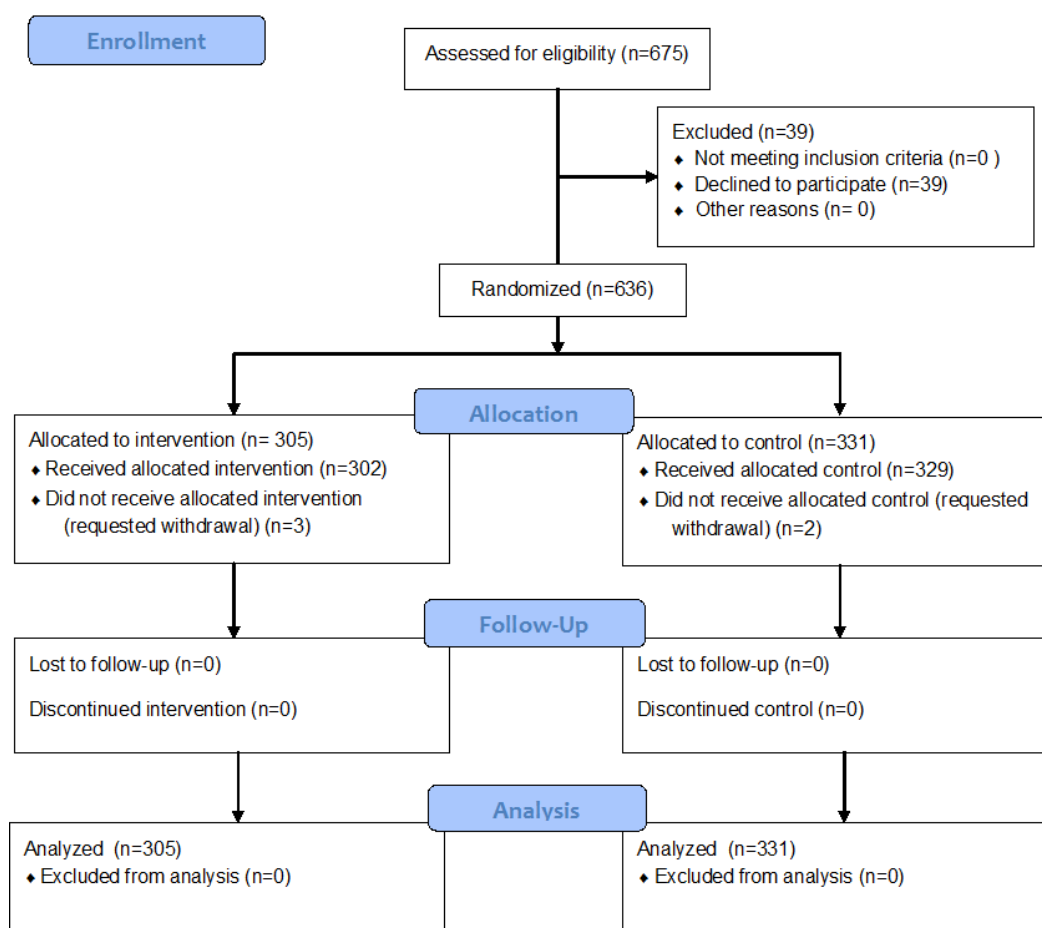
concerns about their mental health, specific features of screening) and acceptability (eg, women's preference) of e-screening. A secondary objective was to identify factors associated with women's preferences for e-screening and their ability to disclose mental health concerns.

Methods

Study Design

The study was a parallel-group, randomized controlled superiority trial (Figure 1). The protocol has been previously published (ClinicalTrials.gov identifier: NCT01899534) [51]. Approval for this study was granted by the Human Research Ethics Board at the University of Alberta.

Figure 1. CONSORT flow diagram.



Eligibility Criteria

Pregnant women were eligible if they were (1) able to speak or read English; (2) willing to be randomized to e-screening; and (3) willing to participate in a follow-up diagnostic interview within 1 week of recruitment. Because the Web-based screening

tool was intended to be completed unassisted, it was designed for use by women with varying degrees of computer literacy.

Setting and Recruitment

Women were recruited from 2 community-based family physician-led maternity clinics, a high-risk antenatal unit in a

tertiary care center, and 2 community hospital-based prenatal classes in Edmonton, Alberta. The high-risk antenatal unit is located in the inner city and serves a demographically diverse population. One maternity clinic was located in a middle-class suburb, while the other was located in a neighborhood with a high proportion of new immigrant families. The hospital-based prenatal classes served women from across the urban center. Thus, the recruitment strategy aimed to include participants with diverse demographic and obstetric characteristics. None of the care providers at the recruitment sites had specialized training in mental health care. Additional details on the recruitment sites are available in the study's published protocol [51].

Trained research assistants at each site used a standardized script to invite women to participate in the study. Once women completed the consent form electronically on a tablet computer, the computer program designed by the Women and Children's Health Research Institute (WCHRI) automatically randomized them (1:1) to the intervention or control group. Thus, the research assistant was blinded to group allocation. Full details on recruitment and consent procedures are found in the trial protocol [51].

The E-Screening Intervention

The e-screening intervention was a full psychosocial assessment including (1) the Antenatal Psychosocial Health Assessment (ALPHA) [21,35] and (2) the Edinburgh Postnatal Depression Scale (EPDS) [52]. This is aligned with international guidelines recommending that screening comprise an evaluation of risks for mental health problems (ALPHA) as well as current symptoms (EPDS) [26,42].

The ALPHA asks women questions on the topics of family life, stressors, feelings about the pregnancy, substance use, abuse, and family of origin. Items are rated as yes or no or 6-point Likert-scale options and are completed in 5 minutes. On the basis of a review of responses in each category, providers subjectively judge whether women are at low, some, or high psychosocial risk. The ALPHA has undergone extensive face and content validity testing, can readily be implemented in routine prenatal care, and pregnant women find it acceptable [35,53].

The 10-item EPDS is one of the most widely used screening instruments for detecting antenatal and postnatal depression symptoms within the previous 7 days [52]. A score of 13 or higher (range 0-30) is a well-established cutoff for clinically significant depression symptoms consistent with meeting the criteria for a major depressive episode [52]. Original psychometric testing resulted in a sensitivity of 85%, specificity of 77%, positive predictive value of 83%, split-half reliability of .8, and Cronbach alpha of .87 [52]. A screenshot of the Web-based version of the EPDS is provided in [Multimedia Appendix 1](#).

Women randomized to the intervention group accessed the intervention for free. They completed the assessment on a single occasion and had no further access to the intervention beyond that time. Research assistants recruiting the women provided technical assistance as needed to assist women to get started on

the tablet computer, although the link was labeled and readily available on the home screen. No specialized training was offered to the intervention group because completion of the Web-based e-screening tool was intended to be a largely self-sufficient, self-report option for clinical settings. No study participants were aware which group was the intervention of interest.

The Paper-Based Screening Control Group

Women in the control group completed paper-based versions of the ALPHA and EPDS, followed by the Web-based baseline questionnaire.

Procedures

Following recruitment, women completed the Web-based e-screening or paper-based version of the screening tools independently (in the recruitment setting) as well as a Web-based baseline questionnaire including (1) demographics (age, parity, marital status, education, income, ethnicity, country of birth, and length of time in Canada); (2) obstetric history (current and past, including use of fertility treatments); (3) mental health history (diagnoses, treatment); (4) level of comfort with computer technology; (5) quality of relationship with perinatal care provider; (6) level of social support, experience of talking with doctor, nurse, or midwife about emotional health; and (7) adverse childhood experiences (using ACE questionnaire [54]). With the exception of the ACE questionnaire, questions were drawn from those used in the All Our Babies Study [55,56] and the Canadian Maternity Experiences Survey [57]. The baseline questionnaire also included Renker and Tonkin's Computer Assessment Evaluation (CAE), which is a feasibility and acceptability assessment. Logos belonging to institutional affiliations were visible on questionnaires and consent forms, as well as in the description of the primary investigator. Women in both groups were telephoned by a trained research assistant (blinded to group allocation), 1 week after recruitment, to complete a MINI International Neuropsychiatric Interview (MINI, version 6.0.0) [58].

No data were stored on the tablets. When women submitted their information, it was sent to a secure server housed in the Faculty of Medicine & Dentistry at the University of Alberta. The Web-based e-screening tool and questionnaire were built using an existing infrastructure offered by Checkbox Survey software provided by the WCHRI Clinical Research Informatics Core. The Web-based e-screening tool was tested for usability and navigability through focus groups as part of a concurrent randomized controlled trial, as described in that trial's published protocol [59]. No changes to the tool were made after initial testing. Data transfer between the tablet and server was encrypted. Data imported to statistical databases for analysis were not identifiable.

Safety Protocol

If women met the criteria for a mood or anxiety disorder on the MINI or if they scored 13 or more on the EPDS, the research assistant made a referral for the women (with their permission) to the hospital-based reproductive mental health support program.

Sample Size

Because no data were available to guide estimation of a minimal clinically important difference in true cases detected through e-screening, we used a confidence interval approach [60]. On the basis of high levels of acceptability and disclosure reported using computer screening [46,47], we established that e-screening would be feasible if, in the intervention group, 85% of women indicated that they were able to tell the truth on all questions on the psychosocial assessment (question 7 of the computer assessment evaluation, CAE). For our other outcomes, we based the sample size calculation on 85% of women scoring 4-8 on the Risk subscale of the Disclosure Expectations Scale and 85% scoring 16-20 on the Utility subscale of the Disclosure Expectations Scale. Using a margin of error of .05 and 25% as an estimated loss to follow-up, we calculated that 261 women per group (N=522) were required (see protocol [51] for calculation). At a final sample size of 636, the study is sufficiently powered to detect differences in the outcomes between groups if they exist.

Measurement of Outcomes

We measured women's views on the feasibility and acceptability of e-screening with an adapted 9-item version of Renker and Tonkin's assessment of the feasibility of computerized screening for interpersonal violence [46,47]. Of the 9 items, 2 items are related to acceptability of e-screening (I did or would like to use a tablet to answer these questions about emotional health; I would prefer answering questions about emotional health on the tablet compared to a paper questionnaire) and the remaining 7 items measure broad areas of feasibility, including disclosure (I was able to tell the truth on all the questions about emotional health; I did not like answering the questions about emotional health on the tablet or paper because it was hard to answer how I really felt), features of screening (privacy; ease of understanding questions; impersonal; time to completion), and comparison with paper-based screening. Women in the intervention group answered questions rating their experience of e-screening (eg, *I liked* answering questions about emotional health on a tablet because it felt private). Those in the control group answered questions to assess their views on e-screening if they had the opportunity (eg, *I would like* answering questions about emotional health on a tablet because it would feel private). Responses were rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree).

Analysis

Intention-to-treat analysis was conducted for all research questions. We used descriptive data (frequencies and 95% CIs; means and SDs) to describe the sample. Baseline differences in groups were compared using independent *t* tests (means) and chi-square tests (%) to determine the extent to which

randomization was successful. Statistical significance for all analyses and final models was set at $P < .05$. We used chi-square tests to compare proportions of women in each group responding affirmatively to questions on feasibility and acceptability.

For analyses related to factors associated with the outcomes of preference for e-screening and disclosure, we conducted bivariate analyses to identify independent factors that were significantly associated with each of the outcomes at $P < .20$, estimating unadjusted odds ratios and their 95% CIs. Those variables were entered in the 2 final multivariable models, where $P < .05$ defined factors that were significantly associated with the outcomes in the final models. All significant independent variables were entered into the multivariable models simultaneously, and each variable was controlled for by all other variables in the model. For the final models, we reported adjusted odds ratios and their 95% CIs.

Results

Sample Characteristics

Of the 675 eligible women who were approached between August 2013 and January 2015, a total of 636 (94.2%) women agreed to participate and were randomized to the intervention (n=305) or control (n=331) group. No women were deemed ineligible after the initial screening for eligibility. A total of 5 women withdrew from the study following group allocation, 3 in the intervention group and 2 in the control group. There were no statistically significant differences in relation to demographic variables between the 2 groups (Table 1).

The majority of pregnant women were 25 to 34 years of age, partnered (married, common-law, or living with a partner), were white, had incomes of Can \$80,000 or more, had at least some postsecondary education, were pregnant with their first child, and were in their first trimester of pregnancy (80.3%, 494/615; see Table 1). More than a quarter of the participants had been diagnosed and treated for anxiety, depression, or another mental health concern before joining the study, and 18.0% (113/629) reported having 4 or more adverse childhood experiences. There were no statistically significant differences between the intervention and control groups in demographic characteristics, comfort with computers, tablets, or mobile phone previous mental health diagnoses or treatment, obstetric history, mean gestational age, or number of adverse childhood experiences. More than two-thirds of women (n=423) were recruited from community-based maternity clinics, 21.0% (131/624) were recruited from community-based prenatal classes, and 11.2% (70/624) were recruited from the high-risk antenatal unit. Missing data were 3.3% (615/636) or less for all variables, with the majority having less than 1.5% (10/636); thus, data imputation was not used.

Table 1. Sample characteristics (N=636).

Characteristics	Full sample (N=636 ^a), n (%)	Paper-based screening group (n=331 ^a), n (%)	E-screening group (n=305 ^a), n (%)	<i>P</i> value ^b
Recruitment site				
Community-based clinics	423 (67.8)	224 (70.0)	199 (65.5)	.47
High-risk antenatal unit at tertiary care center	70 (11.2)	34 (10.6)	36 (11.8)	
Prenatal classes	131 (21.0)	62 (19.4)	69 (22.7)	
Age, years				
<25	88 (13.9)	50 (15.2)	38 (12.5)	.51
25-34	459 (72.2)	233 (70.6)	226 (74.6)	
35+	86 (13.6)	47 (14.2)	39 (12.9)	
Income in Can \$				
Less than \$40,000	97 (15.4)	52 (15.8)	45 (14.9)	.81
\$40,000-\$79,999	139 (22.0)	75 (22.8)	64 (21.2)	
\$80,000 or more	395 (62.6)	202 (61.4)	193 (63.9)	
Education				
High school or less	100 (15.8)	57 (17.3)	43 (14.2)	.29
Some postsecondary or more	531 (84.2)	272 (82.7)	259 (85.8)	
Marital status				
Unpartnered (single, divorced, or separated)	27 (4.3)	14 (4.3)	13 (4.3)	.98
Partnered (married, common-law, or living with a partner)	604 (95.7)	315 (95.7)	289 (95.7)	
Ethnicity				
Not white	169 (26.8)	91 (27.7)	78 (25.8)	.60
White	462 (73.2)	238 (72.3)	224 (74.2)	
Born in Canada				
No	119 (18.9)	66 (20.1)	53 (17.5)	.42
Yes	512 (81.1)	263 (79.9)	249 (82.5)	
Ever diagnosed with depression, anxiety, or any other kind of emotional concern				
Yes	164 (25.9)	86 (26.1)	78 (25.7)	.91
No	470 (74.1)	244 (73.9)	226 (74.3)	
Ever treated for depression, anxiety, or any other kind of emotional concern				
Yes	179 (28.2)	92 (27.9)	87 (28.6)	.84
No	455 (71.8)	238 (72.1)	217 (71.4)	
Pregnant before				
First child	426 (69.3)	213 (68.5)	213 (70.1)	.67
Not first child	189 (30.7)	98 (31.5)	91 (29.9)	
Gestational age, mean (SD)	9.00 (6.46)	8.61 (6.08)	9.39 (6.80)	.22
Used fertility treatments to become pregnant				
Yes	35 (5.5)	17 (5.2)	18 (5.9)	.67
No	599 (94.5)	313 (94.8)	286 (94.1)	
ACE score				

Characteristics	Full sample (N=636 ^a), n (%)	Paper-based screening group (n=331 ^a), n (%)	E-screening group (n=305 ^a), n (%)	P value ^b
Score greater than or equal to 4	113 (18.0)	64 (19.5)	49 (16.3)	.31
Score less than 4	516 (82.0)	265 (80.5)	251 (83.7)	
I am comfortable using a computer or laptop				
Very comfortable	591 (93.7)	311 (94.5)	280 (92.7)	.45
Somewhat comfortable	36 (5.7)	17 (5.2)	19 (6.3)	
Not very comfortable	4 (0.6)	1 (0.3)	3 (1.0)	
I am comfortable using a computer tablet (eg, iPad)				
Very comfortable	530 (84.0)	280 (85.1)	250 (82.8)	.64
Somewhat comfortable	89 (14.1)	44 (13.4)	45 (14.9)	
Not very comfortable	12 (1.9)	5 (1.5)	7 (2.3)	
I am comfortable using a smartphone				
Very comfortable	546 (86.5)	286 (86.9)	260 (86.1)	.32
Somewhat comfortable	70 (11.1)	38 (11.6)	32 (10.6)	
Not very comfortable	15 (2.4)	5 (1.5)	10 (3.3)	

^aSome demographic data missing.

^bComparison of control and intervention groups: chi-square statistic used for variables with 3 or more categories; two-tailed *t* test used for variables with estimated means.

Primary Objectives

Acceptability of E-Screening

More women in the e-screening group strongly or somewhat agreed that they would like to use or did like using a tablet for answering questions on emotional health (57.9%, 175/302 vs 37.2%, 121/325) and would or did prefer using a tablet to paper (46.0%, 139/302 vs 29.2%, 95/325), compared with women in the paper-based screening group. We observed, too, that fewer women who used the tablet answered *does not matter one way or the other* than women who completed paper-based screening on the items *I would or did like to use a tablet to answer questions about emotional health* (30.8%, 93/302 vs 53.5%, 174/325) and *I would prefer answering questions about emotional health on the tablet compared to a paper questionnaire* (40.4%, 122/302 vs 60.9%, 198/325).

Feasibility of E-Screening

Disclosure

Overall, women in both e-screening and paper-based screening groups indicated that they would be able to disclose their concerns about their mental health (Table 2). There was no significant difference between groups on the item *I was able to tell the truth on all the questions about emotional health*, with 94.1% (284/302) of women in the e-screening intervention group and 90.2% (293/325) in the paper-based control group somewhat or strongly agreeing they could tell the truth on all

questions. In addition, few women in both groups indicated that they would find it difficult to answer how they felt with e-screening.

Features of Screening

Women in the e-screening group consistently reported the features of e-screening as superior to paper-based screening. For instance, significantly more women in the e-screening group (vs paper-based screening) reported they would like e-screening because it was private (64.6%, 195/302 vs 31.7%, 103/325) and they perceived the questions to be easier to understand (88.8%, 268/302 vs 87.7%, 285/325). Similarly, fewer women in the e-screening group reported they would find e-screening impersonal (3.6%, 11/302 participants vs 4.9%, 16/325 participants). Finally, significantly more women in the e-screening group reported they did not find screening too time-consuming, compared with women in the paper-based screening group (57.9%, 54/302 vs 40.9%, 40/325).

Comparison With Face-to-Face Screening

In both groups, less than 10% (24/302 and 30/325) of women acknowledged that it would be easier to have a nurse ask questions about emotional health compared with completing questions on a tablet. More than half the women in both groups preferred self-report e-screening compared with face-to-face screening, and this proportion was significantly greater in the e-screening group (58.9%, 178/302 participants vs 50.7%, 165/325).

Table 2. Feasibility and acceptability of e-screening.

Individual items of computer assessment evaluation scale	Full sample (N=627 ^a), n (%)	Group 1: paper-based screening (n=325 ^a), n (%)	Group 2: e-screening (n=302 ^a), n (%)	<i>P</i> value ^d
1. I would or did like to use a tablet to answer these questions about emotional health^b				
Strongly disagree	42 (6.7)	23 (7.1)	19 (6.3)	<.001
Somewhat disagree	22 (3.5)	7 (2.2)	15 (5.0)	
Does not matter one way or the other	267 (42.6)	174 (53.5)	93 (30.8)	
Somewhat agree	65 (10.4)	32 (9.8)	33 (10.9)	
Strongly agree	231 (36.8)	89 (27.4)	142 (47.0)	
2. I found the questions about emotional health easy to understand				
Strongly disagree	42 (6.7)	17 (5.2)	25 (8.3)	.03
Somewhat disagree	13 (2.1)	10 (3.1)	3 (1.0)	
Does not matter one way or the other	19 (3.0)	13 (4.0)	6 (2.0)	
Somewhat agree	122 (19.5)	72 (22.2)	50 (16.6)	
Strongly agree	431 (68.7)	213 (65.5)	218 (72.2)	
3. I would not or did not like answering questions about emotional health on a tablet because it felt or would feel impersonal^b				
Strongly disagree	306 (48.9)	139 (42.9)	167 (55.3)	.01
Somewhat disagree	66 (10.5)	32 (9.9)	34 (11.3)	
Does not matter one way or the other	227 (36.3)	137 (42.3)	90 (29.8)	
Somewhat agree	17 (2.7)	10 (3.1)	7 (2.3)	
Strongly agree	10 (1.6)	6 (1.8)	4 (1.3)	
4. I would prefer answering questions about emotional health on a tablet compared to a paper questionnaire				
Strongly disagree	43 (6.9)	18 (5.5)	25 (8.3)	<.001
Somewhat disagree	30 (4.8)	14 (4.3)	16 (5.3)	
Does not matter one way or the other	320 (51.0)	198 (60.9)	122 (40.4)	
Somewhat agree	63 (10.0)	27 (8.3)	36 (11.9)	
Strongly agree	171 (27.3)	68 (20.9)	103 (34.1)	
5. I would find it easier to answer the questions about emotional health on a tablet rather than having a nurse ask me questions				
Strongly disagree	24 (3.8)	12 (3.7)	12 (4.0)	.02
Somewhat disagree	30 (4.8)	18 (5.5)	12 (4.0)	
Does not matter one way or the other	230 (36.7)	130 (40.0)	100 (33.1)	
Somewhat agree	173 (27.6)	96 (29.5)	77 (25.5)	
Strongly agree	170 (27.1)	69 (21.2)	101 (33.4)	
6. I did not like answering the questions about emotional health on the tablet or paper because it was hard to answer how I really felt^c				
Strongly disagree	251 (40.0)	90 (27.7)	161 (53.3)	<.001

Individual items of computer assessment evaluation scale	Full sample (N=627 ^a), n (%)	Group 1: paper-based screening (n=325 ^a), n (%)	Group 2: e-screening (n=302 ^a), n (%)	P value ^d
Somewhat disagree	89 (14.2)	46 (14.2)	43 (14.2)	
Does not matter one way or the other	218 (34.8)	152 (46.8)	66 (21.9)	
Somewhat agree	60 (9.6)	30 (9.2)	30 (9.9)	
Strongly agree	9 (1.4)	7 (2.2)	2 (0.7)	
7. I would or did like answering questions about emotional health on a tablet because it would or did feel private^b				
Strongly disagree	23 (3.7)	16 (4.9)	7 (2.3)	<.001
Somewhat disagree	35 (5.6)	19 (5.8)	16 (5.3)	
Does not matter one way or the other	271 (43.2)	187 (57.5)	84 (27.8)	
Somewhat agree	155 (24.7)	66 (20.3)	89 (29.5)	
Strongly agree	143 (22.8)	37 (11.4)	106 (35.1)	
8. I did not like answering questions about emotional health on the tablet or on paper because the questions took too long for me to answer^c				
Strongly disagree	213 (34.0)	93 (28.6)	120 (39.7)	<.001
Somewhat disagree	95 (15.2)	40 (12.3)	55 (18.2)	
Does not matter one way or the other	225 (35.9)	152 (46.8)	73 (24.2)	
Somewhat agree	78 (12.4)	32 (9.8)	46 (15.2)	
Strongly agree	16 (2.6)	8 (2.5)	8 (2.6)	
9. I was able to tell the truth on all the questions about emotional health				
Strongly disagree	20 (3.2)	14 (4.3)	6 (2.0)	.25
Somewhat disagree	6 (1.0)	5 (1.5)	1 (0.3)	
Does not matter one way or the other	24 (3.8)	13 (4.0)	11 (3.6)	
Somewhat agree	64 (10.2)	33 (10.2)	31 (10.3)	
Strongly agree	513 (81.8)	260 (80.0)	253 (83.8)	

^aTotal sample missing data=9 (group 1: paper-based screening=6; group 2: e-screening=3).

^bWomen in the intervention group answered questions rating their experience of e-screening (eg, I *liked* answering questions about emotional health on a tablet because it felt private). Those in the control group answered questions to assess their views on e-screening (eg, I *would like* answering questions about emotional health on a tablet because it would feel private).

^cWomen in the intervention group answered questions regarding the experience of e-screening (eg, I *did not like* answering questions about emotional health on the tablet because the questions took too long for me to answer). Those in the control group answered questions regarding the experience of paper-based screening (eg, I *did not like* answering questions about emotional health on paper because the questions took too long for me to answer).

^dComparison of control and intervention groups: chi-square statistic used for all variables.

Secondary Objectives

Factors Associated With Preference for E-Screening

Table 3 presents the results of the bivariate analysis of the association of each independent variable with the outcome, *I would or did like to use the tablet to answer these questions about emotional health*. Independent variables associated with the outcome at $P<.20$ were entered into the final multivariable logistic regression model, including mode of screening, maternal

age, income, and education. In the final model, only mode of screening was significantly associated with the outcome, where the odds of preferring e-screening were 2.29 times greater for women in the e-screening group than women in the paper-based screening group. When we repeated this analysis using the outcome, *I would or did prefer answering questions about emotional health on the tablet compared to a paper questionnaire*, findings were similar to the odds of preferring e-screening estimated at 2.03 (95% CI 1.46-2.84).

Table 3. Factors associated with preference for e-screening based on the outcome, *I would or did like using a tablet to answer these questions about emotional health.*

Independent variables	Strongly or somewhat agree, n (%)	Does not matter, somewhat or strongly disagree ^a , n (%)	UOR ^b (95% CI)	AOR ^c (95% CI)
Mode of report				
Electronic	176 (59.3)	127 (38.4)	2.34 (1.70-3.22) ^d	2.29 (1.66-3.17)
Paper-based	121 (40.7)	204 (61.6)	1.00	1.00
Recruitment site				
Community-based clinic or prenatal class	261 (89.1)	288 (88.9)	1.02 (0.62-1.69)	
High-risk antenatal unit at tertiary care center	32 (10.9)	36 (11.1)	1.00	
Age, years				
Less than 25	31 (10.5)	57 (17.2)	0.56 (0.35-0.90) ^d	0.68 (0.41-1.13)
Greater than or equal to 25	265 (89.5)	274 (82.8)	1.00	1.00
Income				
Less than \$40,000	38 (12.8)	59 (17.8)	0.68 (0.44-1.06) ^d	0.77 (0.49-1.23)
\$40,000 or more	258 (87.2)	272 (82.2)	1.00	1.00
Marital status				
Unpartnered	11 (3.7)	16 (4.8)	0.76 (0.35-1.67)	
Partnered	285 (96.3)	315 (95.2)	1.00	
Diagnosis of depression, anxiety, or any other kind of emotional concern				
Yes	74 (24.9)	88 (26.6)	0.92 (0.64-1.31)	
No	223 (75.1)	243 (73.4)	1.00	
Ever treated for depression, anxiety, or any other kind of emotional concern				
Yes	81 (27.3)	97 (29.3)	0.91 (0.64-1.28)	
No	216 (72.7)	234 (70.7)	1.00	
Ethnicity				
Not white	78 (26.4)	89 (26.9)	0.97 (0.68-1.39)	
White	218 (73.6)	242 (73.1)	1.00	
Born in Canada				
No	61 (20.6)	57 (17.2)	1.25 (0.84-1.86)	
Yes	235 (79.4)	274 (82.8)	1.00	
Pregnant before				
First child	197 (67.7)	228 (70.8)	0.86 (0.61-1.22)	
Not first child	94 (32.3)	94 (29.2)	1.00	
Education				
High school or less	37 (12.5)	63 (19.0)	0.61 (0.39-0.94) ^d	0.73 (0.45-1.18)
Some postsecondary or more	259 (87.5)	268 (81.0)	1.00	1.00
Used fertility treatments to become pregnant				
Yes	18 (6.1)	17 (5.1)	1.19 (0.60-2.36)	
No	279 (93.9)	314 (94.9)	1.00	

Independent variables	Strongly or somewhat agree, n (%)	Does not matter, somewhat or strongly disagree ^a , n (%)	UOR ^b (95% CI)	AOR ^c (95% CI)
ACE score				
Score greater than or equal to 4	49 (16.7)	63 (19.0)	0.85 (0.56-1.28)	
Score less than 4	245 (83.3)	268 (81.0)	1.00	

^aThe categories *does not matter* and *somewhat or strongly disagree* were combined to address low cell sizes in the *somewhat or strongly disagree* category for some variables.

^bUOR: unadjusted odds ratio.

^cAOR: adjusted odds ratio.

^dIndependent variables associated with the outcome at $P < .20$ were entered into the final multivariable logistic regression model, including mode of screening, maternal age, income, and education.

Disclosure

Table 4 presents the results of the bivariate analysis of the association of each independent variable with the outcome, *I was able to tell the truth on all the questions about emotional health*. Independent variables associated with the outcome at

$P < .20$ were entered into the final multivariable logistic regression model, including mode of screening and education. In the final model, no variables were significantly associated with a woman's ability to be honest during mental health screening.

Table 4. Factors associated with disclosure of mental health concerns during screening based on the outcome, *I was able to tell the truth on all questions about emotional health.*

Independent variables	Strongly or somewhat agree, n (%)	Does not matter, somewhat or strongly disagree ^a , n (%)	UOR ^b (95% CI)	AOR ^c (95% CI)
Mode of report				
Electronic	285 (49.3)	18 (36.0)	1.73 (0.95-3.15) ^d	1.69 (0.92-3.08)
Paper-based	293 (50.7)	32 (64.0)	1.00	1.00
Recruitment site				
Community-based clinic or prenatal class	507 (89.1)	42 (87.5)	1.17 (0.48-2.86)	
High-risk antenatal unit at tertiary care center	62 (10.9)	6 (12.5)	1.00	
Age, years				
Less than 25	80 (13.9)	8 (16.0)	0.85 (0.38-1.87)	
Greater than or equal to 25	497 (86.1)	42 (84.0)	1.00	
Income in Can \$				
Less than \$40,000	90 (15.6)	7 (14.0)	1.14 (0.50-2.60)	
\$40,000 or more	487 (84.4)	43 (86.0)	1.00	
Marital status				
Unpartnered	25 (4.3)	2 (4.0)	1.09 (0.25-4.73)	
Partnered	552 (95.7)	48 (96.0)	1.00	
Diagnosis of depression, anxiety, or any other kind of emotional concern				
Yes	150 (26.0)	12 (24.0)	1.11 (0.57-2.18)	
No	428 (74.0)	38 (76.0)	1.00	
Ever treated for depression, anxiety, or any other kind of emotional concern				
Yes	164 (28.4)	14 (28.0)	1.02 (0.54-1.94)	
No	414 (71.6)	36 (72.0)	1.00	
Ethnicity				
Not white	151 (26.2)	16 (32.0)	0.75 (0.40-1.40)	
White	426 (73.8)	34 (68.0)	1.00	
Born in Canada				
No	110 (19.1)	8 (16.0)	1.24 (0.57-2.71)	
Yes	467 (80.9)	42 (84.0)	1.00	
Pregnant before				
First child	394 (70.0)	31 (62.0)	1.43 (0.79-2.60)	
Not first child	169 (30.0)	19 (38.0)	1.00	
Education				
High school or less	87 (15.1)	13 (26.0)	0.51 (0.26-0.99) ^d	0.52 (0.27-1.02)
Some postsecondary or more	490 (84.9)	37 (74.0)	1.00	1.00
Used fertility treatments to become pregnant				
Yes	32 (5.5)	3 (6.0)	0.92 (0.27-3.11)	
No	546 (94.5)	47 (94.0)	1.00	
ACE score				

Independent variables	Strongly or somewhat agree, n (%)	Does not matter, somewhat or strongly disagree ^a , n (%)	UOR ^b (95% CI)	AOR ^c (95% CI)
Score greater than or equal to 4	106 (18.4)	6 (12.0)	1.66 (0.69-3.99)	
Score less than 4	469 (81.6)	44 (88.0)	1.00	

^aThe categories *does not matter* and *somewhat or strongly disagree* were combined to address low cell sizes in the *somewhat or strongly disagree* category for some variables.

^bUOR: unadjusted odds ratio.

^cAOR: adjusted odds ratio.

^dIndependent variables associated with the outcome at $P < .20$ were entered into the final multivariable logistic regression model.

Discussion

Principal Findings

The findings of this study suggest that e-screening is a feasible approach to mental health screening. Overall, women in both e-screening and paper-based screening groups indicated that they would be able to disclose their concerns about their mental health, with women in the e-screening group consistently reporting the features of e-screening as superior to paper-based screening. In the multivariable analyses, we found that women in the e-screening group were more likely to prefer e-screening compared with women in the paper-based screening group. However, none of the independent variables, including demographics, mental health history, mode of screening, or obstetric history, were significantly associated with women's ability to be honest during screening.

Primary Outcomes

Acceptability of E-Screening

More women in the e-screening group favored the use of the tablet for mental health screening and indicated a preference for e-screening over paper than women in the paper-based screening group. This is consistent with our previous studies, in which 86% of pregnant women surveyed indicated that they would be very or somewhat comfortable answering questions on a computer or iPad [61].

It is also interesting that more women who were randomized to the e-screening group responded with a more defined preference for e-screening (eg, higher proportion agreed, lower proportion reported that it did not matter one way or the other), compared with women in the paper-based screening group. This finding may relate to the benefit that those in the e-screening group experienced, compared with women in the paper-based screening group who anticipated their responses to e-screening. This finding is important from an implementation perspective when considering women's initial responses to e-screening. While women who have never completed e-screening may initially be ambivalent, our evidence suggests that the actual experience may prove to be an easier, more comfortable, and more private experience than anticipated. This information (eg, once women experience e-screening, they tend to "like it") may also be valuable to share with women who express ambivalence about e-screening.

Feasibility of E-Screening

Disclosure

The vast majority of women in both groups reported that they were able to tell the truth on all questions about emotional health. We were interested in this aspect of screening because previous qualitative studies of paper-based and face-to-face screening found that some pregnant [62] and postpartum [62,63] women purposefully limited their disclosure of current or previous mental health concerns during the screening process. We wanted to quantify the magnitude of this issue, given that the ability of screening tools to accurately identify women with potential mental health problems depends on women being honest in the first place about their status. Our trial's finding is consistent with a previous cross-sectional study we conducted (N=460) in which 79% of pregnant women indicated they could be completely honest if their prenatal care provider asked them about their mental health [20]. Multivariable analysis in that study also showed that the level of honesty (completely vs somewhat or not at all) women anticipated they would have during screening did not vary depending on whether questions were asked on paper or tablet or computer [20]. We found no other studies that assessed disclosure of mental health issues in pregnancy during e-screening. Our findings align with Renker and Tonkin's conclusions that postpartum women were able to disclose interpersonal violence during a computer-administered interpersonal screening assessment [47].

Features of Screening

All of the features of e-screening that we assessed were viewed more favorably by the e-screening group, suggesting that these women felt that e-screening was a superior approach to screening than paper-based screening. Significantly more women in the e-screening group perceived that screening was not too time-consuming and that e-screening had the benefits of being more private than paper-based screening. These findings are similar to one study of computerized violence screening where postpartum women reported that the features of computer screening made it easier for them to answer questions about violence, compared with face-to-face or written approaches [47]. Fewer women in the e-screening group also indicated that they found e-screening impersonal, which is positive given that, anecdotally, clinicians have expressed concerns that women might find e-screening too impersonal to divulge clinically important information.

We found no other studies with which to compare our findings regarding pregnant women's (or others') perceptions of the features of mental health e-screening. However, women's views on these features of e-screening do not appear to influence its clinical benefit. Although women in the e-screening group found e-screening to have more optimal features than a paper-based approach, both groups indicated they could disclose their mental health concerns, regardless of mode of screening. From another perspective, however, women's perceptions of the benefits of e-screening may impact uptake of screening and therefore screening rates. For instance, if women perceive e-screening to be a personal, private experience of easy-to-manage questions, they may be more willing to consent to screening as part of routine prenatal care. Certainly, issues related to the process of screening have been previously identified as deterrents to screening. For example, in a qualitative study examining the acceptability of face-to-face postnatal depression screening, some women described the lack of privacy during screening as inhibiting and inappropriate, while others described the face-to-face approach as intrusive. In addition, e-screening may eliminate some of the barriers to screening engagement that women have identified regarding poor attitudes and negative judgment of the health care professional conducting the screening [62,63].

Comparison With Face-to-Face Screening

In this study, less than 10.0% (24/302 or 30/325) of women in each group indicated that it was easier to answer questions on emotional health when asked face-to-face, in this case by a nurse (compared with on a tablet), more than half in each group stated a preference for e-screening over nurse-led screening, and less than 5.0% (11/302 or 16/324) reported that they would not like e-screening because it would feel impersonal. Taken together, these findings suggest that women prefer self-report methods over face-to-face approaches for mental health screening. A previous cross-sectional study that we conducted in a sample of 460 pregnant women recruited across the province of Alberta (Canada) showed similar findings in that while more women were comfortable with a variety of self-report options (paper, e-screening, completion at home and sending to clinic, completion in clinic waiting room), the mode of screening that garnered the lowest level of comfort was being called at home by a nurse [61]. Findings of this trial are also consistent with a cross-sectional study conducted by Renker and Tonkin, which reported that, in general, postpartum women "overwhelmingly supported" computer-administered interpersonal screening assessment over face-to-face and written approaches [47].

The qualitative study by Rollans et al [62] reported that, during the process of mental health screening by a midwife, some pregnant women found the midwives' responses when they disclosed sensitive information distressing. At times, it was the response of the midwife or nurse to a woman's disclosure that caused her the most distress. This study also found that women's perceptions of the midwives' approach influenced their level of comfort with screening, especially when they felt like they were "being watched." In light of the fact that women did not view e-screening as impersonal, and perceived a potential risk with face-to-face screening, e-screening may offer a more acceptable approach to screening.

Secondary Outcomes

Factors Associated With Women's Preferences for E-Screening

The finding that none of the characteristics we assessed, including demographics, mental health history, and obstetric history, were significantly related to women's preferences for e-screening suggests that most women would accept being screened using either mode. The only factor significantly associated with women's preference for e-screening was being in the e-screening group. This result suggests that women who actually experienced e-screening were more likely to prefer it. This should offer some reassurance in e-screening implementation in the clinical setting in that it indicates that once women use e-screening they tend to favor it.

Factors Associated With Women's Disclosure During Screening

No factors we assessed were significantly associated with women's ability to be honest during screening, including mode of screening (e-screening vs paper-based screening). As such, we did not identify any subgroups of women who might be less apt to disclose their mental health concerns during screening. This finding suggests that women across all ages, income, and education, those with a mental health history, and those with obstetric complications would be comfortable with being honest during mental health screening. The lack of association between mode of screening and disclosure indicates that both paper-based and e-screening approaches to screening facilitate disclosure. These findings are similar to those of our cross-sectional study, in which we also found that neither e-screening nor paper-based screening were related to disclosure [20]. This is positive, in that it suggests that clinical settings can select whichever mode best suits their service delivery model without hampering disclosure.

Limitations

In order to limit participant burden, we evaluated the e-screening and paper-based versions of the ALPHA and the EPDS. Although these 2 screening tools are widely used in perinatal clinical settings, evaluation of other tools is warranted (eg, Generalized Anxiety Disorder-7 scale, Whooley Questions, Antenatal Risk Questionnaire). Overall, women in our study tended to be well-educated, partnered, and affluent, which may limit the generalizability of the findings. However, in our multivariable analyses, these factors were not important influences in women's perceptions of the acceptability and feasibility of mode of mental health screening. We also excluded women who did not speak English as the first step to trial e-screening. Indeed, one of the most important applications of e-screening may be the feature of having immigrant women answer questions in their preferred language and devising a computer-based algorithm to assess the scale score and interpret it automatically for the English-speaking provider. Future studies should evaluate this application of e-screening.

Conclusions

The final participation rate for the trial was 94.2% (636/675), suggesting that most women were eager to participate in a trial

of screening for mental health concerns. This trial's findings support the feasibility and acceptability of e-screening among pregnant women, suggesting that it is a viable service delivery option for mental health screening in busy, primary care settings. As an implementation consideration, clinic and hospital staff

would require minimal training to support women in accessing the Web-based screening link on the tablet computer. Future studies should evaluate the effectiveness of e-screening on clinical outcomes, including follow-up assessment, linkage to services, and reduction of risk of mental health disorders.

Acknowledgments

DK holds the Lois Hole Hospital for Women Cross-Provincial Chair in Perinatal Mental Health and a Canadian Institutes of Health Research (CIHR) New Investigator Award. M-PA holds the St. John of God Chair of Perinatal and Women's Mental Health. RG is supported by a National Health and Medical Research Council Career Development Fellowship. SDM is supported by a Tier 2 Canada Research Chair. AB holds the Ada Slaight and Slaight Family Directorship in Maternity Care. We are grateful for the funding provided by the CIHR for this study. The funder had no role in the writing, review, or approval of the manuscript for publication.

Authors' Contributions

DK, M-PA, SVZ, and AB designed the study. DK, M-PA, SVZ, SDM, GM, GL, WS, and AB reviewed and contributed to the final grant. DK and SVZ managed the trial, while DK, SVZ, M-PA, and LV monitored trial processes. DK, RG, AB, and LV designed, conducted, and interpreted the analysis. DK led the writing of the manuscript. All authors were involved in interpretation of findings and the critical review of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot of the Edinburgh Postnatal Depression Scale.

[PDF File (Adobe PDF File), 951KB - [jmir_v19i4e88_app1.pdf](#)]

References

1. Andersson L, Sundström-Poromaa I, Wulff M, Åström M, Bixo M. Depression and anxiety during pregnancy and six months postpartum: a follow-up study. *Acta Obstet Gynecol Scand* 2006;85(8):937-944. [doi: [10.1080/00016340600697652](#)] [Medline: [16862471](#)]
2. Kingston D, Heaman M, Fell D, Dzakpasu S, Chalmers B. Factors associated with perceived stress and stressful life events in pregnant women: findings from the Canadian Maternity Experiences Survey. *Matern Child Health J* 2012 Jan;16(1):158-168. [doi: [10.1007/s10995-010-0732-2](#)] [Medline: [21165763](#)]
3. Priest SR, Austin M, Barnett BB, Buist A. A psychosocial risk assessment model (PRAM) for use with pregnant and postpartum women in primary care settings. *Arch Womens Ment Health* 2008 Dec;11(5-6):307-317. [doi: [10.1007/s00737-008-0028-3](#)] [Medline: [18726142](#)]
4. Austin M, Kildea S, Sullivan E. Maternal mortality and psychiatric morbidity in the perinatal period: challenges and opportunities for prevention in the Australian setting. *Med J Aust* 2007 Apr 02;186(7):364-367. [Medline: [17407434](#)]
5. Grant K, McMahon C, Austin M. Maternal anxiety during the transition to parenthood: a prospective study. *J Affect Disord* 2008 May;108(1-2):101-111. [doi: [10.1016/j.jad.2007.10.002](#)] [Medline: [18001841](#)]
6. Horwitz SM, Briggs-Gowan MJ, Storfer-Isser A, Carter AS. Persistence of maternal depressive symptoms throughout the early years of childhood. *J Womens Health (Larchmt)* 2009 May;18(5):637-645 [FREE Full text] [doi: [10.1089/jwh.2008.1229](#)] [Medline: [19445615](#)]
7. Mayberry LJ, Horowitz JA, Declercq E. Depression symptom prevalence and demographic risk factors among U.S. women during the first 2 years postpartum. *J Obstet Gynecol Neonatal Nurs* 2007;36(6):542-549. [doi: [10.1111/j.1552-6909.2007.00191.x](#)] [Medline: [17973697](#)]
8. Kingston D. Policywise. 2011. The effects of prenatal and postpartum maternal psychological distress on child development: A systematic review URL: <https://policywise.com/wp-content/uploads/resources/2016/07/TheEffectsofPrenatalandPostpartumMaternalPsychologicalDistressonChildDevelopmentASystematicReviewpdf-Kingston-Dawn.pdf>
9. van der Waerden J, Galéra C, Larroque B, Saurel-Cubizolles M, Sutter-Dallay A, Melchior M, EDEN Mother-Child Cohort Study Group. Maternal Depression Trajectories and Children's Behavior at Age 5 Years. *J Pediatr* 2015 Jun;166(6):1440-8.e1. [doi: [10.1016/j.jpeds.2015.03.002](#)] [Medline: [25866387](#)]

10. Giallo R, Woolhouse H, Gartland D, Hiscock H, Brown S. The emotional-behavioural functioning of children exposed to maternal depressive symptoms across pregnancy and early childhood: a prospective Australian pregnancy cohort study. *Eur Child Adolesc Psychiatry* 2015 Oct;24(10):1233-1244. [doi: [10.1007/s00787-014-0672-2](https://doi.org/10.1007/s00787-014-0672-2)] [Medline: [25572869](https://pubmed.ncbi.nlm.nih.gov/25572869/)]
11. Kingston D, Tough S, Whitfield H. Prenatal and postpartum maternal psychological distress and infant development: a systematic review. *Child Psychiatry Hum Dev* 2012 Oct;43(5):683-714. [doi: [10.1007/s10578-012-0291-4](https://doi.org/10.1007/s10578-012-0291-4)] [Medline: [22407278](https://pubmed.ncbi.nlm.nih.gov/22407278/)]
12. Coates AO, Schaefer CA, Alexander JL. Detection of postpartum depression and anxiety in a large health plan. *J Behav Health Serv Res* 2004;31(2):117-133. [Medline: [15255221](https://pubmed.ncbi.nlm.nih.gov/15255221/)]
13. Spitzer RL, Williams JB, Kroenke K, Hornyak R, McMurray J. Validity and utility of the PRIME-MD patient health questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *Am J Obstet Gynecol* 2000 Sep;183(3):759-769. [Medline: [10992206](https://pubmed.ncbi.nlm.nih.gov/10992206/)]
14. Dennis C, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. *Birth* 2006 Dec;33(4):323-331. [doi: [10.1111/j.1523-536X.2006.00130.x](https://doi.org/10.1111/j.1523-536X.2006.00130.x)] [Medline: [17150072](https://pubmed.ncbi.nlm.nih.gov/17150072/)]
15. Woolhouse H, Brown S, Krastev A, Perlen S, Gunn J. Seeking help for anxiety and depression after childbirth: results of the Maternal Health Study. *Arch Womens Ment Health* 2009 Apr;12(2):75-83. [doi: [10.1007/s00737-009-0049-6](https://doi.org/10.1007/s00737-009-0049-6)] [Medline: [19214705](https://pubmed.ncbi.nlm.nih.gov/19214705/)]
16. Sword W, Busser D, Ganann R, McMillan T, Swinton M. Women's care-seeking experiences after referral for postpartum depression. *Qual Health Res* 2008 Sep;18(9):1161-1173. [doi: [10.1177/1049732308321736](https://doi.org/10.1177/1049732308321736)] [Medline: [18689530](https://pubmed.ncbi.nlm.nih.gov/18689530/)]
17. Flynn HA, Henshaw E, O'Mahen H, Forman J. Patient perspectives on improving the depression referral processes in obstetrics settings: a qualitative study. *Gen Hosp Psychiatry* 2010;32(1):9-16 [FREE Full text] [doi: [10.1016/j.genhosppsych.2009.07.005](https://doi.org/10.1016/j.genhosppsych.2009.07.005)] [Medline: [20114123](https://pubmed.ncbi.nlm.nih.gov/20114123/)]
18. Reay R, Matthey S, Ellwood D, Scott M. Long-term outcomes of participants in a perinatal depression early detection program. *J Affect Disord* 2011 Mar;129(1-3):94-103. [doi: [10.1016/j.jad.2010.07.035](https://doi.org/10.1016/j.jad.2010.07.035)] [Medline: [20800898](https://pubmed.ncbi.nlm.nih.gov/20800898/)]
19. Kingston D, Austin M, Heaman M, McDonald S, Lasiuk G, Sword W, et al. Barriers and facilitators of mental health screening in pregnancy. *J Affect Disord* 2015 Nov 1;186:350-357. [doi: [10.1016/j.jad.2015.06.029](https://doi.org/10.1016/j.jad.2015.06.029)] [Medline: [26281038](https://pubmed.ncbi.nlm.nih.gov/26281038/)]
20. Kingston DE, Biringer A, Toosi A, Heaman MI, Lasiuk GC, McDonald SW, et al. Disclosure during prenatal mental health screening. *J Affect Disord* 2015 Nov 1;186:90-94. [doi: [10.1016/j.jad.2015.07.003](https://doi.org/10.1016/j.jad.2015.07.003)] [Medline: [26233318](https://pubmed.ncbi.nlm.nih.gov/26233318/)]
21. Carroll JC, Reid AJ, Biringer A, Midmer D, Glazier RH, Wilson L, et al. Effectiveness of the Antenatal Psychosocial Health Assessment (ALPHA) form in detecting psychosocial concerns: a randomized controlled trial. *CMAJ* 2005 Aug 2;173(3):253-259 [FREE Full text] [doi: [10.1503/cmaj.1040610](https://doi.org/10.1503/cmaj.1040610)] [Medline: [16076821](https://pubmed.ncbi.nlm.nih.gov/16076821/)]
22. Mitchell AJ, Coyne J. Screening for postnatal depression: barriers to success. *BJOG* 2009 Jan;116(1):11-14 [FREE Full text] [doi: [10.1111/j.1471-0528.2008.01834.x](https://doi.org/10.1111/j.1471-0528.2008.01834.x)] [Medline: [19016688](https://pubmed.ncbi.nlm.nih.gov/19016688/)]
23. Venkatesh KK, Nadel H, Blewett D, Freeman MP, Kaimal AJ, Riley LE. Implementation of universal screening for depression during pregnancy: feasibility and impact on obstetric care. *Am J Obstet Gynecol* 2016 Oct;215(4):517.e1-517.e8. [doi: [10.1016/j.ajog.2016.05.024](https://doi.org/10.1016/j.ajog.2016.05.024)] [Medline: [27210067](https://pubmed.ncbi.nlm.nih.gov/27210067/)]
24. Yawn BP, Dietrich AJ, Wollan P, Bertram S, Graham D, Huff J, et al. TRIPPD: a practice-based network effectiveness study of postpartum depression screening and management. *Ann Fam Med* 2012;10(4):320-329 [FREE Full text] [doi: [10.1370/afm.1418](https://doi.org/10.1370/afm.1418)] [Medline: [22778120](https://pubmed.ncbi.nlm.nih.gov/22778120/)]
25. NICE. 2014. Antenatal and postnatal mental health: clinical management and service guidance URL: <https://www.nice.org.uk/guidance/cg192>
26. Cope.org. 2011 Feb -1. Clinical practice guidelines for depression and related disorders - anxiety, bipolar disorder and puerperal psychosis - in the perinatal period. A guideline for primary care health professionals URL: <http://cope.org.au/wp-content/uploads/2013/12/Perinatal-Mental-Health-Clinical-Practice-Guidelines.pdf>
27. O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary care screening for and treatment of depression in pregnant and postpartum women: evidence report and systematic review for the US preventive services task force. *J Am Med Assoc* 2016 Jan 26;315(4):388-406. [doi: [10.1001/jama.2015.18948](https://doi.org/10.1001/jama.2015.18948)] [Medline: [26813212](https://pubmed.ncbi.nlm.nih.gov/26813212/)]
28. Siu AL, US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, et al. Screening for depression in adults: US preventive services task force recommendation statement. *J Am Med Assoc* 2016 Jan 26;315(4):380-387. [doi: [10.1001/jama.2015.18392](https://doi.org/10.1001/jama.2015.18392)] [Medline: [26813211](https://pubmed.ncbi.nlm.nih.gov/26813211/)]
29. Austin M, Colton J, Priest S, Reilly N, Hadzi-Pavlovic D. The antenatal risk questionnaire (ANRQ): acceptability and use for psychosocial risk assessment in the maternity setting. *Women Birth* 2013 Mar;26(1):17-25. [doi: [10.1016/j.wombi.2011.06.002](https://doi.org/10.1016/j.wombi.2011.06.002)] [Medline: [21764399](https://pubmed.ncbi.nlm.nih.gov/21764399/)]
30. Buist A, Condon J, Brooks J, Speelman C, Milgrom J, Hayes B, et al. Acceptability of routine screening for perinatal depression. *J Affect Disord* 2006 Jul;93(1-3):233-237. [doi: [10.1016/j.jad.2006.02.019](https://doi.org/10.1016/j.jad.2006.02.019)] [Medline: [16647761](https://pubmed.ncbi.nlm.nih.gov/16647761/)]
31. Matthey S, White T, Phillips J, Taouk R, Chee TT, Barnett B. Acceptability of routine antenatal psychosocial assessments to women from English and non-English speaking backgrounds. *Arch Womens Ment Health* 2005 Sep;8(3):171-180. [doi: [10.1007/s00737-005-0076-x](https://doi.org/10.1007/s00737-005-0076-x)] [Medline: [15915317](https://pubmed.ncbi.nlm.nih.gov/15915317/)]

32. Buist A, Ellwood D, Brooks J, Milgrom J, Hayes BA, Sved-Williams A, et al. National program for depression associated with childbirth: the Australian experience. *Best Pract Res Clin Obstet Gynaecol* 2007 Apr;21(2):193-206. [doi: [10.1016/j.bpobgyn.2006.11.003](https://doi.org/10.1016/j.bpobgyn.2006.11.003)] [Medline: [17175198](https://pubmed.ncbi.nlm.nih.gov/17175198/)]
33. Chew-Graham C, Chamberlain E, Turner K, Folkes L, Caulfield L, Sharp D. GPs' and health visitors' views on the diagnosis and management of postnatal depression: a qualitative study. *Br J Gen Pract* 2008 Mar;58(548):169-176 [[FREE Full text](#)] [Medline: [18399021](https://pubmed.ncbi.nlm.nih.gov/18399021/)]
34. Leiferman JA, Dauber SE, Heisler K, Paulson JF. Primary care physicians' beliefs and practices toward maternal depression. *J Womens Health (Larchmt)* 2008 Sep;17(7):1143-1150. [doi: [10.1089/jwh.2007.0543](https://doi.org/10.1089/jwh.2007.0543)] [Medline: [18657043](https://pubmed.ncbi.nlm.nih.gov/18657043/)]
35. Reid AJ, Biringier A, Carroll JD, Midmer D, Wilson LM, Chalmers B, et al. Using the ALPHA form in practice to assess antenatal psychosocial health. *Antenatal Psychosocial Health Assessment*. *CMAJ* 1998 Sep 22;159(6):677-684 [[FREE Full text](#)] [Medline: [9780969](https://pubmed.ncbi.nlm.nih.gov/9780969/)]
36. Coleman VH, Carter MM, Morgan MA, Schulkin J. Obstetrician-gynecologists' screening patterns for anxiety during pregnancy. *Depress Anxiety* 2008;25(2):114-123. [doi: [10.1002/da.20278](https://doi.org/10.1002/da.20278)] [Medline: [17326098](https://pubmed.ncbi.nlm.nih.gov/17326098/)]
37. Bowen A, Bowen R, Butt P, Rahman K, Muhajarine N. Patterns of depression and treatment in pregnant and postpartum women. *Can J Psychiatry* 2012 Mar;57(3):161-167. [Medline: [22398002](https://pubmed.ncbi.nlm.nih.gov/22398002/)]
38. Chew-Graham CA, Sharp D, Chamberlain E, Folkes L, Turner KM. Disclosure of symptoms of postnatal depression, the perspectives of health professionals and women: a qualitative study. *BMC Fam Pract* 2009;10:7 [[FREE Full text](#)] [doi: [10.1186/1471-2296-10-7](https://doi.org/10.1186/1471-2296-10-7)] [Medline: [19159478](https://pubmed.ncbi.nlm.nih.gov/19159478/)]
39. Kim JJ, La Porte LM, Adams MG, Gordon TE, Kuendig JM, Silver RK. Obstetric care provider engagement in a perinatal depression screening program. *Arch Womens Ment Health* 2009 Jun;12(3):167-172. [doi: [10.1007/s00737-009-0057-6](https://doi.org/10.1007/s00737-009-0057-6)] [Medline: [19277845](https://pubmed.ncbi.nlm.nih.gov/19277845/)]
40. Leddy M, Haaga D, Gray J, Schulkin J. Postpartum mental health screening and diagnosis by obstetrician-gynecologists. *J Psychosom Obstet Gynaecol* 2011 Mar;32(1):27-34. [doi: [10.3109/0167482X.2010.547639](https://doi.org/10.3109/0167482X.2010.547639)] [Medline: [21261561](https://pubmed.ncbi.nlm.nih.gov/21261561/)]
41. Swalm D, Brooks J, Doherty D, Nathan E, Jacques A. Using the Edinburgh postnatal depression scale to screen for perinatal anxiety. *Arch Womens Ment Health* 2010 Dec;13(6):515-522. [doi: [10.1007/s00737-010-0170-6](https://doi.org/10.1007/s00737-010-0170-6)] [Medline: [20574749](https://pubmed.ncbi.nlm.nih.gov/20574749/)]
42. Austin M, Marcé Society Position Statement Advisory Committee. Marcé International Society position statement on psychosocial assessment and depression screening in perinatal women. *Best Pract Res Clin Obstet Gynaecol* 2014 Jan;28(1):179-187. [doi: [10.1016/j.bpobgyn.2013.08.016](https://doi.org/10.1016/j.bpobgyn.2013.08.016)] [Medline: [24138943](https://pubmed.ncbi.nlm.nih.gov/24138943/)]
43. Choo EK, Ranney ML, Aggarwal N, Boudreaux ED. A systematic review of emergency department technology-based behavioral health interventions. *Acad Emerg Med* 2012 Mar;19(3):318-328. [doi: [10.1111/j.1553-2712.2012.01299.x](https://doi.org/10.1111/j.1553-2712.2012.01299.x)] [Medline: [22435865](https://pubmed.ncbi.nlm.nih.gov/22435865/)]
44. Renker PR. Breaking the barriers: the promise of computer-assisted screening for intimate partner violence. *J Midwifery Womens Health* 2008;53(6):496-503. [doi: [10.1016/j.jmwh.2008.07.017](https://doi.org/10.1016/j.jmwh.2008.07.017)] [Medline: [18984505](https://pubmed.ncbi.nlm.nih.gov/18984505/)]
45. MacMillan HL, Wathen CN, Jamieson E, Boyle M, McNutt L, Worster A, McMaster Violence Against Women Research Group. Approaches to screening for intimate partner violence in health care settings: a randomized trial. *J Am Med Assoc* 2006 Aug 2;296(5):530-536. [doi: [10.1001/jama.296.5.530](https://doi.org/10.1001/jama.296.5.530)] [Medline: [16882959](https://pubmed.ncbi.nlm.nih.gov/16882959/)]
46. Renker PR, Tonkin P. Women's views of prenatal violence screening: acceptability and confidentiality issues. *Obstet Gynecol* 2006 Feb;107(2 Pt 1):348-354. [doi: [10.1097/01.AOG.0000195356.90589.c5](https://doi.org/10.1097/01.AOG.0000195356.90589.c5)] [Medline: [16449123](https://pubmed.ncbi.nlm.nih.gov/16449123/)]
47. Renker P, Tonkin P. Postpartum women's evaluations of an audio/video computer-assisted perinatal violence screen. *Comput Inform Nurs* 2007;25(3):139-147. [doi: [10.1097/01.NCN.0000270040.14541.37;00024665-200705000-00008](https://doi.org/10.1097/01.NCN.0000270040.14541.37;00024665-200705000-00008)] [Medline: [17496478](https://pubmed.ncbi.nlm.nih.gov/17496478/)]
48. Le HN, Perry DF, Sheng X. Using the internet to screen for postpartum depression. *Matern Child Health J* 2009 Mar;13(2):213-221. [doi: [10.1007/s10995-008-0322-8](https://doi.org/10.1007/s10995-008-0322-8)] [Medline: [18278545](https://pubmed.ncbi.nlm.nih.gov/18278545/)]
49. Locke SE, Kowaloff HB, Hoff RG, Safran C, Popovsky MA, Cotton DJ, et al. Computer interview for screening blood donors for risk of HIV transmission. *MD Comput* 1994;11(1):26-32. [Medline: [8145632](https://pubmed.ncbi.nlm.nih.gov/8145632/)]
50. Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 1998 May 8;280(5365):867-873 [[FREE Full text](#)] [Medline: [9572724](https://pubmed.ncbi.nlm.nih.gov/9572724/)]
51. Kingston D, McDonald S, Biringier A, Austin M, Hegadoren K, McDonald S, et al. Comparing the feasibility, acceptability, clinical-, and cost-effectiveness of mental health e-screening to paper-based screening on the detection of depression, anxiety, and psychosocial risk in pregnant women: a study protocol of a randomized, parallel-group, superiority trial. *Trials* 2014 Jan 02;15:3 [[FREE Full text](#)] [doi: [10.1186/1745-6215-15-3](https://doi.org/10.1186/1745-6215-15-3)] [Medline: [24383441](https://pubmed.ncbi.nlm.nih.gov/24383441/)]
52. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987 Jun;150:782-786. [Medline: [3651732](https://pubmed.ncbi.nlm.nih.gov/3651732/)]
53. Blackmore ER, Carroll J, Reid A, Biringier A, Glazier RH, Midmer D, et al. The use of the Antenatal Psychosocial Health Assessment (ALPHA) tool in the detection of psychosocial risk factors for postpartum depression: a randomized controlled trial. *J Obstet Gynaecol Can* 2006 Oct;28(10):873-878. [Medline: [17140502](https://pubmed.ncbi.nlm.nih.gov/17140502/)]

54. Chan A, 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](https://doi.org/10.7326/0003-4819-158-3-201302050-00583)] [Medline: [23295957](#)]
55. McDonald SW, Lyon AW, Benzie KM, McNeil DA, Lye SJ, Dolan SM, et al. The All Our Babies pregnancy cohort: design, methods, and participant characteristics. *BMC Pregnancy Childbirth* 2013;13(Suppl 1):S2 [FREE Full text] [doi: [10.1186/1471-2393-13-S1-S2](https://doi.org/10.1186/1471-2393-13-S1-S2)] [Medline: [23445747](#)]
56. Gracie SK, Lyon AW, Kehler HL, Pennell CE, Dolan SM, McNeil DA, et al. All Our Babies Cohort Study: recruitment of a cohort to predict women at risk of preterm birth through the examination of gene expression profiles and the environment. *BMC Pregnancy Childbirth* 2010 Dec 30;10:87 [FREE Full text] [doi: [10.1186/1471-2393-10-87](https://doi.org/10.1186/1471-2393-10-87)] [Medline: [21192811](#)]
57. Phac-aspc. 2009. What Mothers Say: The Canadian Maternity Experiences Survey URL: <http://www.phac-aspc.gc.ca/rhs-ssg/survey-eng.php>
58. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *Br Med J* 2013;346:e7586 [FREE Full text] [Medline: [23303884](#)]
59. Kingston D, Austin MP, Hegadoren K, McDonald S, Lasiuk G, McDonald S, et al. Study protocol for a randomized, controlled, superiority trial comparing the clinical and cost- effectiveness of integrated online mental health assessment-referral-care in pregnancy to usual prenatal care on prenatal and postnatal mental health and infant health and development: the Integrated Maternal Psychosocial Assessment to Care Trial (IMPACT). *Trials* 2014;15:72 [FREE Full text] [doi: [10.1186/1745-6215-15-72](https://doi.org/10.1186/1745-6215-15-72)] [Medline: [24597683](#)]
60. Thabane L, Ma J, Chu R, Cheng J, Ismail A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol* 2010;10:1 [FREE Full text] [doi: [10.1186/1471-2288-10-1](https://doi.org/10.1186/1471-2288-10-1)] [Medline: [20053272](#)]
61. Kingston DE, Biringer A, McDonald SW, Heaman MI, Lasiuk GC, Hegadoren KM, et al. Preferences for mental health screening among pregnant women: a cross-sectional study. *Am J Prev Med* 2015 Oct;49(4):e35-e43. [doi: [10.1016/j.amepre.2015.03.026](https://doi.org/10.1016/j.amepre.2015.03.026)] [Medline: [26143952](#)]
62. Rollans M, Schmied V, Kemp L, Meade T. Digging over that old ground: an Australian perspective of women's experience of psychosocial assessment and depression screening in pregnancy and following birth. *BMC Womens Health* 2013 Apr 09;13:18 [FREE Full text] [doi: [10.1186/1472-6874-13-18](https://doi.org/10.1186/1472-6874-13-18)] [Medline: [23570282](#)]
63. Shakespeare J, Blake F, Garcia J. A qualitative study of the acceptability of routine screening of postnatal women using the Edinburgh Postnatal Depression Scale. *Br J Gen Pract* 2003;53(493):614-619.

Abbreviations

ALPHA: Antenatal Psychosocial Health Assessment
CIHR: Canadian Institutes of Health Research
EPDS: Edinburgh Postnatal Depression Scale
MINI: Mini-International Neuropsychiatric Interview
WCHRI: Women and Children's Health Research Institute

Edited by C Gibbons; submitted 22.10.16; peer-reviewed by N Highet, J Slomian, L Bacchus; comments to author 16.11.16; revised version received 15.01.17; accepted 25.01.17; published 07.04.17.

Please cite as:

Kingston D, Austin MP, Veldhuyzen van Zanten S, Harvalik P, Giallo R, McDonald SD, MacQueen G, Vermeyden L, Lasiuk G, Sword W, Biringer A

Pregnant Women's Views on the Feasibility and Acceptability of Web-Based Mental Health E-Screening Versus Paper-Based Screening: A Randomized Controlled Trial

J Med Internet Res 2017;19(4):e88

URL: <http://www.jmir.org/2017/4/e88/>

doi: [10.2196/jmir.6866](https://doi.org/10.2196/jmir.6866)

PMID: [28389421](https://pubmed.ncbi.nlm.nih.gov/28389421/)

©Dawn Kingston, Marie-Paule Austin, Sander Veldhuyzen van Zanten, Paula Harvalik, Rebecca Giallo, Sarah D McDonald, Glenda MacQueen, Lydia Vermeyden, Gerri Lasiuk, Wendy Sword, Anne Biringer. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 07.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Seeking Web-Based Information About Attention Deficit Hyperactivity Disorder: Where, What, and When

Sara Rosenblum^{1*}, PhD; Elad Yom-Tov^{2*}, PhD

¹Laboratory of Complex Human Activity and Participation (CHAP), Department of Occupational Therapy, University of Haifa, Haifa, Israel

²Microsoft Research, Herzliya, Israel

* all authors contributed equally

Corresponding Author:

Sara Rosenblum, PhD

Laboratory of Complex Human Activity and Participation (CHAP)

Department of Occupational Therapy

University of Haifa

Eshkol Bldg, 9th Fl

Abba Khoushy Ave 199

Haifa, 3498838

Israel

Phone: 972 4 824 0474

Fax: 972 4 824 9753

Email: rosens@research.haifa.ac.il

Abstract

Background: Attention Deficit Hyperactivity Disorder (ADHD) is a common neurodevelopmental disorder, prevalent among 2-10% of the population.

Objective: The objective of this study was to describe where, what, and when people search online for topics related to ADHD.

Methods: Data were collected from Microsoft's Bing search engine and from the community question and answer site, Yahoo Answers. The questions were analyzed based on keywords and using further statistical methods.

Results: Our results revealed that the Internet indeed constitutes a source of information for people searching the topic of ADHD, and that they search for information mostly about ADHD symptoms. Furthermore, individuals personally affected by the disorder made 2.0 more questions about ADHD compared with others. Questions begin when children reach 2 years of age, with an average age of 5.1 years. Most of the websites searched were not specifically related to ADHD and the timing of searches as well as the query content were different among those prediagnosis compared with postdiagnosis.

Conclusions: The study results shed light on the features of ADHD-related searches. Thus, they may help improve the Internet as a source of reliable information, and promote improved awareness and knowledge about ADHD as well as quality of life for populations dealing with the complex phenomena of ADHD.

(*J Med Internet Res* 2017;19(4):e126) doi:[10.2196/jmir.6579](https://doi.org/10.2196/jmir.6579)

KEYWORDS

attention deficit hyperactivity disorder; Internet; search engine; coping behavior; parents

Introduction

Background

Attention deficit hyperactivity disorder (ADHD) is a common, chronic, and pervasive pediatric neurodevelopmental psychiatric disorder that is prevalent among 2-10% of the population and has a complex etiology [1,2]. In the United States, approximately 5.4 million or 9.8% of children aged between 4 and 17 years are diagnosed with ADHD, with a 21.8% increase rate of

diagnosis between 2003 and 2007 [3]. ADHD is characterized by a developmentally inappropriate performance level, impulsivity, and inability to sustain attention and concentration [4].

Previous literature has indicated the mobility of the real-world daily manifestations of ADHD from early childhood to adulthood as well as gender-related differences [5,6]. ADHD has longstanding neuropsychological impairments and therefore, parents of children suspected or diagnosed with ADHD often

experience daily distress [7-9]. Having a child with ADHD has several implications on family life, and requires parental management of their child's daily behavior. Parents often need to take on the role of case manager or "gatekeepers" and need to manage the home, school, and social behaviors, as well as medications, the daily routine, and therapy sessions, including communication with the therapist, if in treatment [3,10]. Not surprisingly, family members of children with ADHD face daily strains and search for information and solutions to these issues [11]. Thus, it is of interest to identify what, where, and when information is searched for related to ADHD and its management.

ADHD symptoms often appear in early childhood, yet the clinical guidelines of the American Academy of Pediatrics recommend diagnosing ADHD only from the age of 4 [12]. Although parents may feel differences in their child's behavior in comparison to other children from an early age, most children are diagnosed with ADHD only at the age of 7-8 years when their behavior can be validated by school teachers' reports. Consequently, ADHD diagnosis is likely to be influenced by the child's social and school environment as well as the child's external characteristics [13-15].

Therefore, from the child's early years of development up to school years and even during adulthood, parents worldwide may search for resources and solutions related to their child's behavior.

A previous study of the information needs of parents of children with ADHD revealed that parents' information resources were related to pediatricians (86%), books (76%), general practitioners (65%), schools (61%), the Internet (59%), and other forms of media (54%) [16]. Parents were mainly interested in the causes and symptoms at the time of ADHD diagnosis and preferred to receive verbal information from professionals (69%), whereas pediatricians were rated highest as a useful, trusted, easy-to-understand, and up-to-date information source.

Nowadays, the Internet is a popular source of information and medical information is one of the most popular Internet topics thus enabling children's diagnosis or self-diagnosis as well as support at different coping stages [17-20]. However, due to the

limitations of the information supplied online, it was found that online search complements rather than opposes the primary role of the doctor-patient relationship for delivering health and medical information from a professional trusted expert [16,21-23]. The Internet may address people's sense of self-responsibility and their "everyday" needs for health information searches through experience in navigating a multiplicity of Web-based sources [21].

In that context, it was found that parents of children with psychiatric diagnoses respond favorably to the Internet as a source of information [24]. More specifically, parents of children with ADHD found it to be a preferred source of information among other resources, including physicians or educators [16,25]. Accordingly, the extent of parental use of the Internet and the type of information sought after is of interest to health care providers [18].

Aim of the Study

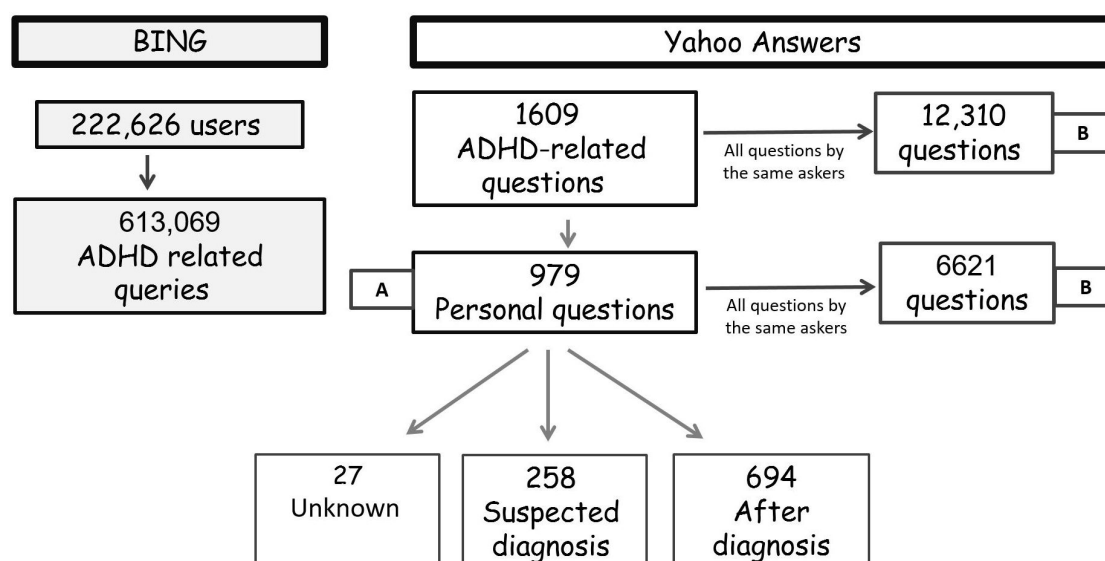
The aim of this study was to explore the dynamic process of Web-based information searches related to ADHD over a 3-month period. The specific goals were to identify *where*, *what*, and *when* information related to ADHD is sought after. Improving insight about people's information-seeking behavior may have implications on improving health care literacy and health outcomes through Web-based solutions for this population.

Methods

Data Collection

Data were collected from two separate sources: (1) queries made in the United States using the Microsoft Bing search engine (following findings [26]) and (2) questions posted on the community question-answering website, Yahoo Answers. The latter site was chosen both for its overall popularity, as well as its popularity specifically among parents of children with suspected or diagnosed ADHD (eg, [27]). The study was approved by the ethics committee of the University of Haifa, Faculty of Social and Health Sciences, approval number 368/15.

Figure 1 represents the information extraction process.

Figure 1. Analysis phases and number of questions analyzed by data source. ADHD: attention deficit hyperactivity disorder.

Bing Queries

All English language queries submitted to the Bing search engine by users of desktop computers in the United States for the 3-month period between January 2015 and March 2015 were extracted. Queries were filtered to maintain only those that contained the terms “ADHD” or “attention deficit” and “disorder.” A total of 613,069 queries from 222,626 users were identified.

For each query, the text of the query (without use of any automated spelling correction), time, and date were extracted, in addition to a list of pages shown to the user and the pages clicked on by the user. In order to maintain privacy, user identifiers, which were based on browser cookies, were anonymized before access by the investigators. Queries were aggregated preceding analysis preventing individual-level user datum from being examined by the researchers.

We note that the market share of Bing in the United States is around 19%, according to recent estimates [28]. The correlation between the number of Bing users per county in the United States and the number of people in that county according to the 2010 US Census is $R^2=.83$ ($P=.001$). Furthermore, an analysis of behavioral data collected from an opt-in consumer panel recruited by an Internet analytics company comScore, which includes age (in 5-year increments) and gender, shows a correlation of .79 ($P=.004$) between the fraction of users at each age and gender between Bing and Google users. Thus, it is estimated that Bing users are a representative sample of the US population.

Yahoo Answers Questions

Questions about ADHD posted on Yahoo Answers, an online community question and answer website, were collected. Bing was used to query questions from Yahoo Answers that contained the term “attention deficit hyperactivity disorder” or the term “ADHD.” This resulted in 1609 unique questions.

The questions extracted from Yahoo were then each labeled by at least three human assessors from the crowdsourcing platform CrowdFlower, who were asked to label the question according to whether the question refers primarily to a person who (1) was diagnosed with ADHD, (2) is suspected of having ADHD, and (3) none of the above.

As presented in Figure 1, we refer to questions with the first label as “diagnosed,” the second as “suspected diagnosis,” and the third as “unknown.” The assessors agreed on a majority of the questions’ labels (79%) of the questions.

The authors examined a sample of the questions and identified four main categories. Two MSc in occupational therapy research assistants further classified the questions into one of the following four groups:

1. Personal account
2. General question about ADHD
3. Question about ADHD among their pets
4. Other

A question was defined as “personal” if it was asked in the first person form and included words such as “I” or “my son” or “daughter” or “we” or “my wife,” which reflects the asker’s personal involvement within the question being asked.

Phase A

A total of 979 personal account type of questions were found and analyzed by the two research assistant raters. Each question was examined to detect any additional information about the person discussed in the question, including the age and gender, whether the person was diagnosed with ADHD or not. Furthermore, the raters analyzed whether the question addressed problems in (1) daily function, related to activities that children do throughout the day, including activities of daily living (such as self-hygiene, dressing, showering, eating, walking), social activities, and play; (2) behavioral manifestations, related to how the child interacts with people in his or her environment at home, school, or at social events; and finally, (3) academic

performance, related to issues pertaining to school or spontaneous learning, based on predetermined keywords and examples (see [Multimedia Appendix 1](#)). The keywords and examples were compiled by the first author based on initial examination and content analysis of the extracted questions. In this process, the main area of the question was recognized and keywords identified. Additionally, an extensive literature review of these areas and keywords was performed and accordingly, the guidelines for classification by the raters were established. In case of disagreement between the two raters, the issues were discussed with the first author and the classification was made according to the final decision of this discussion.

Phase B

All other ADHD-related questions that were posted by anyone in the *personal account* or the *general questions* participant's category were extracted, resulting in 12,310 questions from 1130 users.

Results

Bing Queries

Bing queries were analyzed while focusing on the *what* and *where* queries. Namely, *what* is the main concept that people

use for the search and *where* or in which sites do they search for answers to their inquiries. Initially, the 613,069 queries made on Bing were classified into whether they referred to individual persons if they contained one of the following phrases: "my son," "my daughter," "my child," or "my kid;" approximately 0.4% of the queries contained these phrases, and thus are referred to henceforth as "Individual Persons Query" (IPQ). People, who made at least one IPQ made an average of 5.4 queries, compared with an average of 2.7 queries among people who did not make an IPQ. Accordingly, users who referred to individual persons tended to search for significantly more information about ADHD. The 10 most popular ADHD-related queries submitted to Bing that included specific search terms are presented in [Table 1](#). Of these 10 queries, those referring to similar content were grouped together resulting in 6 query categories as described in [Table 1](#). As a whole, the 10 queries detailed in [Table 1](#) were used by 21.80% of the searchers. Within the queries, the most frequently used term in the question was ADHD (5.71%), whereas 61.5% related to the general queries about (1) ADHD and (2) ADHD and symptoms categories.

Table 1. The 10 most popular queries submitted to Bing.

Query category	Query	Total N=613,069, n (%)
General	ADHD ^a	35,010 (5.71)
	Attention deficit disorder	4616 (0.75)
	What is ADHD?	2960 (0.48)
Symptoms	ADHD symptoms	10,945 (1.79)
	Attention deficit disorder symptoms	2891 (0.47)
Test	ADHD test	2321 (0.37)
Medications	ADHD medications	6228 (6.9)
Children	ADHD in children	4051 (0.66)
Adults	Adult ADHD	3255 (0.53)
	ADHD in adults	3007 (0.49)

^aADHD: attention deficit hyperactivity disorder.

The 20 most popular websites displayed as a result of the ADHD queries, and their rank in website popularity displayed to the users as a result of their search in the Bing search result engine, are presented in [Table 2](#). Also presented is the popularity ranking of the websites in terms of the likelihood of the users to click on these websites when they are displayed in search results

([Table 2](#)). Thus, for example, the site www.cdc.gov is the 6th most popular site displayed in response to an ADHD-related query, but is not among the 20 sites most clicked in response to ADHD-related queries. Both general information websites and those specializing in ADHD are present in both rankings.

Table 2. Most popular websites displayed and clicked on in Bing attention deficit hyperactivity disorder (ADHD) queries.

Website	Rank in display	Rank in click	ADHD ^a -specific	HoN ^b approved
webmd.com	1	1	No	Yes
additudemag.com	2	2	Yes	
psychcentral.com	3	3	No	Yes
wikipedia.org	4	4	No	
helpguide.org	5	7	No	
cdc.gov	6	>20	No	
nimh.nih.gov	7	15	No	
mayoclinic.org	8	10	No	
medicinenet.com	9	>20	No	Yes
healthline.com	10	8	No	Yes
chadd.org	11	18	Yes	
add.org	12	13	Yes	
add.about.com	13	6	Yes	Yes
drugs.com	14	16	No	Yes
kidshealth.org	15	20	No	Yes
ncbi.nlm.nih.gov	16	9	No	Yes
livestrong.com	17	11	No	
answers.yahoo.com	18	12	No	
healthcentral.com	19	17	No	Yes
help4adhd.org	20	19	Yes	

^aADHD: attention deficit hyperactivity disorder.

^bHoN: Health on the Net.

The Health on the Net (HoN) foundation is an organization that certifies websites that provide health-related information and certification if they meet specific reliability standards. Interestingly, although 45% of the 20 popular websites that appear in Table 2 are HoN-certified, only 35% of the clicked websites are certified.

As presented in Table 2, Yahoo Answers is the 18th most popular website displayed to users in response to ADHD queries on Bing and the 12th most popular site clicked by them. However, Yahoo Answers is also the only question-answering site on the list.

People who made at least one IPQ on Bing also differed in the association of their browsing behavior with Yahoo Answers. The average number of Bing queries made by people in the IPQ population who had a Yahoo Answers result displayed to them was 5.7, compared with 5.4 in the remaining population. However, people in the IPQ population who clicked on a Yahoo Answers result made an average of 14.5 queries on Bing compared with 7.2 made by other users in this sample. Thus, people who have a strong personal interest in ADHD tend to search more on Bing and click more on Yahoo Answers results. Consequently, the following sections will focus on questions from Yahoo Answers.

Yahoo Answers Data

Analysis of the questions in Yahoo answers focused on *when* people ask questions concerning *what* trouble them, and whether the main issue that troubles them is associated with the need to ask more questions.

Phase A

In total, 979 personal questions were found in Yahoo Answers in Phase A, as presented in Figure 1. Of these 979 questions, 26.35% (258/979) reported that the child was suspected of having an ADHD diagnosis, 70.89% (694/979) reported that the child was diagnosed, and the remaining did not report this information. From the entire sample, 69.97% (685/979) of questions were about males, whereas only 23.49% (230/979) of questions were about females, and the remaining 6.54% (64/979) of questions did not state the child's gender. The age of children was given in 98.06% (960/979) of questions. The distribution of the reported age of children, stratified by whether they were reported as being of suspected diagnosis or diagnosed is presented in Figure 2.

Moreover, children who were diagnosed with ADHD were significantly older than those suspected of having ADHD. The average age for the latter was 5.1 years (range 0.5-20, SD 2.5), compared with 7.9 years for diagnosed children (range 3-20, SD 2.7; ranksum, $P<.001$).

Figure 3 displays the age distribution of children suspected of ADHD, whereas Figure 4 displays the age distribution of children diagnosed with ADHD, both stratified by reported gender. The average age is 5.1 years for males, who are suspected of having ADHD, and 5.2 years for females (not statistically significant, ranksum test).

The relative frequency of each of the three types of issues reported in the questions (daily function, behavioral, academic), stratified by age is presented in Figure 5 concerning children suspected for ADHD and in Figure 6 concerning children diagnosed as ADHD.

In order to gather more information about the query process, a Cox hazard regression model [29] was constructed to assess the correlation between variables of the original first question and the likelihood of asking additional questions related to ADHD. The variables in the model included the following:

1. demographics: age and gender
2. whether the question discussed functional, behavioral, or academic issues related to ADHD

3. diagnosis status: whether the question discussed a child who was suspected to have ADHD, or already diagnosed with ADHD
4. whether the person inquiring appeared to be emotional, helpless, or positive
5. whether the person inquiring asked for information
6. whether the question mentioned that the child was receiving medication and whether the child had additional diagnoses

For this analysis, only 97.2% (952/979) of questions in which the diagnosis status was known were considered. The statistically significant variables ($P < .05$) were the diagnosis status, meaning whether a query appeared in children suspected of ADHD or after diagnosis (hazard ratio: 0.60, $P = .01$), and whether the asker revealed helplessness (hazard ratio: 1.45, $P = .05$). Thus, additional ADHD-related questions are correlated with the child being before diagnosis (hazard lower than 1) and with the person inquiring feeling helpless (eg, a hazard ratio higher than 1).

Figure 2. Distributions of children's age for the 952 questions where askers describe whether their child is suspected of attention deficit hyperactivity disorder (ADHD) or diagnosed with it, in Yahoo Answers questions.

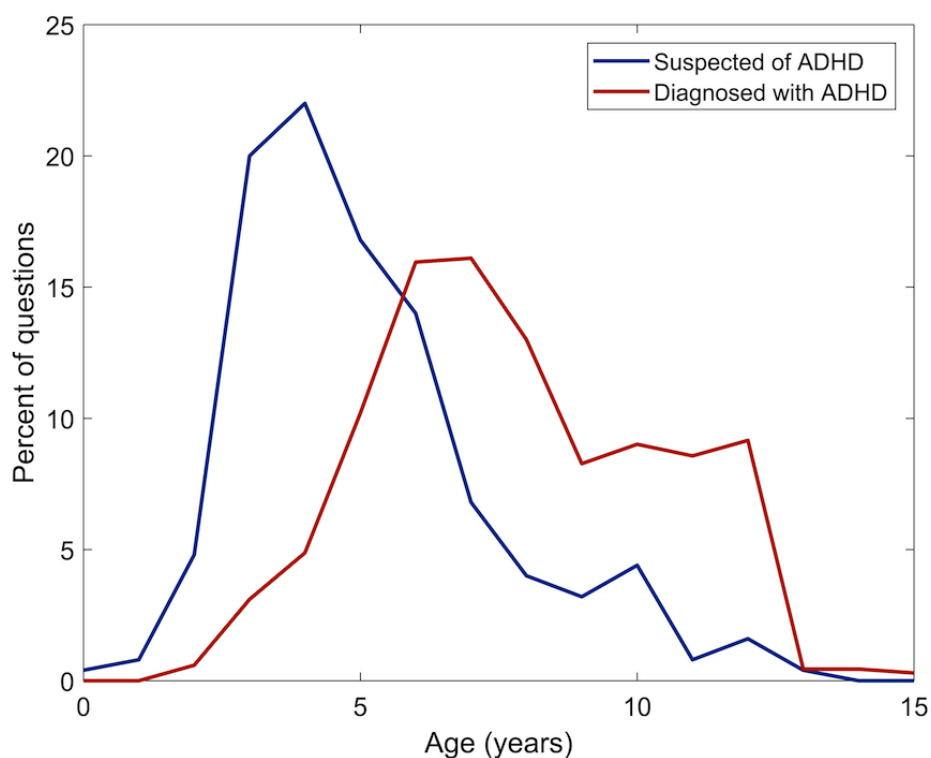


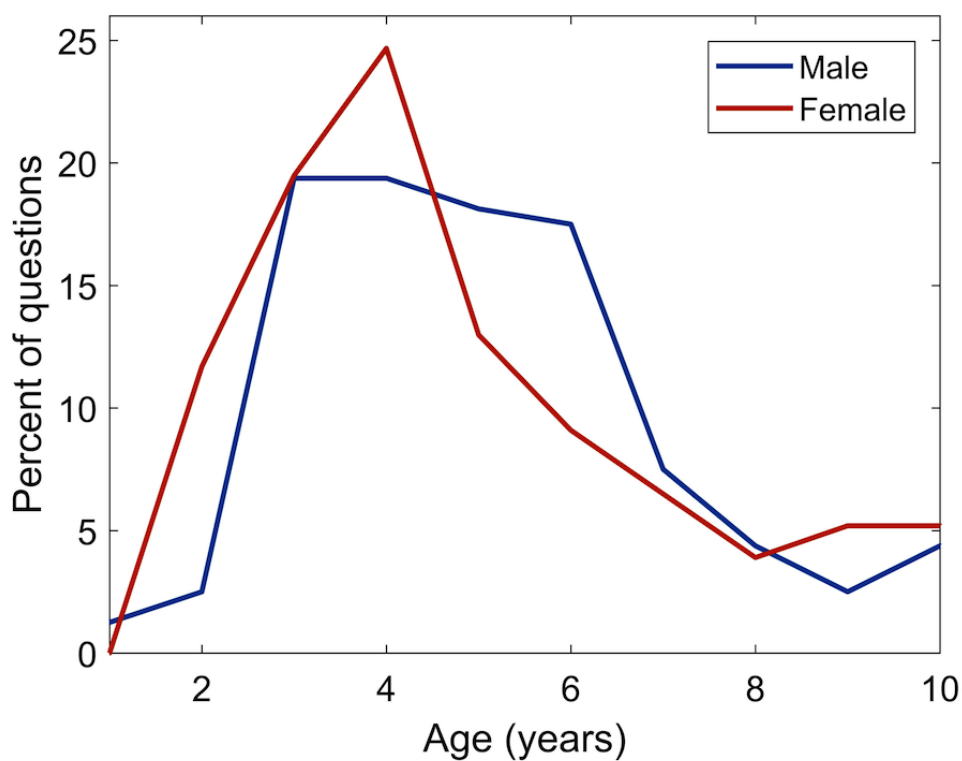
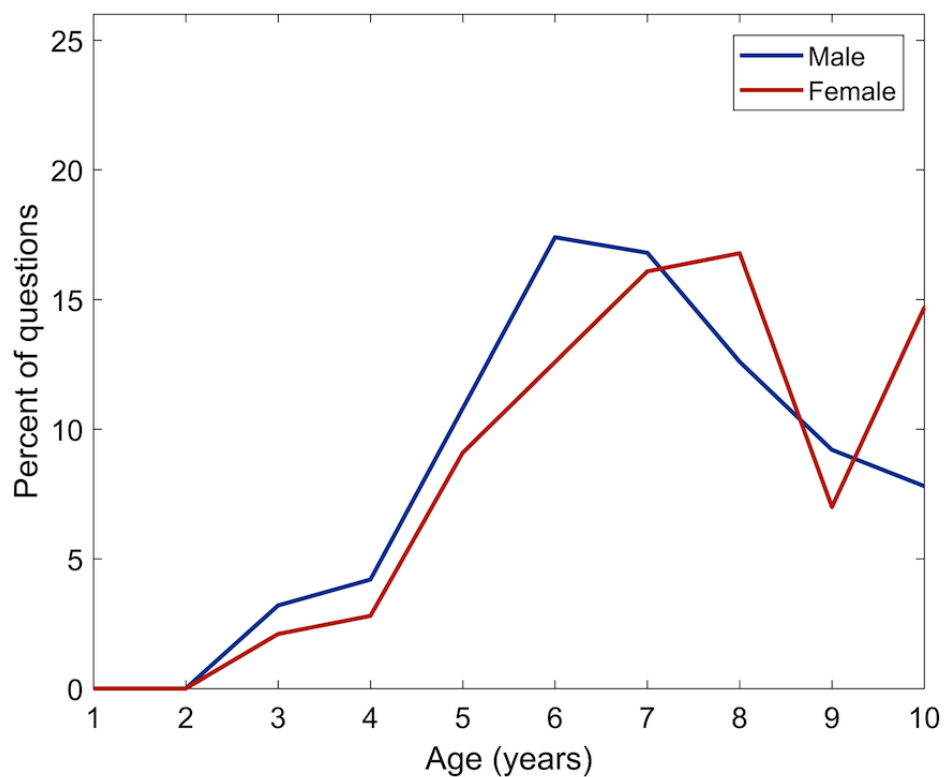
Figure 3. Age of children suspected of having attention deficit hyperactivity disorder (ADHD), stratified by gender.**Figure 4.** Age of children diagnosed with attention deficit hyperactivity disorder (ADHD), stratified by gender.

Figure 5. Frequency of issues reported in children suspected of attention deficit hyperactivity disorder (ADHD), stratified by age.

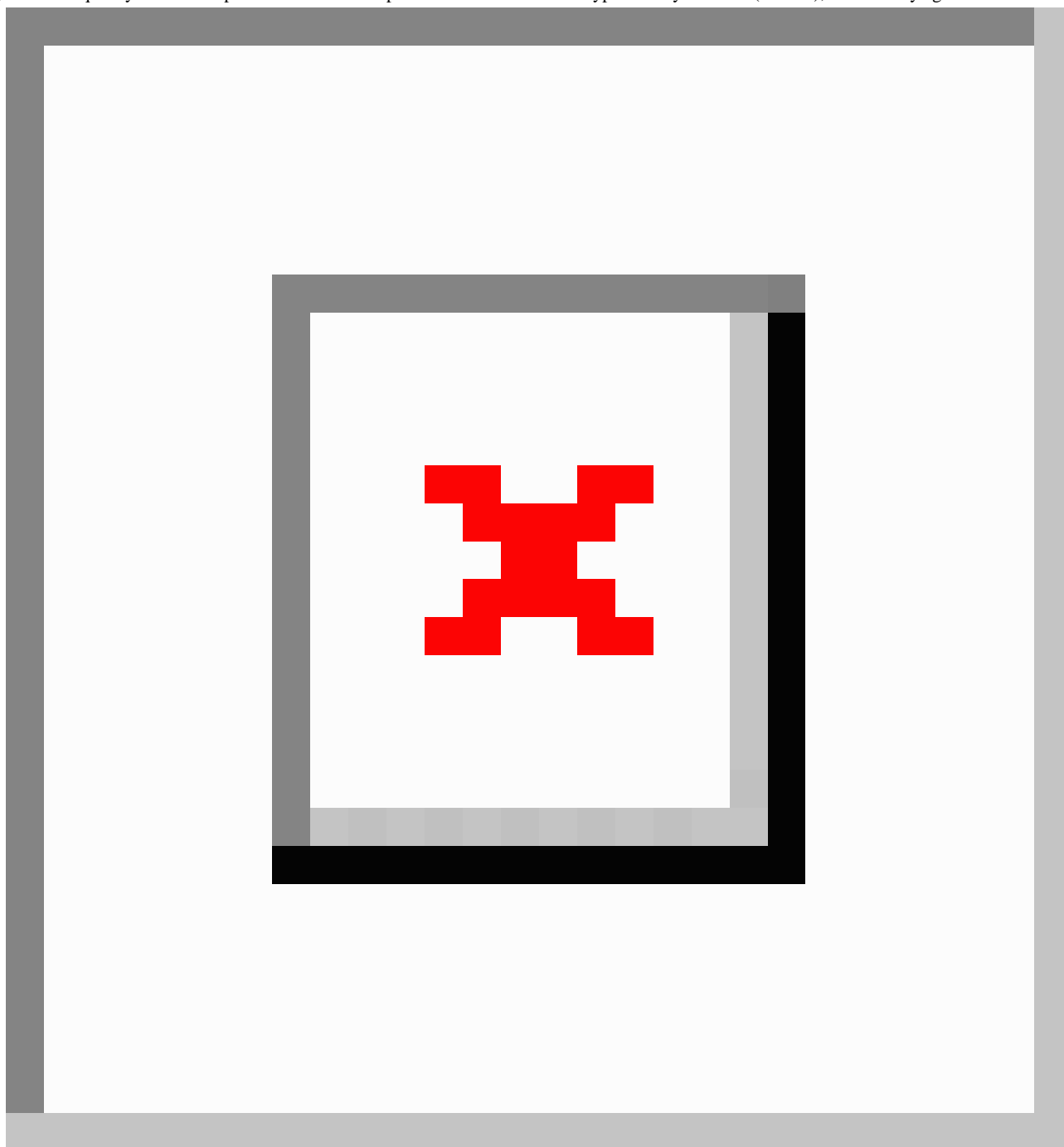
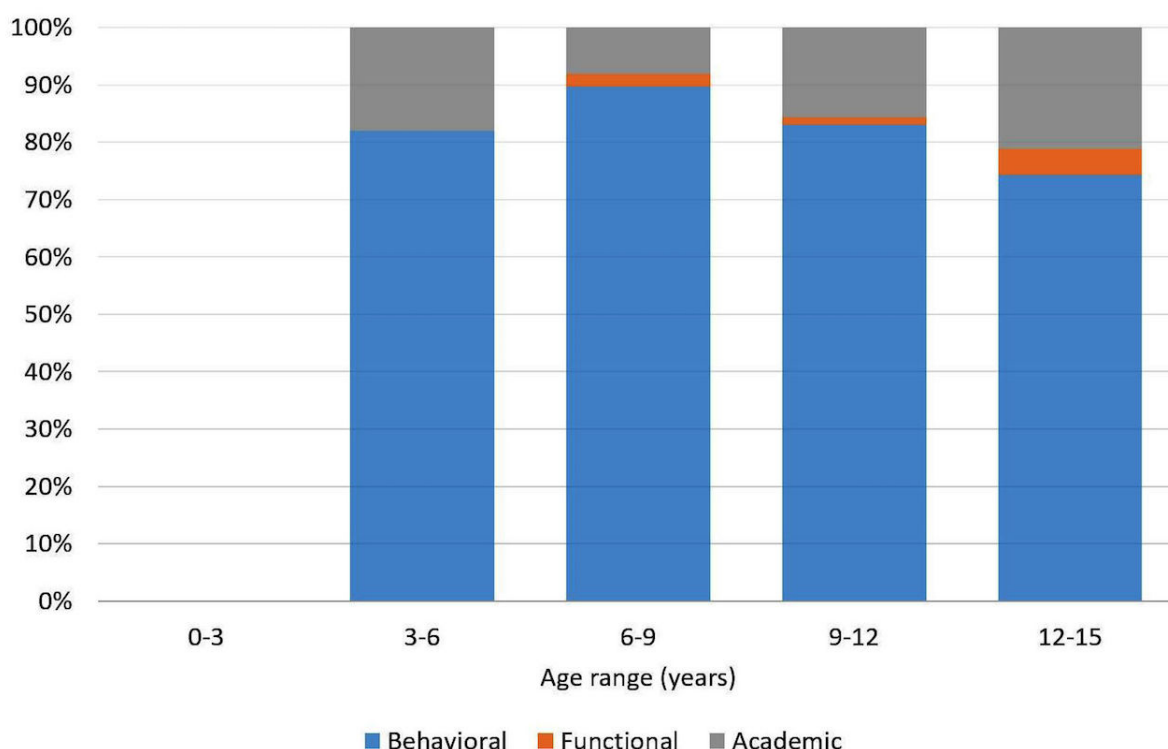


Figure 6. Frequency of issues reported in children diagnosed as attention deficit hyperactivity disorder (ADHD), stratified by age.

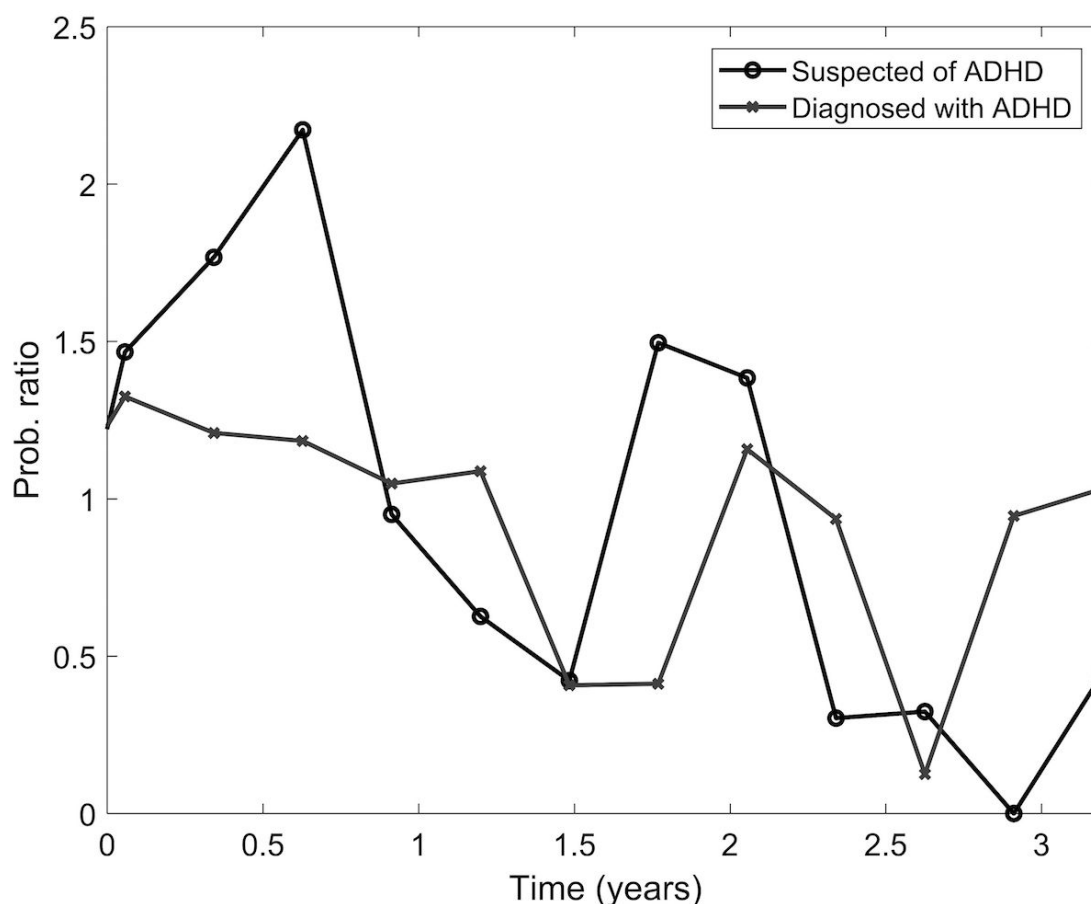
Phase B

In Phase B, 13.07% (1609/12,310) of ADHD-related questions by users in our sample were extracted from all 12,310 Yahoo Answers questions. The question was labeled as ADHD-related if the text contained the word ADHD, the name of a drug used to treat ADHD (adderall, amphetamine, buspirone, clonidine, ethylphenidate, melatonin, methamphetamine, methylphenidate, oxycodone, or rember), or common comorbidities of ADHD (developmental deficit malnutrition, autism, learning disability, dyslexia, bipolar, autistic spectrum disorder, mood disorder, Tourette, Asperger's syndrome, speech impairment, or obsessive compulsive disorder). Of the entire set of questions, 53.79% (6621/12,310) of questions were asked by people in the labeled set of 979 personal account questions. Among the 979 personal questions, there were 13.0% (127/979) of questions labeled as ADHD-related in this set by people who asked a question about a child prediagnosis, and 33.6% (329/979) questions about children postdiagnosis.

Considering all the 6621 questions, the ratio between the probability of asking an ADHD-related question, divided by the probability of asking any type of question, as a function of time is presented in Figure 7.

The questions were stratified by whether they were asked by a user who discussed a child prediagnosis in the labeled set, or a child postdiagnosis. As presented in Figure 7, a different trend is observed among people who asked about a child before diagnosis, where two spikes represent questions after approximately 7 months and after approximately 2 years. We attribute these peaks to further stages in the care for the child, for which parents require additional information. The postdiagnosis curve not only shows a downward trend ($R^2=.30$, $P=.005$), indicating that people are less likely to ask an ADHD-related question after diagnosis, but also shows a peak around approximately 2 years, as for the undiagnosed situation.

Figure 7. The probability ratio of asking an additional attention deficit hyperactivity disorder (ADHD)–related question, divided by suspected of ADHD or already diagnosed.



Discussion

Principal Findings

The aim of this study was to obtain better insights about *where*, *what*, and *when* people seek Web-based ADHD-related information. The results indicate that the Internet indeed constitutes a source of information, as reflected by 613,069 queries about ADHD by 222,626 Bing users and 1609 questions in Yahoo Answers in the defined time of the study. These results are supported by prior studies, which suggest that the Internet is a source of information for people concerned with mental health and developmental disabilities. People, including 96% of parents, tend to seek out information on incidence, recurrence, or increased severity of the disability [26,30,31]. More specifically, both parents of adolescents with ADHD and the adolescents themselves have expressed similar strong preferences for ADHD information sources from the Internet (49% and 51%, respectively) and from a doctor (40% and 27%, respectively) [25]. Parents looking for information about their child's symptoms and behavior seek to obtain tools for managing the condition, in light of the lack of information on the various aspects of handling ADHD [3,32].

This study also found that people, who are interested in ADHD due to a close family member's diagnosis as reflected in IPQ queries, perform a more extensive search process compared

with others. This is in line with search characteristics for other medical conditions [33]. Specifically, the study results emphasize the necessity for more knowledge in order to cope with the existing misconceptions concerning the treatment of children with ADHD and assist families who need to cope with dilemmas, confusion, and uncertainty [3,34-37].

Currently, the literature about Web-based information seeking among parents of children with ADHD is scarce. To the best of our knowledge, there is no knowledge in the existing literature related to which specific information parents seek for. Results of this study that are presented in Table 1 indicate that among the 10 most frequent queries, amounting in over 12% of the queries, 61.5% were interested in ADHD or ADHD symptoms. These are in line with previous findings regarding the uncertainty of parents regarding the nature of ADHD [38]. Providing parents with more accurate and reliable information about early symptoms of ADHD is important. These early symptoms stand out as the most important risk factor for later antisocial development and impaired daily functioning as well as learning difficulties [39,40].

The next highest category (6.9%) of queries was related to medication. Nowadays, medication use is highly common worldwide and is still the mainstay of ADHD treatment in the United States [3]. Consequently, this finding supports this population's need for more research and knowledge about

addressing dilemmas bound with taking medication, which is a source of daily stress [41-43]. Furthermore, as previously revealed, 37 of 57 ADHD websites were in fact funded by drug companies. Hence, it is not clear whether the information provided on the Web indeed addresses this population's concerns and needs or commercial interests [44].

Aside from searching for information about ADHD symptoms and medication, 4.5% of the queries asked were related to children and 6.9% related to adults with ADHD. This finding is in accordance with the phenomena of ADHD as a life-span longitudinal disorder, affecting 2-4% of adults [45-47].

The study results provided not only a response to the question of *what* information was sought after, but also *where* the search was done. The list of the most popular websites that appears in Table 2 indicates that only 25% of the sites are ADHD-specific websites, whereas the remaining websites deal with other issues. The fact that only 35% of the clicked websites have been authorized by HoN indicates a preference of users for less authoritative websites, though possibly ones which provide more social support [48,49] or websites that serve the users' prior beliefs. As a result of this pattern, research has found that both parents and experts agree that the quality of the Web-based information regarding treatment choices for ADHD is generally poor and does not address parents' needs [27,48,50].

In the personal questions category on Yahoo Answers, where data about gender were available, the gender ratio was 3:1. Respectively, after addressing the questions of *what* information and *where* they searched for the information, the current results addressed the question of about *whom* and *when* on the time axis the question was asked. In line with previous findings that reported higher frequency of ADHD in males compared with females, with a gender ratio of 3:1 and 2.3:1, as well as greater impairment among boys [51,52], most questions in this study indeed were related to boys. Concerning the question of *when*, analysis revealed that questions by parents who suspected that their child had ADHD were asked already at the age of 2 years, whereas the average age in all those 258 questions is 5.1 years. This means that parental concerns about their child's behavior and daily function in the preschool years begin well before the child enters the educational system, although children with ADHD are usually diagnosed during school years at just over the age of 7. As retrospective reporting of earlier symptoms is difficult for parents, the Web may be a useful tool for capturing information, which can support and establish the validity of the criteria of previous manifestations for future diagnosis [53].

Moreover, it was interesting to obtain information not only about *which*, *where*, and *when*, but also about *what*? What indeed troubles the parents? What information is sought after? The *what* question was analyzed among those who declared having a diagnosis and those who did not declared so. As presented in Figures 4 and 5, the most frequent issue raised in both of those groups was the behavioral aspect of ADHD. However, it is interesting to note that although 20% of the questions asked among families whose child's diagnosis is unknown concerned functional daily issues, in the already diagnosed group, this category did not appear in relation to younger children. Nonetheless, there is evidence in the literature

that functional deficits occur among this population over the years in the areas of self-hygiene, morning and evening organization, and play and social relationships [54-57]. It is unclear as to why these issues do not reflect in the Web-based searches. It may be that the parents do not expect to find an appropriate solution to such problems or that they feel that these issues are trivial and as parents, they are expected to cope with them regardless to the ADHD diagnosis.

Another interesting finding is that in both the group that declared and the group that did not declare a diagnosis, 20% or more of the questions around the age of 12 years were related to the academic queries category.

Academic deficits among children with ADHD have long been documented (eg, [58-60]). As academic requirements increase over the years, it is not surprising that this issue troubles parents and that they search the Web for information about how to deal with these difficulties [58]. However, it is surprising that more questions related to academic performance were asked in the group who did not declare a diagnosis. Perhaps these questions are part of the parent's dilemma as to whether the symptoms they see are indeed indicators of ADHD, whereas among those who are already diagnosed, questions about academic performance are possibly discussed with educational and therapeutic professionals or on specific ADHD parents' blogs or forums.

Further to achieving insight about the *which*, *when*, *where*, and *what*, our purpose was to identify who are the people that ask additional questions following their first question. The results showed that a question asked among those with no declared diagnosis and included signs of helplessness, is associated with asking more questions online. As mentioned above, parents' daily confrontations with the disorder serves as the motivation for Web-based use. Therefore, it is understandable that parents, who are more worried and more stressed as reflected in their first question, continue to search for additional assistance.

When analyzing the phenomena of whether people whose children were diagnosed asked more questions, it was found that these people are less likely to continue to ask questions online. Perhaps this tendency occurs because at this stage, parents refer their questions to their doctor in light of the findings of preference of this source among this population [25]. Another explanation may be that parents who have a question concerning ADHD may find it easier to ask other people for solutions, rather than seek for answers in preprepared information pages. However, the peak in amount of questions asked in both groups that appeared after approximately 2 years may indicate that as the school environment requirements increase, parents again try to search for information regarding this demanding phase in life.

Summary

In summary, our results indicate that the Internet is indeed a source of information about ADHD and tracking the search process performed by the users on the Web sheds light on people's needs. Besides the high priority given by parents of adolescents with ADHD to receiving information from the

doctor, regardless the time constrains [25], they indeed use further Web-based searching [3,25,32].

Thus, as ADHD is a chronic long-life phenomenon, expert professions may use the Web as a useful source of information for supplying accurate information over time according to this population's daily needs [25]. Acquiring information may prevent emotional consequences and social difficulties which in turn burden costs for individuals, health services, and governments (eg, [9]). Negative life outcomes and underachievement may be prevented by understanding the specific kind of assistance is required and by providing the help in a timely manner [9].

Despite the interesting results, this study has several limitations. As far as known, this is the first study in which a quantitative analysis of the search extent and content was performed. The classification criteria as well as interrater validity need to further be discussed and improved.

Further studies that include more accurate criteria as well as a qualitative analysis of the queries made within international Internet forums are required to achieve better insight about this population's needs and the existing barriers to seeking help (eg, [17,61]).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples to words that appear in each category.

[PDF File (Adobe PDF File), 36KB - [jmir_v19i4e126_app1.pdf](#)]

References

1. Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry* 2007 Jun;164(6):942-948. [doi: [10.1176/ajp.2007.164.6.942](#)] [Medline: [17541055](#)]
2. Russell AE, Ford T, Russell G. Socioeconomic associations with ADHD: findings from a mediation analysis. *PLoS One* 2015;10(6):e0128248 [FREE Full text] [doi: [10.1371/journal.pone.0128248](#)] [Medline: [26030626](#)]
3. Pajo B, Stuart P. A comparative review of "how to" books for parents of ADHD children and "how to" books for parents of typical children. *Child Youth Serv Rev* 2012 Apr;34(4):826-833. [doi: [10.1016/j.childyouth.2012.01.011](#)]
4. Cormier E. Attention deficit/hyperactivity disorder: a review and update. *J Pediatr Nurs* 2008 Oct;23(5):345-357. [doi: [10.1016/j.pedn.2008.01.003](#)] [Medline: [18804015](#)]
5. Camilleri N, Makhoul S. ADHD: from childhood into adulthood. *Malta Medical Journal* 2013 Jan;25(1):2-7 [FREE Full text]
6. Rutter M, Kim-Cohen J, Maughan B. Continuities and discontinuities in psychopathology between childhood and adult life. *J Child Psychol Psychiatry* 2006;47(3-4):276-295. [doi: [10.1111/j.1469-7610.2006.01614.x](#)] [Medline: [16492260](#)]
7. Harpin VA. The effect of ADHD on the life of an individual, their family, and community from preschool to adult life. *Arch Dis Child* 2005 Feb;90(Suppl 1):i2-i7 [FREE Full text] [doi: [10.1136/adc.2004.059006](#)] [Medline: [15665153](#)]
8. Ma J, Lai K, Pun S. Parenting distress and parental investment of Hong Kong Chinese parents with a child having an emotional or behavioural problem: a qualitative study. *Child Fam Soc Work* 2002 May;7(2):99-106. [doi: [10.1046/j.1365-2206.2002.00232.x](#)]
9. Newark PE, Stieglitz R. Therapy-relevant factors in adult ADHD from a cognitive behavioural perspective. *Atten Defic Hyperact Disord* 2010 Jun;2(2):59-72. [doi: [10.1007/s12402-010-0023-1](#)] [Medline: [21432591](#)]
10. Dubouloz CJ, King J, Ashe B, Paterson B, Chevrier J, Moldoveanu M. The process of transformation in rehabilitation: what does it look like? *Int J Therapy Rehab* 2010 Nov;17(11):604-615. [doi: [10.12968/ijtr.2010.17.11.79541](#)]
11. Bussing R, Gary FA, Mills TL, Garvan CW. Cultural variations in parental health beliefs, knowledge, and information sources related to attention-deficit/hyperactivity disorder. *J Fam Issues* 2007 Mar 01;28(3):291-318. [doi: [10.1177/0192513X06296117](#)]
12. Subcommittee on Attention-Deficit/Hyperactivity Disorder, Steering Committee on Quality Improvement and Management, Wolraich M, Brown L, Brown RT, DuPaul G, et al. ADHD: clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics* 2011 Nov;128(5):1007-1022. [doi: [10.1542/peds.2011-2654](#)] [Medline: [22003063](#)]
13. Morgan P, Staff J, Hillemeier M, Farkas G, Maczuga S. Racial and ethnic disparities in ADHD diagnosis from kindergarten to eighth grade. *Pediatrics* 2013 Jul;132(1):85-93. [Medline: [23796743](#)]
14. Ogundele MO, De Doysa R. Is ADHD a disease of affluence? A local authority experience in north-west of England. *Arch Dis Child* 2011 Apr;96(Supplement 1):A25-A26. [doi: [10.1136/adc.2011.212563.52](#)]
15. Schneider H, Eisenberg D. Who receives a diagnosis of attention-deficit/hyperactivity disorder in the United States elementary school population? *Pediatrics* 2006 Apr;117(4):e601-e609. [doi: [10.1542/peds.2005-1308](#)] [Medline: [16585277](#)]

16. Sciberras E, Iyer S, Efron D, Green J. Information needs of parents of children with attention-deficit/hyperactivity disorder. *Clin Pediatr (Phila)* 2010 Feb;49(2):150-157. [doi: [10.1177/0009922809346730](https://doi.org/10.1177/0009922809346730)] [Medline: [20080521](https://pubmed.ncbi.nlm.nih.gov/20080521/)]
17. Terbeck S, Chesterman LP. Parents, ADHD and the internet. *Atten Defic Hyperact Disord* 2012 Sep;4(3):159-166. [doi: [10.1007/s12402-012-0081-7](https://doi.org/10.1007/s12402-012-0081-7)] [Medline: [22736404](https://pubmed.ncbi.nlm.nih.gov/22736404/)]
18. Goldman RD, Macpherson A. Internet health information use and e-mail access by parents attending a paediatric emergency department. *Emerg Med J* 2006 May;23(5):345-348 [FREE Full text] [doi: [10.1136/emj.2005.026872](https://doi.org/10.1136/emj.2005.026872)] [Medline: [16627833](https://pubmed.ncbi.nlm.nih.gov/16627833/)]
19. Hogue MB, Doran E, Henry DA. A prompt to the web: the media and health information seeking behaviour. *PLoS One* 2012;7(4):e34314 [FREE Full text] [doi: [10.1371/journal.pone.0034314](https://doi.org/10.1371/journal.pone.0034314)] [Medline: [22509289](https://pubmed.ncbi.nlm.nih.gov/22509289/)]
20. Reichow B, Gelbar NW, Mouradjian K, Shefcyk A, Smith IC. Characteristics of international websites with information on developmental disabilities. *Res Dev Disabil* 2014 Oct;35(10):2293-2298. [doi: [10.1016/j.ridd.2014.05.028](https://doi.org/10.1016/j.ridd.2014.05.028)] [Medline: [24952371](https://pubmed.ncbi.nlm.nih.gov/24952371/)]
21. Kivits J. Informed patients and the internet: a mediated context for consultations with health professionals. *J Health Psychol* 2006 Mar;11(2):269-282. [doi: [10.1177/1359105306061186](https://doi.org/10.1177/1359105306061186)] [Medline: [16464924](https://pubmed.ncbi.nlm.nih.gov/16464924/)]
22. Chiu Y. Probing, impelling, but not offending doctors: the role of the internet as an information source for patients' interactions with doctors. *Qual Health Res* 2011 Dec;21(12):1658-1666. [doi: [10.1177/1049732311417455](https://doi.org/10.1177/1049732311417455)] [Medline: [21799204](https://pubmed.ncbi.nlm.nih.gov/21799204/)]
23. 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](https://doi.org/10.1177/1049732310368404)] [Medline: [20442347](https://pubmed.ncbi.nlm.nih.gov/20442347/)]
24. Cunningham CE, Deal K, Rimas H, Chen Y, Buchanan DH, Sdao-Jarvie K. Providing information to parents of children with mental health problems: a discrete choice conjoint analysis of professional preferences. *J Abnorm Child Psychol* 2009 Nov;37(8):1089-1102. [doi: [10.1007/s10802-009-9338-9](https://doi.org/10.1007/s10802-009-9338-9)] [Medline: [19629676](https://pubmed.ncbi.nlm.nih.gov/19629676/)]
25. Bussing R, Zima BT, Mason DM, Meyer JM, White K, Garvan CW. ADHD knowledge, perceptions, and information sources: perspectives from a community sample of adolescents and their parents. *J Adolesc Health* 2012 Dec;51(6):593-600 [FREE Full text] [doi: [10.1016/j.jadohealth.2012.03.004](https://doi.org/10.1016/j.jadohealth.2012.03.004)] [Medline: [23174470](https://pubmed.ncbi.nlm.nih.gov/23174470/)]
26. Reichow B, Shefcyk A, Bruder MB. Quality comparison of websites related to developmental disabilities. *Res Dev Disabil* 2013 Oct;34(10):3077-3083. [doi: [10.1016/j.ridd.2013.06.013](https://doi.org/10.1016/j.ridd.2013.06.013)] [Medline: [23891722](https://pubmed.ncbi.nlm.nih.gov/23891722/)]
27. Montoya A, Hernández S, Massana MP, Herreros O, Garcia-Giral M, Cardo E, et al. Evaluating Internet information on attention-deficit/hyperactivity disorder (ADHD) treatment: parent and expert perspectives. *Educ Health (Abingdon)* 2013;26(1):48-53 [FREE Full text] [doi: [10.4103/1357-6283.112801](https://doi.org/10.4103/1357-6283.112801)] [Medline: [23823673](https://pubmed.ncbi.nlm.nih.gov/23823673/)]
28. Yom-Tov E, Muennig P, El-Sayed A. Web-based antismoking advertising to promote smoking cessation: a randomized controlled trial. *J Med Internet Res* 2016 Nov;18(11):e306. [doi: [10.2196/jmir.6563](https://doi.org/10.2196/jmir.6563)] [Medline: [27872032](https://pubmed.ncbi.nlm.nih.gov/27872032/)]
29. Andersen P, Gill R. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982 Dec;10(4):1100-1120 [FREE Full text] [doi: [10.1214/aos/1176345976](https://doi.org/10.1214/aos/1176345976)]
30. Moseley KL, Freed GL, Goold SD. Which sources of child health advice do parents follow? *Clin Pediatr (Phila)* 2011 Jan;50(1):50-56. [doi: [10.1177/0009922810379905](https://doi.org/10.1177/0009922810379905)] [Medline: [20837618](https://pubmed.ncbi.nlm.nih.gov/20837618/)]
31. Powell J, Clarke A. Internet information-seeking in mental health: population survey. *Br J Psychiatry* 2006 Sep;189:273-277 [FREE Full text] [doi: [10.1192/bjp.bp.105.017319](https://doi.org/10.1192/bjp.bp.105.017319)] [Medline: [16946364](https://pubmed.ncbi.nlm.nih.gov/16946364/)]
32. dosReis S, Mychailyszyn MP, Myers M, Riley AW. Coming to terms with ADHD: how urban African-American families come to seek care for their children. *Psychiatr Serv* 2007 May;58(5):636-641. [doi: [10.1176/appi.ps.58.5.636](https://doi.org/10.1176/appi.ps.58.5.636)] [Medline: [17463344](https://pubmed.ncbi.nlm.nih.gov/17463344/)]
33. Ofra Y, Paltiel O, Pelleg D, Rowe JM, Yom-Tov E. Patterns of information-seeking for cancer on the internet: an analysis of real world data. *PLoS One* 2012;7(9):e45921 [FREE Full text] [doi: [10.1371/journal.pone.0045921](https://doi.org/10.1371/journal.pone.0045921)] [Medline: [23029317](https://pubmed.ncbi.nlm.nih.gov/23029317/)]
34. Charach A, Skyba A, Cook L, Antle BJ. Using stimulant medication for children with ADHD: what do parents say? A brief report. *J Can Acad Child Adolesc Psychiatry* 2006 May;15(2):75-83 [FREE Full text] [Medline: [18392197](https://pubmed.ncbi.nlm.nih.gov/18392197/)]
35. Hansen DL, Hansen EH. Caught in a balancing act: parents' dilemmas regarding their ADHD child's treatment with stimulant medication. *Qual Health Res* 2006 Nov;16(9):1267-1285. [doi: [10.1177/1049732306292543](https://doi.org/10.1177/1049732306292543)] [Medline: [17038757](https://pubmed.ncbi.nlm.nih.gov/17038757/)]
36. Perera S, Katangodage D. G509(P) Managing parental expectations in paediatric adhd clinics – a new model of care. *Arch Dis Child* 2015 Apr 27;100(Suppl 3):A219.1-A2A219 [FREE Full text] [doi: [10.1136/archdischild-2015-308599.462](https://doi.org/10.1136/archdischild-2015-308599.462)]
37. Scuitto MJ. ADHD knowledge, misconceptions, and treatment acceptability. *J Atten Disord* 2015 Feb;19(2):91-98. [doi: [10.1177/1087054713493316](https://doi.org/10.1177/1087054713493316)] [Medline: [23893531](https://pubmed.ncbi.nlm.nih.gov/23893531/)]
38. Taylor E. Developing ADHD. *J Child Psychol Psychiatry* 2009 Jan;50(1-2):126-132. [doi: [10.1111/j.1469-7610.2008.01999.x](https://doi.org/10.1111/j.1469-7610.2008.01999.x)] [Medline: [19076263](https://pubmed.ncbi.nlm.nih.gov/19076263/)]
39. Czamara D, Tiesler CM, Kohlböck G, Berdel D, Hoffmann B, Bauer C, et al. Children with ADHD symptoms have a higher risk for reading, spelling and math difficulties in the GINIplus and LISAPLUS cohort studies. *PLoS One* 2013;8(5):e63859 [FREE Full text] [doi: [10.1371/journal.pone.0063859](https://doi.org/10.1371/journal.pone.0063859)] [Medline: [23724008](https://pubmed.ncbi.nlm.nih.gov/23724008/)]
40. Norén SE, Molero Y, Lichtenstein P, Larson T, Lundström S, Anckarsäter H, et al. Childhood symptoms of ADHD overrule comorbidity in relation to psychosocial outcome at age 15: a longitudinal study. *PLoS One* 2015;10(9):e0137475 [FREE Full text] [doi: [10.1371/journal.pone.0137475](https://doi.org/10.1371/journal.pone.0137475)] [Medline: [26360378](https://pubmed.ncbi.nlm.nih.gov/26360378/)]

41. Clavenna A, Bonati M. Safety of medicines used for ADHD in children: a review of published prospective clinical trials. *Arch Dis Child* 2014 Sep;99(9):866-872. [doi: [10.1136/archdischild-2013-304170](https://doi.org/10.1136/archdischild-2013-304170)] [Medline: [24748641](https://pubmed.ncbi.nlm.nih.gov/24748641/)]
42. Prosser B, Reid R. Changes in use of psychostimulant medication for ADHD in South Australia (1990-2006). *Aust N Z J Psychiatry* 2009 Apr;43(4):340-347. [doi: [10.1080/00048670902721129](https://doi.org/10.1080/00048670902721129)] [Medline: [19296289](https://pubmed.ncbi.nlm.nih.gov/19296289/)]
43. Thorell LB, Dahlström K. Children's self-reports on perceived effects on taking stimulant medication for ADHD. *J Atten Disord* 2009 Mar;12(5):460-468. [doi: [10.1177/1087054708320430](https://doi.org/10.1177/1087054708320430)] [Medline: [18685139](https://pubmed.ncbi.nlm.nih.gov/18685139/)]
44. Mitchell J, Read J. Attention-deficit hyperactivity disorder, drug companies and the internet. *Clin Child Psychol Psychiatry* 2012 Jan;17(1):121-139. [doi: [10.1177/1359104510396432](https://doi.org/10.1177/1359104510396432)] [Medline: [21429977](https://pubmed.ncbi.nlm.nih.gov/21429977/)]
45. Ginsberg Y, Hirvikoski T, Lindefors N. Attention Deficit Hyperactivity Disorder (ADHD) among longer-term prison inmates is a prevalent, persistent and disabling disorder. *BMC Psychiatry* 2010 Dec 22;10:112 [FREE Full text] [doi: [10.1186/1471-244X-10-112](https://doi.org/10.1186/1471-244X-10-112)] [Medline: [21176203](https://pubmed.ncbi.nlm.nih.gov/21176203/)]
46. Kessler RC, Adler L, Barkley R, Biederman J, Conners CK, Demler O, et al. The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. *Am J Psychiatry* 2006 Apr;163(4):716-723 [FREE Full text] [doi: [10.1176/ajp.2006.163.4.716](https://doi.org/10.1176/ajp.2006.163.4.716)] [Medline: [16585449](https://pubmed.ncbi.nlm.nih.gov/16585449/)]
47. Levy S, Katusic SK, Colligan RC, Weaver AL, Killian JM, Voigt RG, et al. Childhood ADHD and risk for substance dependence in adulthood: a longitudinal, population-based study. *PLoS One* 2014;9(8):e105640 [FREE Full text] [doi: [10.1371/journal.pone.0105640](https://doi.org/10.1371/journal.pone.0105640)] [Medline: [25162629](https://pubmed.ncbi.nlm.nih.gov/25162629/)]
48. Akram G, Thomson AH, Boyter AC, Morton MJS. Characterisation and evaluation of UK websites on attention deficit hyperactivity disorder. *Arch Dis Child* 2008 Aug 01;93(8):695-700. [doi: [10.1136/adc.2007.130708](https://doi.org/10.1136/adc.2007.130708)] [Medline: [18337283](https://pubmed.ncbi.nlm.nih.gov/18337283/)]
49. Pelleg D, Yom-Tov E, Maarek Y. Can you believe an anonymous contributor? On truthfulness in Yahoo! Answers. 2012 Presented at: ASE/IEEE International Conference on Social Computing and ASE/IEEE International Conference on Privacy, Security, Risk and Trust; Sept. ; Amsterdam, The Netherlands; September 3-5, 2012; Amsterdam p. 411-420.
50. Reavley NJ, Jorm AF. The quality of mental disorder information websites: a review. *Patient Educ Couns* 2011 Nov;85(2):e16-e25. [doi: [10.1016/j.pec.2010.10.015](https://doi.org/10.1016/j.pec.2010.10.015)] [Medline: [21087837](https://pubmed.ncbi.nlm.nih.gov/21087837/)]
51. Bauermeister JJ, Shrout PE, Chávez L, Rubio-Stipec M, Ramírez R, Padilla L, et al. ADHD and gender: are risks and sequela of ADHD the same for boys and girls? *J Child Psychol Psychiatry* 2007 Aug;48(8):831-839. [doi: [10.1111/j.1469-7610.2007.01750.x](https://doi.org/10.1111/j.1469-7610.2007.01750.x)] [Medline: [17683455](https://pubmed.ncbi.nlm.nih.gov/17683455/)]
52. Bruchmüller K, Margraf J, Schneider S. Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *J Consult Clin Psychol* 2012 Feb;80(1):128-138. [doi: [10.1037/a0026582](https://doi.org/10.1037/a0026582)] [Medline: [22201328](https://pubmed.ncbi.nlm.nih.gov/22201328/)]
53. Polanczyk G, Caspi A, Houts R, Kollins SH, Rohde LA, Moffitt TE. Implications of extending the ADHD age-of-onset criterion to age 12: results from a prospectively studied birth cohort. *J Am Acad Child Adolesc Psychiatry* 2010 Mar;49(3):210-216. [Medline: [20410710](https://pubmed.ncbi.nlm.nih.gov/20410710/)]
54. Brinkman WB, Sherman SN, Zmitrovich AR, Visscher MO, Crosby LE, Phelan KJ, et al. Parental angst making and revisiting decisions about treatment of attention-deficit/hyperactivity disorder. *Pediatrics* 2009 Aug;124(2):580-589. [doi: [10.1542/peds.2008-2569](https://doi.org/10.1542/peds.2008-2569)] [Medline: [19651580](https://pubmed.ncbi.nlm.nih.gov/19651580/)]
55. Sayal K, Goodman R, Ford T. Barriers to the identification of children with attention deficit/hyperactivity disorder. *J Child Psychol Psychiatry* 2006 Jul;47(7):744-750. [doi: [10.1111/j.1469-7610.2005.01553.x](https://doi.org/10.1111/j.1469-7610.2005.01553.x)] [Medline: [16790009](https://pubmed.ncbi.nlm.nih.gov/16790009/)]
56. Miller M, Hinshaw SP. Does childhood executive function predict adolescent functional outcomes in girls with ADHD? *J Abnorm Child Psychol* 2010 Apr;38(3):315-326 [FREE Full text] [doi: [10.1007/s10802-009-9369-2](https://doi.org/10.1007/s10802-009-9369-2)] [Medline: [19960365](https://pubmed.ncbi.nlm.nih.gov/19960365/)]
57. Wehmeier PM, Dittmann RW, Schacht A, Helsberg K, Lehmkuhl G. Morning and evening behavior in children and adolescents treated with atomoxetine once daily for attention-deficit/hyperactivity disorder (ADHD): findings from two 24-week, open-label studies. *Child Adolesc Psychiatry Ment Health* 2009 Feb 09;3(1):5 [FREE Full text] [doi: [10.1186/1753-2000-3-5](https://doi.org/10.1186/1753-2000-3-5)] [Medline: [19203355](https://pubmed.ncbi.nlm.nih.gov/19203355/)]
58. Daley D, Birchwood J. ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? *Child Care Health Dev* 2010 Jul;36(4):455-464. [doi: [10.1111/j.1365-2214.2009.01046.x](https://doi.org/10.1111/j.1365-2214.2009.01046.x)] [Medline: [20074251](https://pubmed.ncbi.nlm.nih.gov/20074251/)]
59. Loe IM, Feldman HM. Academic and educational outcomes of children with ADHD. *J Pediatr Psychol* 2007 Jul;32(6):643-654 [FREE Full text] [doi: [10.1093/jpepsy/jsl054](https://doi.org/10.1093/jpepsy/jsl054)] [Medline: [17569716](https://pubmed.ncbi.nlm.nih.gov/17569716/)]
60. Raggi VL, Chronis AM. Interventions to address the academic impairment of children and adolescents with ADHD. *Clin Child Fam Psychol Rev* 2006 Jun;9(2):85-111. [doi: [10.1007/s10567-006-0006-0](https://doi.org/10.1007/s10567-006-0006-0)] [Medline: [16972189](https://pubmed.ncbi.nlm.nih.gov/16972189/)]
61. Tzang R, Chang Y, Chen C. Barriers to seeking help among children with attention deficit hyperactivity disorder in Taiwan. *Asia Pac Psychiatry* 2014 Dec;6(4):373-378. [doi: [10.1111/appy.12064](https://doi.org/10.1111/appy.12064)] [Medline: [23857893](https://pubmed.ncbi.nlm.nih.gov/23857893/)]

Abbreviations

ADHD: attention deficit hyperactivity disorder
HoN: Health on the Net
IPQ: Individual Persons Query

Edited by G Eysenbach; submitted 30.08.16; peer-reviewed by I Weber, M Torok, TR Soron, Y Peng; comments to author 19.11.16; revised version received 14.12.16; accepted 14.02.17; published 21.04.17.

Please cite as:

Rosenblum S, Yom-Tov E

Seeking Web-Based Information About Attention Deficit Hyperactivity Disorder: Where, What, and When

J Med Internet Res 2017;19(4):e126

URL: <http://www.jmir.org/2017/4/e126/>

doi: [10.2196/jmir.6579](https://doi.org/10.2196/jmir.6579)

PMID: [28432038](https://pubmed.ncbi.nlm.nih.gov/28432038/)

©Sara Rosenblum, Elad Yom-Tov. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 21.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Usability, Acceptability, and Adherence to an Electronic Self-Monitoring System in Patients With Major Depression Discharged From Inpatient Wards

Lise Lauritsen^{1*}, MD; Louise Andersen^{1*}, MD; Emilia Olsson^{1*}, MD; Stine Rauff Søndergaard^{1*}, MD; Lasse Benn Nørregaard^{2*}, MSc; Philip Kaare Løventoft^{3*}, MSc; Signe Dunker Svendsen^{1*}, MSc; Erik Frøkjær^{4*}, MSc; Hans Mørch Jensen^{1*}, MD; Ida Hageman¹, MD; Lars Vedel Kessing^{1*}, MD, DMSc; Klaus Martiny^{1*}, PhD, MD, DMSc

¹Psychiatric Center Copenhagen, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

²Monsenso App, Copenhagen, Denmark

³Rehfeld Medical, Copenhagen, Denmark

⁴Department of Computer Science, Copenhagen University, Copenhagen, Denmark

*these authors contributed equally

Corresponding Author:

Klaus Martiny, PhD, MD, DMSc

Psychiatric Center Copenhagen

Rigshospitalet

University of Copenhagen

Edel Sautes Alle 10

Copenhagen, 2100 Ø

Denmark

Phone: 45 38647100

Fax: 45 38647106

Email: klaus.martiny@regionh.dk

Abstract

Background: Patients suffering from depression have a high risk of relapse and readmission in the weeks following discharge from inpatient wards. Electronic self-monitoring systems that offer patient-communication features are now available to offer daily support to patients, but the usability, acceptability, and adherence to these systems has only been sparsely investigated.

Objective: We aim to test the usability, acceptability, adherence, and clinical outcome of a newly developed computer-based electronic self-assessment system (the Daybuilder system) in patients suffering from depression, in the period from discharge until commencing outpatient treatment in the Intensive Outpatient Unit for Affective Disorders.

Methods: Patients suffering from unipolar major depression that were referred from inpatient wards to an intensive outpatient unit were included in this study before their discharge, and were followed for four weeks. User satisfaction was assessed using semiquantitative questionnaires and the System Usability Scale (SUS). Patients were interviewed at baseline and at endpoint with the Hamilton depression rating scale (HAM-D17), the Major Depression Inventory (MDI), and the 5-item World Health Organization Well-Being Index (WHO-5). In this four-week period patients used the Daybuilder system to self-monitor mood, sleep, activity, and medication adherence on a daily basis. The system displayed a graphical representation of the data that was simultaneously displayed to patients and clinicians. Patients were phoned weekly to discuss their data entries. The primary outcomes were usability, acceptability, and adherence to the system. The secondary outcomes were changes in: the electronically self-assessed mood, sleep, and activity scores; and scores from the HAM-D17, MDI, and WHO-5 scales.

Results: In total, 76% of enrolled patients (34/45) completed the four-week study. Five patients were readmitted due to relapse. The 34 patients that completed the study entered data for mood on 93.8% of the days (872/930), sleep on 89.8% of the days (835/930), activity on 85.6% of the days (796/930), and medication on 88.0 % of the days (818/930). SUS scores were 86.2 (standard deviation [SD] 9.7) and 79% of the patients (27/34) found that the system lived up to their expectations. A significant improvement in depression severity was found on the HAM-D17 from 18.0 (SD 6.5) to 13.3 (SD 7.3; $P<.01$), on the MDI from 27.1 (SD 13.1) to 22.1 (SD 12.7; $P=.006$), and in quality of life on the WHO-5 from 31.3 (SD 22.9) to 43.4 (SD 22.1; $P<.001$) scales, but not on self-assessed mood ($P=.08$). Mood and sleep parameters were highly variable from day-to-day. Sleep-offset

was significantly delayed from baseline, averaging 48 minutes (standard error 12 minutes; $P < .001$). Furthermore, when estimating delay of sleep-onset (with sleep quality included in the model) during the study period, this showed a significant negative effect on mood ($P = .03$).

Conclusions: The Daybuilder systems performed well technically, and patients were satisfied with the system and had high adherence to self-assessments. The dropout rate and the gradual delay in sleep emphasize the need for continued clinical support for these patients, especially when considering sleep guidance.

(*J Med Internet Res* 2017;19(4):e123) doi:[10.2196/jmir.6673](https://doi.org/10.2196/jmir.6673)

KEYWORDS

depressive disorder, major; electronic monitoring; graph drawing; sleep; self-assessment; observational study; inpatients; patient participation; chronotherapeutics

Introduction

Major Depression

Major depression is estimated by the World Health Organization to top the list of the 20 most financially burdensome disorders in the Western world [1], and a substantial part of these costs cover hospital treatment for patients with severe depression who have long-standing admissions, and who are often readmitted due to relapse [2,3]. Thus, in Denmark, these patients have readmission rates between 10-30% in the months following discharge [4]. Furthermore, there is an increased risk of suicide in the immediate period following discharge [5]. Therefore, new tools and treatments are needed to prevent readmission and suicide after discharge from inpatients wards.

Electronic Monitoring

Increased access to reliable and fast Internet services and the development of interactive systems have spurred interest in the use of electronic monitoring in medicine [6], including psychiatry, where electronic monitoring is increasingly being used as a clinical tool [7,8]. Existing systems differ in mode of function and complexity. Some systems offer interactive self-help with psychotherapy/psychoeducation [9], some include clinical backup, some are designed for use on smartphones, and others are designed for computers or tablets [10]. A small number of systems use a bidirectional feedback loop between patients and clinicians [11,12], in which data entered by patients can be seen in real-time by the clinician, and responses can occur immediately if needed. These systems make it easier for patients and clinicians to notice changes in conditions over time, be it behavioral changes such as activity or medication, or psychopathological symptoms (eg, mood and sleep). Using these systems, worsening of the condition can be acted upon by patients and clinicians in collaboration, via comonitoring [13]. In general, data entered in these systems are graphically presented to enable patients and researchers/clinicians to visualize relationships between the entered data and development of the measured variables over time.

When used in research, electronic monitoring has a number of advantages: it partly eliminates the need for pen and paper data collection, it makes data immediately available for analyses, and it secures day-to-day information that might otherwise be lost due to cognitive dysfunction that is prevalent in patients with major depression [14]. If the clinical practice can match the opportunities given in interactive systems by supplying

feedback, patients will be enabled to make earlier adjustments in treatment and behavior by having an earlier and more focused response to their condition.

High dropout rates are a common problem in electronic monitoring [15], although studies using systems that connect patients and clinicians have higher adherence rates [16]. Some applications have a multitude of options and can be difficult to use for patients with depression. Simplicity and ease of use are considered essential for users [17]. The system that we evaluated in this study (the Daybuilder system) was developed with this in mind, having a reduced number of assessment parameters and a graphical representation that is central in the screen. The Daybuilder system has been developed in close collaboration with users and clinicians in psychiatry [18].

Summary

In summary, by using interactive electronic monitoring we can: obtain day-to-day information of clinical state; obtain a measure of day-to-day variability; and enable patients and clinicians to discover time-trends and relationships between variables that facilitate early and more focused intervention, via the graphical representation of data over time. A pilot test with the Daybuilder system showed the system to be easy to use and stable [18]. In this study, the aim was to test the usability, acceptability, and adherence to the Daybuilder system in patients suffering from depression who were discharged from inpatient psychiatric wards, and patients were followed for four weeks. We also report relapse, the time courses of mood, sleep, and activity, and the interrelatedness of these factors.

Methods

Design

The study design was a single arm observational study. The study consisted of three distinct periods: Phase 1 was the period from inclusion to discharge, Phase 2 was the period from discharge to commencing treatment at the Intensive Outpatient Unit for Affective Disorders (IOA) service, and Phase 3 was the period from the start at IOA to the end of each patient's four-week study period.

Participants

Patients suffering from unipolar major depression that were referred from inpatient wards to an IOA were asked to participate in the study before they were discharged from psychiatric inpatient wards. The scientific ethical committee

for the Capital Region of Copenhagen was informed about the study in writing (Journal nr. H-3-2013-FSP32) and decided that the study did not require review by the committee. The study was approved by Psychiatric Center Copenhagen and the Danish Data Agency (RHP-2013-023, I-Suite number: 02470). The patient identification list was kept under double lock. Baseline procedures consisted of a psychometric assessment and an introduction to the Daybuilder system. Informed consent was obtained from all participants after oral and written information had been given about the study content and possible consequences.

Eligibility Criteria

Inclusion criteria included: major depression as defined in the Diagnostic and Statistical Manual of Mental Disorders 4th edition, and age >18 years. Exclusion criteria included: suicidality (corresponding to a score of 2 or above on the Hamilton depression rating scale [HAM-D₁₇] item 3, or if the investigators were unable to assess the degree of suicidality), abuse of alcohol or other substances that could influence the use of the Daybuilder system, bipolar illness, psychotic depression for the last two weeks prior to inclusion, and comorbid dementia or other organic brain damage that could influence the participant's ability to use the Daybuilder system. Criteria for leaving the study included: patients wishing to leave the study or admittance for a somatic illness that would potentially influence the ability to use the Daybuilder system. Patients were allowed to continue in the study if readmitted to an inpatient psychiatric ward.

Psychometric Assessment

Sociodemographics were collected through interviews and from case files. Diagnostic confirmation was done by use of the Mini-International Neuropsychiatric Interview instrument [19]. Baseline and endpoint depression severity were assessed by the investigator-administered HAM-D₁₇ scale [20] which covers the full spectrum of depression symptoms, the Hamilton six-item subscale (HAM-D₆) [21] which covers the core symptoms of depression, the Bech-Rafaelsen melancholia scale (MES) which includes items covering symptoms of psychomotor retardation [22], and paper-and-pen self-assessment was done with the

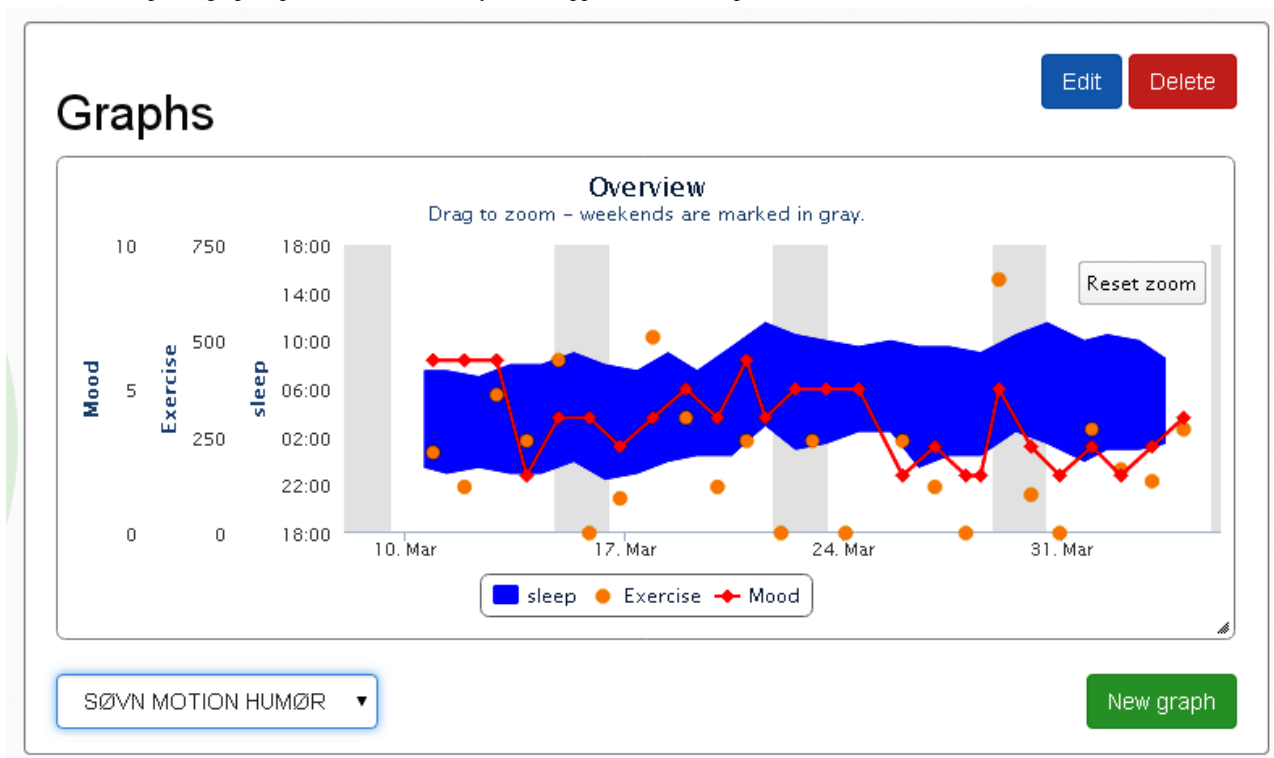
Major Depression Inventory (MDI) [23], and the 5-item World Health Organization Well-Being Index (WHO-5) scale [24].

Patients answered semiquantitative questions regarding the usability of the system at baseline and at the endpoint. These questions covered expectations on the use of the system at inclusion, and reflection on experiences with the system at the endpoint. Patients were asked to fill in the System Usability Scale (SUS) at endpoint [25].

Daybuilder Procedures

A patient-specific profile was created on the Daybuilder webpage for each patient, and each person was assigned a study number and an email address to enable them to log into the system. Patients were instructed on how to use Daybuilder, and how to enter the following variables on all days of the four-week study period: sleep-onset, sleep-offset, number of awakenings at night, quality of sleep, naps (time and duration), mood (morning and evening), activity (number of minutes outside the psychiatric ward, or when discharged as minutes outside their home as an estimate of activity), and medication (whether daily medication were taken or not). Mood and quality of sleep were entered on a Visual Analog Scale (VAS; 0=worst depression/worst sleep ever; 10=no depression/best sleep). Participants were instructed to enter mood scores in the morning and evening. Patients began data entry in the Daybuilder system on the day of inclusion. Study investigators phoned patients weekly to aid with any problems related to Daybuilder and to discuss outcomes of data monitoring, as seen in the Daybuilder graphs. Patients were seen by investigators at a final visit after four weeks. Text messages were used during the project to help patients remember that they had to enter their data. Patients' data were not seen between telephone calls.

All entries into the Daybuilder system were done through the computer (personal computer or Mac), except mood values which could also be entered through short message service (SMS) text messaging. A reminder was sent twice daily over SMS texting to register mood. When entering values in Daybuilder, the system automatically generated a graphic display of all variables to aid understanding of evolving patterns and relationships between variables (eg, between mood, sleep, and activity; see Figure 1).

Figure 1. An example of graphic presentation in the Daybuilder application for sleep, mood, and exercise.

Outcomes

The primary outcomes were usability, acceptability, and adherence, which were assessed by SUS scores, semiquantitative questions, and adherence to data entry in the electronic system. Secondary outcomes were changes in self-assessed daily scores of mood, sleep, activity, and medication adherence.

Statistics

Usability, acceptability, and adherence measures were analyzed on completers, who were defined as patients clinically assessed at endpoint visit. Adherence was calculated as a percentage of data entry days in relation to the planned four-week study period, for each data entry parameter (eg, sleep, mood). Mood, sleep, and activity outcomes were analyzed using available data from all included patients. A mean daily mood score was calculated for those with more than one mood entry per day. Correlation between Hamilton scores and mood scores was calculated using the Pearson correlation procedure.

Daily continuous scale scores that were entered, including sleep scores, were analyzed in a random linear regression model using available data from all included patients with *intercept* and *day* as random effects. Results are given as estimated values, confidence limits (CLs), standard errors (SEs), and *P*-values. For explorative analyses on sleep parameters, the model only included time (day) as a covariate. For explorative analyses on mood the model included time (day), sleep-onset, sleep-offset, sleep quality, activity, and interactions between sleep-onset and

day, sleep-offset and day, sleep quality and day, and activity and day. Summary means for sociodemographics, usability, adherence, and acceptability measures are given with standard deviations (SDs). All time points are in the form of hour:minutes.

Based on the paper by Bech et al [26] showing that when comparing MDI depression scores with a VAS scale (0=no depression; 100=worst depression), the cut-off for severe depression on the MDI (scores 31-50, higher scores indicating worse depression) corresponded to a VAS score of 58.4 ($MDI = 0.49 \times VAS + 2.4$). The VAS score used in this study was reversed with 0 as the worst score and 10 as the best, so we translated the VAS score of 58.4 (approximated to 60) to a score of 4 and below as a signifier of severe depression. The level of statistical significance was set at 5%, and was two-sided. Analyses were performed by SAS software (SAS system for Windows, release 9.4., SAS Inst., Cary, NC, USA).

Results

Participants

In total, 230 patients were referred to the IOA in the inclusion period from September 2013 to March 2015. Only 89 patients were asked to participate, as the remaining did not fulfill inclusion criteria or fulfilled exclusion criteria, or were already discharged at the time of screening. A total of 45 patients accepted the invitation to join the study. Sociodemographic data is detailed in Table 1.

Table 1. Sociodemographic data.

Sociodemographics	% or mean (SD)
Gender, females	55.6%
Age, years	35.9 (10.8)
Duration of current episode, months	10.2 (11.7)
Number of previous episodes	2.5 (5.7)
Sick leave in current episode	73.3%
Suicide attempt in current episode	8.9%
Number of patients with a self-perceived eliciting factor for actual episode	84.4%
Smoking	22.2%
Electro Convulsive Treatment in current episode	22.2%
Alcohol consumption, number of drinks per month	8.1 (15.6)

Mean age was 35.9 years (SD 10.8). Patients had 2.5 previous episodes of major depression (SD 5.7) and a mean duration of current depression of 10.2 months (SD 11.7). Most patients were on sick leave. Only a few patients had attempted suicide in the current episode before admission. Patients were treated with 1.9 drugs (SD 0.9; range 0-4): antidepressants included Selective Serotonin Reuptake Inhibitors (n=12), Serotonin-Norepinephrine Reuptake Inhibitors (n=15), Noradrenergic and Specific Serotonergic Antidepressants (n=11), mianserin (n=11), Tricyclic Antidepressants (n=15), isocarboxazid (n=1), and agomelatine (n=1); antipsychotics included quetiapine (n=8) and olanzapine (n=2); mood stabilizers included lithium (n=5) and lamotrigine (n=4); benzodiazepines (n=3); hypnotics (n=3); and melatonin (n=2).

Usability, Acceptability, and Adherence

In total, 76% of enrolled patients (34/45) completed the four-week study; six patients dropped out during study Phase 1, and five patients during study Phase 2. The causes of dropout included worsening of depression for six patients, and miscellaneous nonillness related issues for the remaining five patients. An additional five patients were readmitted to an inpatient ward due to worsening of depression (all in study Phase 2), all of whom continued their self-monitoring and were evaluated at endpoint. Six patients were not discharged on the last day of data entry. Thus, the readmission rate was 13% (5/39). Mean days in study Phase 1 was 6.4 (7.9; range 0-28), study Phase 2 was 7.6 (7.1; range 0-26), and study Phase 3 was 9.9 (9.6; range 0-28; $P=.14$). Electro Convulsive Treatment (ECT) had been used for 22% (10/45) of all patients, 36% (4/11) of the dropouts, and 18% (6/34) of the completers ($P=.23$). Patients who dropped out entered data into Daybuilder for 12.5 days (11.9; range 1-28).

In general, patients found that the Daybuilder system lived up to their expectations; however, when evaluated at endpoint, patients found that they had registered less data than they had anticipated at baseline (Table 2). Fifty-nine percent of patients (20/34) believed that the system could detect a relapse, and 50% of patients (17/34) believed that the system could influence the course of their illness.

Other semiquantitative questions at baseline showed that patients expected the Daybuilder system to enable visualization of their condition, give support and structure, enable positive expectations, or give a hope of recovery. Several additional self-monitoring items were suggested, such as social activity, appetite, meals, anxiety, cognitive function, medication, and side effects. At endpoint, only 50% of patients (17/34) felt that the system had covered their needs for self-monitoring. A total of 33 patients filled in the SUS scale, with a mean value of 86.2 (SD 9.7; range 65-100). Patients not receiving ECT (n=27) had an SUS score of 86.9 (8.9; range 65-100) and patients receiving ECT (n=6) had a score of 82.9 (13.0; range 65-97.5; $P=.40$).

The importance of the weekly phone calls was rated on a scale from 0 to 10 (10=highest importance): 58% of patients (19/33) rated in the interval 8-10, 33% (11/33) in the interval 5-7, and only 9% (3/33) in the interval 0-4. The frequency of phone calls was deemed appropriate by 35% of the patients (12/34). Twenty-six percent of patients (9/34) would have liked more frequent phone calls or a combination of consultation and telephone contact, 12% (4/34) suggested being contacted when the Daybuilder registrations showed signs of deterioration or if data entry was missing, 9% (3/34) felt a need for a more flexible and individualized design that would depend on their mental state, and 18% (6/34) did not answer this question.

Table 2. Usability data from semiquantitative questions asked at baseline and endpoint.

Theme		Baseline, % (n=43)	Endpoint, % (n=34)
Did the Daybuilder system live up to expectations?	Yes		79 (27)
	No		12 (4)
	Uncertain		9 (3)
	No response		0 (0)
Do you think you will be able to make all registrations/did you make all registrations?	Yes	98 (42)	74 (25)
	No	0 (0)	26 (9)
	Uncertain	0 (0)	0 (0)
	No response	2 (1)	0 (0)
Do you expect that self-monitoring of mood will influence your mood/did self-monitoring influence your mood?	Yes	44 (19)	32 (11)
	No	37 (16)	65 (22)
	Uncertain	19 (8)	0 (0)
	No response	0 (0)	3 (1)
Is the need for self-monitoring covered in the Daybuilder system?	Yes	53 (23)	50 (17)
	No	30 (13)	41 (14)
	Uncertain	12 (5)	0 (0)
	No response	5 (2)	9 (3)
Do you expect that the Daybuilder system can detect/did detect a relapse of depression?	Yes	65 (28)	59 (20)
	No	2 (1)	18 (6)
	Uncertain	30 (13)	18 (6)
	No response	2 (1)	6 (2)
Do you expect that the Daybuilder system will influence the course of your illness/did the system influence the course of your illness?	Yes	56 (24)	50 (17)
	No	35 (15)	47 (16)
	Uncertain	2 (1)	0 (0)
	No response	7 (3)	3 (1)

Only 9% of patients (3/34) were worried about technical problems and 79% of patients (27/34) found that the system lived up to their expectations. Thirty-two percent of patients (11/34) felt that their self-assessment of mood in the Daybuilder system had influenced their mood, but only one patient felt that it had a negative effect. Adherence with data entry into the Daybuilder application for the 34 completers was high: mood on 93.8% of the days (872/930), sleep on 89.8% of the days (835/930), activity on 85.6% of the days (796/930), and medication on 88.0% of the days (818/930).

Mood, Sleep, and Activity Outcomes From the Daybuilder System

Table 3 details the self-assessment scores from the Daybuilder system. Self-assessed mood was not significantly improved during the four-week study period ($P=.08$). Additional analyses showed that the frequency of mood scores <4 (equivalent to severe depression), were prevalent in all three phases: 44% in Phase 1, 31% in Phase 2, and 32% in Phase 3. **Figure 2** shows each patient's self-assessed mood scores from the Daybuilder system, with day of discharge inserted (marked DS in the figure), and illustrates the high day-to-day variability. Inspection

of the patients that were readmitted showed no substantial worsening of mood prior to readmission.

Table 3. Estimated self-assessment scores from the Daybuilder system.

Day	Mood score (SE)	Sleep-onset hh:mm (SE, min)	Sleep-offset hh:mm (SE, min)	Sleep-midpoint hh:mm (SE, min)	Sleep Quality score (SE)	Activity minutes (SE)
Day 1	4.7 (0.3)	23:30 (9)	7:42 (9)	3:36 (8)	5.2 (0.2)	156.9 (17.6)
Day 8	4.9 (0.2)	23:38 (9)	7:54 (9)	3:46 (8)	5.4 (0.2)	168.6 (15.0)
Day 15	5.0 (0.3)	23:46 (10)	8:07 (10)	3:56 (10)	5.6 (0.2)	180.2 (15.2)
Day 22	5.1 (0.3)	23:53 (15)	8:19 (12)	4:07 (11)	5.8 (0.2)	191.9 (18.0)
Day 28	5.3 (0.3)	24:00 (14)	8:30 (14)	4:15 (13)	5.9 (0.3)	201.9 (21.8)
Change	0.5 (0.3)	00:29 (10)	00:48 (12)	00:39 (10)	0.7 (0.3)	45.1 (25.6)
<i>P</i> -value	.08	.006	<.001	<.001	.04	.09

Sleep-offset at endpoint was delayed to 48 minutes (SE 12) compared to baseline sleep-offset ($P<.001$; $t=4.0$, CL 24.6-74.7) and was mostly prevalent in Phase 2. The range of sleep-onset was from 19:00 to 06:30, and sleep-offset was from 24:00 to 14:30. Sleep duration was 8:12 (hours:minutes) at baseline and 8:30 (hours:minutes) at endpoint ($P=.10$). Sleep quality was significantly improved during the four-week period from 5.2 (SE 0.2) to 5.9 (SE 0.3; $P=.04$; $t=2.2$; CL 0.05-1.4). Naps were only taken on 6% of study days (56/931). The mean number of awakenings was 1.1 (1.9; range 0-20) per night.

The mean duration of activity was 156.9 minutes at baseline (SE 17.6) and 201.9 minutes at endpoint (SE 21.8; $P=.09$). Explorative analyses on the effect of sleep parameter on self-reported mood showed no significant effect for sleep-onset, sleep-midpoint, or sleep-offset, but sleep quality was significantly positively associated with mood (parameter estimate 0.15, $P<.001$, $t=6.9$; CL 0.11-0.19). Furthermore, when estimating delay of sleep-onset (with sleep quality included in the model) during the study period, this showed a significant

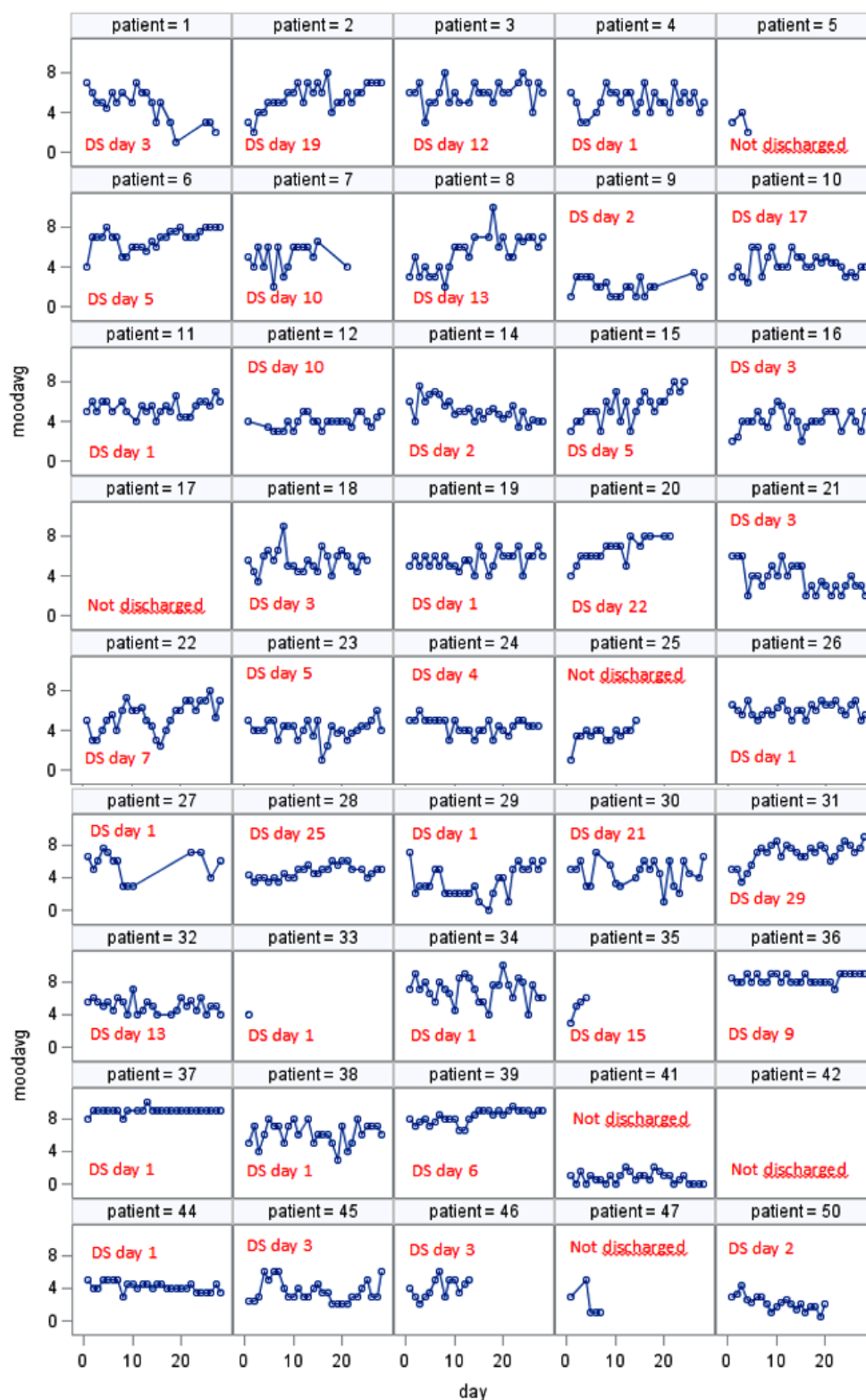
negative effect on mood (combined effect of sleep-onset and the interaction between sleep-onset and day). Thus, a three-hour delay in sleep-onset reduced mood by 0.4 points compared to no delay (SE 0.2; $P=.03$; $t=2.2$; CL 0.04-0.77). Explorative analyses showed no influence of study phase on self-monitored mood and sleep quality ($P=.93$).

Interview and Self-Assessment Scores From Paper-and-Pen Questionnaires

During the study period, a statistically significant reduction in the degree of depression was seen on all depression scales, and an increase in scores was observed on the WHO-5 quality of life scale (Table 4). Correlation between HAM-D₁₇ and self-reported mood was 0.51 ($P<.001$) at baseline, and 0.44 ($P=.02$) at endpoint. Linear regression showed a negative impact of HAM-D₁₇ baseline scores on adherence to sleep and mood registrations (sleep parameters, $R^2=0.18$, $P=.01$; mood scores $R^2=0.10$, $P=.07$).

Table 4. Scores from depression and quality of life scales.

Scale (n)	Baseline n (SD)	Endpoint n (SD)	<i>P</i> -value
Hamilton Depression Rating Scale 17 item version (34)	18.0 (6.5)	13.3 (7.3)	<.001
Hamilton Depression Rating Scale 6 item version (34)	9.9 (3.0)	7.1 (3.7)	<.001
Bech-Rafaelsen Melancholia Scale (34)	17.9 (5.7)	13.3 (7.0)	<.001
Major Depression Inventory (33)	27.1 (13.1)	22.2 (12.7)	.006
5-item World Health Organization Well-Being Index (32)	31.3 (22.9)	43.4 (22.1)	<.001

Figure 2. Individual daily self-assessed mood scores with day of discharge (DS) inserted.

Discussion

We expected the Daybuilder system to be manageable for this patient category, and that there would be great day-to-day variability in mood and sleep; both of these expectations were confirmed. Based on prior results [11] we expected that a subset

of patients would feel that assessment of mood would worsen their mood. However, this was not the case in our study, as only one patient had this experience.

Principal Results

Based on user reviews, the usability, acceptability, and adherence of the Daybuilder system were found to be good.

Sleep onset and offset varied greatly from day-to-day, and the sleep-wake cycle was delayed significantly from baseline to endpoint. Subjective mood scores entered in the Daybuilder system also varied greatly from day-to-day, and demonstrated no improvement from baseline to endpoint.

Considering the short study period, a dropout rate of 24% (11/45) in a four-week study period must be considered as large, and was mainly caused by worsening of depression, although the present study did not reveal the exact causes for terminating self-monitoring. Due to IOA's referral rules, this patient group is expected to be more severely depressed and unstable than the typical inpatient with depression, and from this background some dropout is to be expected. Development in this area should focus on mechanisms that will keep patients from dropping out, especially when experiencing a deterioration of mood. This goal could be accomplished by implementing alarm systems activated at mood drops, or when assessments are missing for more than one day.

The secondary outcomes showed that patients' sleep drifted to later in the day after discharge, and that this drift was associated with worsening of mood. This finding, along with high readmittance rates, calls for an improved system with clinician responses based on visual inspection of daily assessments. The results from this study have prompted us to develop a randomized study that focuses on mood scores and also aims to prevent sleep drift by observing sleep data on a daily basis, and contacting patients in cases of sleep drift or sleep irregularity.

Limitations

Patients referred to IOA probably belong to a more severely depressed subset of inpatients, and thus do not reflect the general group of inpatients with depression. We do not believe that the inclusion and exclusion criteria had any major impact on patients that were included into the study. The most common reason for noninclusion was due to patients being discharged before we could inform them of the project.

The study design, with a single arm, fails to determine the effects that the Daybuilder system has on depression. This information would require a randomized controlled trial. We cannot know from our data whether an active clinician intervention could have prevented dropout and readmission. This approach would require incorporation of timely and active clinician help, and necessitate that patient data is coupled with an automatic system that alerts clinicians when deterioration is detected, and/or that clinicians view patient data daily. Due to the low sample size, negative results could easily be due to lack of power and randomized studies in this field should include far larger sample sizes. Conversely, we cannot rule out spurious positive findings.

ECT treatment is associated with a clinically recognizable retrograde and anterograde amnesia, but this did not seem to influence the use of the Daybuilder system in a significant way. This finding is possibly a tribute to the user-friendliness of the Daybuilder system. The present study design did not allow clinicians to access any day-to-day assessment of patients' data, meaning that a worsening of patients' conditions could only be detected once a week, in relation to the planned telephone

consultation. This limitation puts the need for clinical intervention into perspective.

Patients were instructed to enter morning and evening mood scores to assess diurnal variation. However, not all patients succeeded in this; thus, mood data reflects different time points (morning or evening) and some of the day-to-day variation is probably caused by diurnal variation. The registered difficulty in entering a mood score more than once per day is probably due to the way that mood data was entered into the Daybuilder system, which used tabs inside the user interface. Caution should also be taken when interpreting mean mood scores, as patients who dropped out most likely experienced a deterioration of mood, thus inducing a bias.

The low correlation between self-assessed mood and HAM-D₁₇ scores at baseline (0.51; $P<.001$) and endpoint (0.44; $P=.02$) points to patients reporting aspects of their illness with a different content than the items covered by the Hamilton score. Postpublication analyses from our earlier study, using a similar depression VAS scale (Preskorn) also showed a low correlation (Spearman) of 0.51 ($P<.001$) with HAM-D₁₇ scores [27]. However, this finding does not imply that self-assessed mood scores are less valuable than the Hamilton scores, but only that self-assessed scores report other aspects of the depressive illness (such as negative and positive affect), and are possibly more akin to a patient's own experiences of their condition. We must also consider whether a higher daily mood sampling frequency would be better to track mood fluctuation.

The results from sleep scores confirm that patients with depression had a dysregulated sleep-wake cycle, with large day-to-day variations and a substantial delay in sleep timing when discharged. The finding that a delay in sleep-onset had a negative impact on mood points to the possibility that an intervention to prevent sleep delay could improve mood and theoretically prevent relapse.

It was a surprise that patients' self-assessed mood scores did not significantly deteriorate in the days after discharge. It was also unexpected that the self-assessed mood scores for those that were readmitted did not deteriorate in the days preceding readmission. These two results call for speculation on whether the mood assessment that was used should be supplemented with assessments more aimed at patient security, such as monitoring of suicidal ideation or by using depression scales. This consideration has prompted us to change our upcoming study, which also uses electronic self-monitoring, such that the wording of self-assessed *mood* is replaced by self-assessed *depression* severity. Through communication with the patients, we will aim at a common understanding of the word *depression*. The Daybuilder system could be improved if it is developed as an app for smartphones or tablets. This progression would eliminate the need for patients to be near a computer, and make data entry more flexible.

Conclusions

In conclusion, patients were satisfied with the Daybuilder system and study completers had a high adherence to the Daybuilder application. The dropout rate and the gradual delay in sleep emphasize the need for continued clinical support for patients

discharged from psychiatric wards, especially concerning sleep guidance.

Improvement of the current Daybuilder system could be done using daily clinician monitoring and daily responses to data entry. This approach could be restricted to patients with suspected suicidal risk or high risk of relapse, such as patients treated with ECT [28]. Additionally, alarms connected to mood ratings could be incorporated into the software to alert clinicians

to patients that are deterioration or not registering data. Other improvements include: development of a smartphone app, the use of chat systems, and SMS options to facilitate communication between patients and clinicians. Such an improved full version would add to patient empowerment and autonomy [29,30]. The present study, focusing on usability, highlights the need for interaction between the clinician and the patient.

Acknowledgments

All authors contributed substantially to the conception, design, and drafting of the article, and approved the submitted version. We would like to thank the following people for their contributions to the study: all patients for their contributions, and all staff from the inpatients' wards at the Mental Health Center Copenhagen referring patients to the study; and Lene Theil Skovgaard, Associate Professor (Department of Public Health, University of Copenhagen, Denmark) for statistical support.

The principal investigator, Klaus Martiny, has full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analyses. The study was supported by an unrestricted grant from TrygFonden. Trygfonden was not involved in any part of the project.

Conflicts of Interest

Lasse Benn Nørregaard and Philip Kaare Løventoft are directors of Daybuilder Solutions, were involved in the development of the study version of the Daybuilder system that was used, and have assisted with technical support during the study. Lasse Benn Nørregaard and Phillip Løventoft have not been involved in any clinical interaction or assessment of participants in the study. Lars Vedel Kessing was a consultant for Lundbeck and Astra Zeneca within the last three years. All other authors report no conflicts of interest.

References

1. Olesen J, Gustavsson A, Svensson M, Wittchen H, Jönsson B, CDBE2010 SG, et al. The economic cost of brain disorders in Europe. *Eur J Neurol* 2012 Jan;19(1):155-162. [doi: [10.1111/j.1468-1331.2011.03590.x](https://doi.org/10.1111/j.1468-1331.2011.03590.x)] [Medline: [22175760](https://pubmed.ncbi.nlm.nih.gov/22175760/)]
2. Kessing LV, Hansen MG, Andersen PK. Course of illness in depressive and bipolar disorders. Naturalistic study, 1994-1999. *Br J Psychiatry* 2004 Nov;185:372-377 [FREE Full text] [doi: [10.1192/bjp.185.5.372](https://doi.org/10.1192/bjp.185.5.372)] [Medline: [15516544](https://pubmed.ncbi.nlm.nih.gov/15516544/)]
3. Kessing LV, Hansen HV, Hvenegaard A, Christensen EM, Dam H, Gluud C, Early Intervention Affective Disorders (EIA) Trial Group. Treatment in a specialised out-patient mood disorder clinic v. standard out-patient treatment in the early course of bipolar disorder: randomised clinical trial. *Br J Psychiatry* 2013 Mar;202(3):212-219 [FREE Full text] [doi: [10.1192/bjp.bp.112.113548](https://doi.org/10.1192/bjp.bp.112.113548)] [Medline: [23349295](https://pubmed.ncbi.nlm.nih.gov/23349295/)]
4. Hansen HV, Christensen EM, Dam H, Gluud C, Wetterslev J, Kessing LV, Early Intervention Affective Disorders (EIA) Trial Group. The effects of centralised and specialised intervention in the early course of severe unipolar depressive disorder: a randomised clinical trial. *PLoS One* 2012;7(3):e32950 [FREE Full text] [doi: [10.1371/journal.pone.0032950](https://doi.org/10.1371/journal.pone.0032950)] [Medline: [22442673](https://pubmed.ncbi.nlm.nih.gov/22442673/)]
5. Qin P, Nordentoft M. Suicide risk in relation to psychiatric hospitalization: evidence based on longitudinal registers. *Arch Gen Psychiatry* 2005 Apr;62(4):427-432. [doi: [10.1001/archpsyc.62.4.427](https://doi.org/10.1001/archpsyc.62.4.427)] [Medline: [15809410](https://pubmed.ncbi.nlm.nih.gov/15809410/)]
6. El-Gayar O, Timsina P, Nawar N, Eid W. A systematic review of IT for diabetes self-management: are we there yet? *Int J Med Inform* 2013 Aug;82(8):637-652. [doi: [10.1016/j.ijmedinf.2013.05.006](https://doi.org/10.1016/j.ijmedinf.2013.05.006)] [Medline: [23792137](https://pubmed.ncbi.nlm.nih.gov/23792137/)]
7. Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE, Kessing LV. Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: a systematic review of the validity and evidence. *BMC Psychiatry* 2016;16:7 [FREE Full text] [doi: [10.1186/s12888-016-0713-0](https://doi.org/10.1186/s12888-016-0713-0)] [Medline: [26769120](https://pubmed.ncbi.nlm.nih.gov/26769120/)]
8. Wade AG, Häring J. A review of the costs associated with depression and treatment noncompliance: the potential benefits of online support. *Int Clin Psychopharmacol* 2010 Sep;25(5):288-296. [Medline: [20715299](https://pubmed.ncbi.nlm.nih.gov/20715299/)]
9. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
10. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. *J Ment Health* 2011 Dec;20(6):509-524. [doi: [10.3109/09638237.2011.608746](https://doi.org/10.3109/09638237.2011.608746)] [Medline: [21988230](https://pubmed.ncbi.nlm.nih.gov/21988230/)]
11. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. *Psychol Med* 2015 Oct;45(13):2691-2704. [doi: [10.1017/S0033291715000410](https://doi.org/10.1017/S0033291715000410)] [Medline: [26220802](https://pubmed.ncbi.nlm.nih.gov/26220802/)]

12. Espinosa HD, Carrasco A, Moessner M, Cáceres C, Gloger S, Rojas G, et al. Acceptability study of “Ascenso”: an online program for monitoring and supporting patients with depression in Chile. *Telemed J E Health* 2016 Jul;22(7):577-583. [doi: [10.1089/tmj.2015.0124](https://doi.org/10.1089/tmj.2015.0124)] [Medline: [26741190](https://pubmed.ncbi.nlm.nih.gov/26741190/)]
13. 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](https://doi.org/10.1016/j.cct.2013.08.005)] [Medline: [23974036](https://pubmed.ncbi.nlm.nih.gov/23974036/)]
14. Papakostas GI. Cognitive symptoms in patients with major depressive disorder and their implications for clinical practice. *J Clin Psychiatry* 2014 Jan;75(1):8-14. [doi: [10.4088/JCP.13r08710](https://doi.org/10.4088/JCP.13r08710)] [Medline: [24345473](https://pubmed.ncbi.nlm.nih.gov/24345473/)]
15. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
16. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? *Clin Psychol Rev* 2011 Feb;31(1):89-103. [doi: [10.1016/j.cpr.2010.09.008](https://doi.org/10.1016/j.cpr.2010.09.008)] [Medline: [21130939](https://pubmed.ncbi.nlm.nih.gov/21130939/)]
17. 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](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
18. Løventoft PK, Nørregård LB, Frøkjær E. Designing Daybuilder: an experimental app to support people with depression. 2012 Jun 15 Presented at: Proceedings of the 12th Participatory Design Conference exploratory Papers, Workshop Descriptions, Industry Cases - Volume 2. Association for Computing Machinery; June 15, 2012; Roskilde University, Denmark p. a-4.
19. 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. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
20. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960 Feb;23:56-62 [FREE Full text] [Medline: [14399272](https://pubmed.ncbi.nlm.nih.gov/14399272/)]
21. Bobes J, Bulbena A, Luque A, Dal-Re R, Ballesteros J, Ibarra N, et al. The sufficiency of the HAM-D6 as an outcome instrument in the acute therapy of antidepressants in the outpatient setting. *Int J Psychiatry Clin Pract* 2007;11(2):146-150. [doi: [10.1080/13651500600971646](https://doi.org/10.1080/13651500600971646)] [Medline: [24937560](https://pubmed.ncbi.nlm.nih.gov/24937560/)]
22. Licht RW, Qvitzau S, Allerup P, Bech P. Validation of the Bech-Rafaelsen Melancholia Scale and the Hamilton Depression Scale in patients with major depression; is the total score a valid measure of illness severity? *Acta Psychiatr Scand* 2005 Feb;111(2):144-149. [doi: [10.1111/j.1600-0447.2004.00440.x](https://doi.org/10.1111/j.1600-0447.2004.00440.x)] [Medline: [15667434](https://pubmed.ncbi.nlm.nih.gov/15667434/)]
23. Bech P, Christensen E, Vinberg M, Østergaard S, Martiny K, Kessing L. The performance of the revised Major Depression Inventory for self-reported severity of depression--implications for the DSM-5 and ICD-11. *Psychother Psychosom* 2013;82(3):187-188. [Medline: [23548877](https://pubmed.ncbi.nlm.nih.gov/23548877/)]
24. Bech P. Measuring the dimensions of psychological general well-being by the WHO-5. *QoL Newsletter* 2004;32:15-16.
25. Tullis T, Albert B. Measuring the user experience, 2nd edition. Waltham, MD: Morgan Kaufmann; 2013.
26. Bech P, Timmerby N, Martiny K, Lunde M, Søndergaard S. Psychometric evaluation of the Major Depression Inventory (MDI) as depression severity scale using the LEAD (Longitudinal Expert Assessment of All Data) as index of validity. *BMC Psychiatry* 2015 Aug 05;15:190 [FREE Full text] [doi: [10.1186/s12888-015-0529-3](https://doi.org/10.1186/s12888-015-0529-3)] [Medline: [26242577](https://pubmed.ncbi.nlm.nih.gov/26242577/)]
27. Martiny K, Refsgaard E, Lund V, Lunde M, Sørensen L, Thougard B, et al. A 9-week randomized trial comparing a chronotherapeutic intervention (wake and light therapy) to exercise in major depressive disorder patients treated with duloxetine. *J Clin Psychiatry* 2012 Sep;73(9):1234-1242. [doi: [10.4088/JCP.11m07625](https://doi.org/10.4088/JCP.11m07625)] [Medline: [23059149](https://pubmed.ncbi.nlm.nih.gov/23059149/)]
28. Martiny K, Larsen E, Licht R, Nielsen C, Damkier P, Refsgaard E, Danish University Antidepressant Group (DUAG*). Relapse prevention in major depressive disorder after successful acute electroconvulsive treatment: a 6-month double-blind comparison of three fixed dosages of escitalopram and a fixed dose of nortriptyline - lessons from a failed randomised trial of the Danish University Antidepressant Group (DUAG-7). *Pharmacopsychiatry* 2015 Nov;48(7):274-278. [doi: [10.1055/s-0035-1565063](https://doi.org/10.1055/s-0035-1565063)] [Medline: [26529118](https://pubmed.ncbi.nlm.nih.gov/26529118/)]
29. Simons CJ, Hartmann JA, Kramer I, Menne-Lothmann C, Höhn P, van Bommel AL, et al. Effects of momentary self-monitoring on empowerment in a randomized controlled trial in patients with depression. *Eur Psychiatry* 2015 Nov;30(8):900-906. [doi: [10.1016/j.eurpsy.2015.09.004](https://doi.org/10.1016/j.eurpsy.2015.09.004)] [Medline: [26647864](https://pubmed.ncbi.nlm.nih.gov/26647864/)]
30. Välimäki M, Kurki M, Hätönen H, Koivunen M, Selander M, Saarijärvi S, et al. Developing an internet-based support system for adolescents with depression. *JMIR Res Protoc* 2012;1(2):e22 [FREE Full text] [doi: [10.2196/resprot.2263](https://doi.org/10.2196/resprot.2263)] [Medline: [23612485](https://pubmed.ncbi.nlm.nih.gov/23612485/)]

Abbreviations

CL: confidence limit
ECT: Electro Convulsive Treatment
HAM-D: Hamilton depression rating scale
IOA: Intensive Outpatient Unit for Affective Disorders
MDI: Major Depression Inventory

MES: melancholia scale
SD: standard deviation
SE: standard error
SMS: short message service
SUS: System Usability Scale
VAS: Visual Analog Scale
WHO-5: 5-item World Health Organization Well-Being Index

Edited by G Eysenbach; submitted 22.09.16; peer-reviewed by M Faurholt-Jepsen, C Simons; comments to author 13.10.16; revised version received 21.12.16; accepted 11.02.17; published 21.04.17.

Please cite as:

Lauritsen L, Andersen L, Olsson E, Søndergaard SR, Nørregaard LB, Løventoft PK, Svendsen SD, Frøkjær E, Jensen HM, Hageman I, Kessing LV, Martiny K
Usability, Acceptability, and Adherence to an Electronic Self-Monitoring System in Patients With Major Depression Discharged From Inpatient Wards
J Med Internet Res 2017;19(4):e123
URL: <http://www.jmir.org/2017/4/e123/>
doi: [10.2196/jmir.6673](https://doi.org/10.2196/jmir.6673)
PMID: [28432040](https://pubmed.ncbi.nlm.nih.gov/28432040/)

©Lise Lauritsen, Louise Andersen, Emilia Olsson, Stine Rauff Søndergaard, Lasse Benn Nørregaard, Philip Kaare Løventoft, Signe Dunker Svendsen, Erik Frøkjær, Hans Mørch Jensen, Ida Hageman, Lars Vedel Kessing, Klaus Martiny. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 21.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Gender-Specific Determinants and Patterns of Online Health Information Seeking: Results From a Representative German Health Survey

Eva Baumann¹, PhD; Fabian Czerwinski¹, MSc; Doreen Reifegerste¹, PhD

Hanover Center for Health Communication, Department of Journalism and Communication Research, Hanover University of Music, Drama, and Media, Hanover, Germany

Corresponding Author:

Doreen Reifegerste, PhD

Hanover Center for Health Communication

Department of Journalism and Communication Research

Hanover University of Music, Drama, and Media

Expo Plaza 12

Hanover, 30539

Germany

Phone: 49 (05 11) 31 00 ext 448

Fax: 49 (05 11) 3100400

Email: doreen.reifegerste@ijk.hmtm-hannover.de

Abstract

Background: Online health information-seeking behavior (OHISB) is currently a widespread and common behavior that has been described as an important prerequisite of empowerment and health literacy. Although demographic factors such as socioeconomic status (SES), age, and gender have been identified as important determinants of OHISB, research is limited regarding the gender-specific motivational determinants of OHISB and differences between women and men in the use of online resources for health information purposes.

Objective: The aim of this study was to identify gender-specific determinants and patterns of OHISB by analyzing data from a representative German sample of adults (N=1728) with special attention to access and frequency of use as well as topics and sources of OHISB.

Methods: We employed a 2-step analysis, that is, after exploring differences between users and nonusers of online health information using logistic regression models, we highlighted gender-specific determinants of the frequency of OHISB by applying zero-truncated negative binomial models.

Results: Age (odds ratio, OR for females=0.97, 95% CI 0.96-0.99) and degree of satisfaction with one's general practitioner (GP) (OR for males=0.73, 95% CI 0.57-0.92) were gender-specific determinants of access to OHISB. Regarding the frequency of OHISB, daily Internet use (incidence rate ratio, IRR=1.67, 95% CI 1.19-2.33) and a strong interest in health topics (IRR=1.45, 95% CI 1.19-1.77) were revealed to be more important predictors than SES (IRR for high SES=1.25, 95% CI 0.91-1.73).

Conclusions: Users indicate that the Internet seems to be capable of providing a valuable source of informational support and patient empowerment. Increasing the potential value of the Internet as a source for health literacy and patient empowerment requires need-oriented and gender-specific health communication efforts, media, and information strategies.

(*J Med Internet Res* 2017;19(4):e92) doi:[10.2196/jmir.6668](https://doi.org/10.2196/jmir.6668)

KEYWORDS

health information seeking; social media; gender differences; frequency of seeking; Internet

Introduction

The Relevance of Health Information Seeking

Patients today are increasingly challenged to make informed choices regarding their health care and to play an active role in health-related decisions [1,2], a change which has been described as *empowerment* [3]. However, a relevant precondition of empowerment is health literacy (ie, the skills and competencies to find and evaluate health information [4-6]). In contrast to the established body of research concerning the concept of health literacy and corresponding measures—which includes a constantly growing body of evidence on the determinants of health literacy and the programs that are effective in enhancing it—research on health information seeking, a behavior closely linked to health literacy, has just begun to evolve over the past few years [7]. The dearth of research in this area is all the more astonishing as health information seeking behavior is known to have a strong influence on health-related behavioral intentions, decisions, and outcomes [8].

The Internet represents an increasingly important source of health information [3,9,10], and health is one of the most common topics in online information seeking [9]. Although there is increasing health information available [11], some populations do not sufficiently benefit from the available resources of health information, due to limited access or low media literacy [12]; such population-specific differences raise the risk of increasing health inequalities, commonly referred to as the “digital health divide” [13,14]. In order to improve access to health information [6], we not only need comprehensive research on the gap between *health-onliners* (people who use the Internet to search for health information) and *health-offliners* (people who use channels other than the Internet to search for health information), but also information regarding the determinants of the frequency with which people seek health information online. Although online health information seeking has been analyzed mostly as a binary yes-no outcome, such studies have made only limited contributions to the research about the determinants of the frequency of seeking. Frequency of online health information-seeking behavior (OHISB) is becoming increasingly important as more people use the Internet [9], and an in-depth analysis of the major determinants and outcomes of OHISB is needed.

In light of this clear need—and in parallel to the discourse on the digital divide [14]—the focus of research on OHISB shifts from formerly relevant questions of access to and availability of mere technology toward a deeper understanding of usage frequencies, including demographic, motivational, and health-related factors influencing the frequency of OHISB [15]. Among these factors, gender differences have been frequently reported as relevant for OHISB and health outcomes [9,16], but little is known about the underlying reasons for such differences. Apart from gender differences in general Internet usage [17], reasons for gender differences in OHISB might include the existence of different patterns concerning topics and sources of health information seeking [18] or the lower interest of men in health: Because men tend to be comparably less willing and

motivated to engage with health topics [19], they might search for online health information less frequently than do women.

It has been argued that gender differences in OHISB might be concealed by differing motives for seeking health information: Whereas women are more interested in health issues and emotional support, men are more interested in informational support [20]. Men’s higher interest in and earlier acceptance of technology [21] has also resulted in higher mHealth adoption intentions compared with women [22]. Gender differences have also been reported in mobile phone gratifications [23], social media usage [17], and activity in social support groups [20].

Our aim was therefore to understand gender-specific determinants and patterns of OHISB. This understanding will allow us to gain insight into gender-specific preferences regarding content and sources, and to draw conclusions regarding gender-specific targeting strategies for the development of health-related online media. To date, no representative data on gender-specific OHISB for Germany has been analyzed using multivariate statistics [24], making this paper the first such contribution. Our research for this paper investigated the correlates of health-related online information seeking with special regard to gender differences, conducting a secondary analysis of the German Bertelsmann Health Care Monitor 2015. We conclude with a discussion of implications with regard to health communication theory and practice.

Theories of Health Information Seeking

The models that are frequently used to explain health information seeking—such as the theory of planned behavior (TPB [25]), the theory of motivated information management [26], the risk perception attitude framework [27], or the model of risk information seeking and processing [28]—primarily concentrate on psychological variables (eg, risk perception, subjective norms, control beliefs, or personal experience) or content criteria of the media as determinants of health information-seeking behavior (HISB) [29,30]. As a result, these models and the studies referring to them neglect the direct impact of gender on health information seeking, as well as the related reasons underlying this effect.

Although research has shown that females are more likely to conduct HISB than males, integration of this finding into theory is still lacking [31]. For example, the TPB includes gender as a relevant external variable that influences intentional and attitude-related processes, but the model does not specify the influence of gender [32]; as a result, the TPB only allows researchers to draw limited conclusions about gender-specific health information strategies. One possible explanation for gender-specific patterns of HISB might be found in social role theory [33], which posits that whereas the male gender role casts men as agentic (ie, task-oriented), women are expected to be more socially engaged, with activities such as staying in contact with family members or receiving understanding and feedback from others [23]. These different social roles may contain gender-specific health-related tasks such as care for children or elderly family members, which are in turn associated with an increased demand for health information [12]. These different social roles have also been associated with different use of media channels [17,34] and might be related to different

goals for OHISB. In addition to the sociocultural theory, Meyers-Levy and Loken [35] described evolutionary theory, hormone exposure of the brain, and selectivity hypothesis as further theoretical approaches to explain gender differences. These more biological approaches might be especially relevant to explain differences between men and women in the seeking and provision of social support [20], technical affinity [21], motivations for mediated communication [17], and information processing strategies [35].

To date, there has been only minor exploration as to if and how these differences in motives and channel usage are also relevant for online health information seeking. Gaining a better understanding of gender differences in HISB would help health communication scholars to develop gender-specific health communication interventions. Regarding such considerations, our analysis may contribute to the iterative junction of theoretical approaches on HISB and gender differences.

Gender Differences in General Internet Use and Frequency

Today, the vast majority of the population across Europe and North America has access to the Internet [36], including 86.2% of German residents. Due to the ever-decreasing proportion of Internet nonusers, the discussion on differences between users and nonusers has shifted from access to skills [37]. In general, differences in Internet usage are consistently reported to highly depend on education, age, and gender [38], and socioeconomic status (SES)—a combined measure of education, income, and social position—is strongly correlated with both frequency and patterns of Internet use [39]. Age has been found to be negatively associated with both Internet access and frequency of use [40,41].

Regarding gender differences, findings are somewhat inconsistent, that is, no significant differences in general Internet use have been detected in the United States [40]. The same is true for many other similarly developed countries (eg, Sweden, Norway, the United Kingdom) across Europe, where only minor differences have been found [42]. However, a higher proportion of men than women in Germany reported using the Internet “at least occasionally” (83.0% vs 76.0%, respectively) or “daily” (68.3% vs 58.0%, respectively) [43]; similar numbers have been reported from some other European countries (eg, Austria, Italy, Switzerland), although the significance of these differences has not yet been determined [42]. In addition, gender differences in Internet use might be interacting with age. Although they are evident in older cohorts, they tend to be smaller in younger age groups [7,44]. Men and women seem to differ in both their motivations for and utilization of multiple forms of mediated online communication [17], referring both to topics they search for and to the ways they communicate. These results indicate that women, compared with men, prefer and more frequently engage in interpersonal communication online, using tools such as social networking sites to maintain relationships [45].

Online Health Information-Seeking Behavior

The Internet’s already-prominent role in HISB continues to increase, that is, in the United States, 59% of the adult population (ie, more than 72% of adult Internet users) seeks

online information concerning health topics [9], and the numbers for Germany and other European countries are similar [46]. The Internet is such a popular source of health information primarily because it is an active information channel with a wide range of information on health content, health communities, and health provision [47]. The Internet as a health information source has been found to be especially important for people suffering from chronic diseases [15,48] and for those who are newly diagnosed with a medical condition or health problem [49]. For them, OHISB is a way to obtain more in-depth information, as well as a way to seek out support and contact with other people affected by the same medical condition or diagnosis. This access to social, informational, and emotional support on specific topics then empowers people to manage their health and to take a more active role when interacting with their physicians [3,50,51]. However, significant disparities still exist regarding access to and the ability to process health information online, with older and less educated people being less likely to take advantage of this resource [2,16,52,53].

Gender differences have not only been reported for general Internet use, but also for general health-related behaviors and outcomes, with men having higher mortality and morbidity rates, engaging in more risky behaviors (eg, smoking, alcohol abuse), and taking part in fewer health-promoting behaviors than women [54]. Men also tend to underestimate their health risks, which can lead to avoidance and reactance toward traditional risk information messages; however, despite these differences, little is known about effective gender-specific health communication strategies [55].

With regard to gender-specific HISB, many studies show that women are more engaged in health information seeking in general, as well as on the Internet, specifically. Being female is among the strongest predictors of conducting OHISB [15,48,49]. Whereas women report to be more interested in health information and show more active search activities [56], men are less likely to read health information [57]. This gender gap in OHISB was found to be stable over time when analyzing six waves of Health Information National Trends Survey (HINTS)—data from 2002 to 2013 [31], and was also proven in a robust meta-analysis on US adults [7]; in contrast, recent research from a German sample indicated only a minor gender gap in frequency of OHISB, which did not reach statistical significance [24]. Results have also been inconsistent regarding the channels utilized: Bidmon and Terlutter [24] found that women used health forums, blogs, and search engines as sources more frequently than men, whereas men used apps for OHISB more frequently. In contrast, other studies have reported that men use health-related apps [58] and track health-related indicators as often as women do [9].

Our main goal was to analyze gender-specific determinants and patterns of OHISB. Our first question was which sociodemographic—including gender—and health-related user characteristics explain general utilization of health information on the Internet (RQ1). The second question was whether—among those who use the Internet for health-related purposes—the same factors determine the frequency of OHISB; to address this question, we ask which sociodemographic and

health-related user characteristics—in relation to gender—explain the frequency of OHISB (RQ2).

Among the health-onliners, we are also interested in the gender-specific health-related topics they are most interested in and the online media they prefer to use as sources of health information (RQ3).

Methods

Data Collection and Sample Size

Data were taken from the Bertelsmann Health Care Monitor 2015, a representative national German health survey (available as open access files) conducted by the Bertelsmann Foundation in cooperation with the Barmer GEK, a statutory health insurance (see [59] for data and further information). This survey assesses health-related knowledge, attitudes, and behaviors, and is similar to the annual American HINTS [12] or the Pew polls [9]. The Bertelsmann Health Care Monitor has been conducted annually via mail since 2001. Its average response rate is about 70%, and it has been established as an important data source in the field of health research, with a number of key publications based on it [60]. The basic population for the survey consists of persons living in private households, aged 18-79 years, in Germany. Samples are drawn from the GfK (German Association for Consumer Research) Mail Panel and the sample is representative concerning gender, age, the German federal states, income, and education compared with the data of the Statistical Yearbook of Germany [61,62]. Although the 2015 Bertelsmann Health Care Monitor comprised 1728 German adults aged 18-79 years, for this study, we excluded respondents who reported not to use the Internet at all, since our main interest is to explain OHISB. Thus, the remaining sample size was $N=1219$.

The excluded Internet nonusers showed statistically significant differences for several demographic criteria: They were much older (mean 64.7 years, SD 12.1) than the Internet users (mean 46.0, SD 15.0) with a higher proportion of female respondents (59.9% [299/499] vs 50.04% [610/1219]) and lower SES (31.2% [150/481] vs 15.86% [177/1116]). These findings confirmed prior research concerning demographic determinants of general Internet use [39,40,43].

Measures

Online Health Information Seeking Behavior (OHISB)

Our main analyses were based on participant responses to the question “How many times did you use the Internet for seeking health information within the last 12 months?” Answers ranged from 0 to 130 with a mean of 4.37 (SD 9.44); answers were strongly right skewed (skewness=5.85, SE 0.07). To address our first analytical goal of uncovering the gender-specific determinants of utilization of OHISB, we created a dummy variable to separate health-offliners (OHISB=0) from health-onliners (OHISB \geq 1). To meet our second goal of assessing gender-specific determinants of the frequency of OHISB, we left the responses on their original scale but excluded the health-offliners from the analysis, as they showed

no variance in their HISB frequency. This resulted in a final sample of 643 health-onliners.

The third objective—assessing gender differences in health-related topics and information sources—was achieved by analyzing the frequencies of the health-related topics and websites the health-onliners used. Respondents were asked to select the topics on which they searched for or received information from a list of 14 items. These items ranged from very specific (eg, “drugs and their pharmacological interactions”) to more general (eg, “fitness, well-being”) topics. These items were then grouped into three categories by content: “disease and health care,” “health care policy and health care system,” and “health and well-being.”

Respondents were then asked to select the sources they used when conducting OHISB, that is, they were given the 10 items to choose from popular sources (eg, “online dictionary”) and more specific sources (eg, “websites of noncommercial health organizations”).

Predictor Variables of Online Health Information Seeking Behavior (OHISB)

Demographic Variables

Participants were asked to provide their age in years and gender (female or male), whereas SES was assessed by summing up participants’ responses on their education, occupation, and income (weighted by household size) to a score ranging between 3 and 27 following the standard procedure for the Health Care Monitor [61]. Due to the application of conventional formats, data on SES were transformed from the original 27-point scale to a 3-point scale analog indicating “low,” “medium,” and “high” SES.

Variable Related to General Internet Use

The frequency of general Internet usage was measured using a 3-point ordinal scale of “at least sometimes per month,” “several times per week,” or “daily.”

Health-Related Variables

Patient status was measured using a 4-point scale ranging from 1 (“currently not affected”) to 4 (“chronically ill”). We classified the responses from “mildly or not affected” to “severely or chronically ill,” because OHISB patterns of healthy and mildly affected respondents should be quite similar, whereas severely or chronically affected people were expected to show fundamentally different patterns. Participants’ perceived relevance of understanding somatic processes, their health-consciousness, and the satisfaction with their GP were all measured using 5-point scales, that is, to measure health-consciousness, participants were asked how much attention they generally paid to their health, which they rated from 1 (“Generally, I don’t take care of my health”) to 5 (“Generally, I take good care of my health”). Satisfaction with their GP was scored from 1 (“very dissatisfied”) to 5 (“very satisfied”), their perceived relevance of understanding somatic processes was assessed by their degree of agreement—from 1 (“totally disagree”) to 5 (“totally agree”)—concerning the statement that “patients diagnosed with an illness should understand exactly what is going on.”

The extent to which respondents reported being interested in information concerning health topics in general was originally measured on a 3-point scale indicating weak, medium, or strong interest. We transferred these answers into a dummy variable, contrasting “low or medium level of interest” with “high level of interest” to create reasonably equal group sizes ($n=739$ and $n=450$, respectively). Looking at health-onliners only, their motivations to conduct HISB were assessed using 12 dummy indicators covering a broad range of potential goals (eg, “to find general health information about health risks and diseases” or “determining the best treatment options”). On the basis of social support theory [63], we categorized these 12 indicators into 3 indices representing aspects of “esteem support” (5 items), “informational support” (5 items), and “emotional support” (2 items). The more items participants agreed within each index, the higher their score (one point per item).

Items are given in [Multimedia Appendix 1](#).

Statistical Analysis

To answer RQ1, a logistic regression model was conducted to analyze the influence of sociodemographic, motivational, and health-related factors on differences between health-onliners and health-offliners. Regarding RQ2, Poisson regression models are traditionally used to model data like the frequency of OHISB, as such models are suited to fulfilling the technical needs of an outcome consisting of positive integers. However, the application of Poisson models requires a data structure that is seldom found in reality, that is, the mean is equal to the variance [64,65]. As the variance in real data is often much bigger than the mean, “overdispersion” tends to occur, leading to biased variance estimates and associated inferential problems [66,67]. This was certainly the case for our data, as the variance (136.9) is about 16 times greater than the mean (8.32), indicating severe problems due to strong overdispersion. Furthermore, since we excluded all health-offliners (with an OHISB frequency=0), our data contain no zeros, and the application of

a standard negative binomial model—which tries to predict zeros—should therefore be avoided [68]. We therefore conducted zero-truncated negative binomial regression models to explain the frequency of health information searching. Missing values were deleted listwise for all multivariate analyses. All analyses were conducted using SPSS 22 (IBM Corporation) except the zero-truncated negative binomial models, which were estimated using Stata 11.2 (Stata Corp LLP).

Results

Sample Description

Among all of the 1219 participants who used the Internet, 643 (52.75%; *health-onliners*) searched for health information online and 576 did not (47.25%; *health-offliners*; see [Table 1](#)). The health-offliners showed no significant differences from health-onliners regarding age and gender, but significantly fewer health-offliners had high SES ($P=.001$). Health-onliners rated their own health status more often as “chronically or severely ill” ($n=142$ respondents or 22.1% vs $n=103$ or 17.8% in health-onliners), but the difference was not significant ($P=.06$). Health-onliners’ satisfaction with their GP was slightly lower than health-offliners’ (mean 4.02, SD 0.86 vs mean 4.13, SD 0.86, respectively; $P=.03$). Significant differences between health-onliners and health-offliners were found for several health-related variables, with perceived relevance of understanding somatic processes (mean 4.34, SD 0.83 vs mean 4.16, SD 0.90; $P<.001$) and health-consciousness (mean 3.59, SD 0.71 vs mean 3.34, SD 0.81; $P<.001$) higher among the health-onliners. Additionally, health-onliners were more likely to report being strongly interested in information concerning health topics ($n=297$ or 46.2% vs $n=164$ or 28.5%; $P<.001$) and generally used the Internet more often than health-offliners, with 65.3% versus 50.5% (corresponding to $n=420$ vs $n=291$ respondents) reporting using the Internet “daily” ($P<.001$).

Table 1. Sample characteristics of health-onliners and health-offliners.

Variable	Total sample n=1219	Health-onliners n=643	Health-offliners n=576	Difference onliners versus offliners <i>F</i> or chi-square statistics ^{a,b}	Degree of freedom	<i>P</i> value
Age						
Range	18-79	18-79	18-78			
Mean (SD) ^c	46.02 (14.96)	45.35 (14.55)	46.78 (15.38)	2.812 ^a	1	.09
Gender, n (%)						
Female	610 (50.04)	338 (52.6)	272 (47.2)	3.5 ^b	1	.06
Male	609 (49.96)	305 (47.4)	304 (52.8)			
Socioeconomic status, n (%)						
Low	177 (15.86)	79 (13.4)	98 (18.6)			
Medium	658 (58.96)	339 (57.5)	319 (60.6)	13.1 ^b	2	.001
High	281 (25.18)	172 (29.2)	109 (20.7)			
General Internet use, n (%)						
At least sometimes per month	189 (15.50)	68 (10.6)	121 (21)			
Several times per week	319 (26.17)	155 (24.1)	164 (28.5)	34.9 ^b	2	<.001
Daily	711 (58.33)	420 (65.3)	291 (50.5)			
Patient status, n (%)						
Chronically or severely ill	243 (20.05)	141 (22.1)	102 (17.8)	3.4 ^b		.06
Mildly or not affected	969 (79.95)	498 (77.9)	471 (82.2)		1	
Perceived relevance of understanding somatic processes ^d , mean (SD)	4.25 (0.86)	4.34 (0.83)	4.16 (0.90)	13.012 ^a	1	<.001
Interest in information concerning health topics, n (%)						
Low or medium	739 (62.15)	336 (53.8)	403 (71.5)	39.5 ^b	1	<.001
Strong	450 (37.85)	289 (46.2)	161 (28.5)			
Goals for HISB^{e,f}, mean (SD)						
Esteem support	0.24 (0.26)	0.24 (0.26)	-	N/A ^g		
Emotional support	0.06 (0.20)	0.06 (0.20)	-	N/A		
Informational support	0.37 (0.23)	0.37 (0.23)	-	N/A		
Health-consciousness ^h , mean (SD)	3.47 (0.77)	3.59 (0.71)	3.34 (0.81)	33.431 ^a	1	<.001
Satisfaction with general practitioner ⁱ , mean (SD)	4.07 (0.86)	4.02 (0.86)	4.13 (0.86)	4.492	1	.03

^a*F* values derived from analysis of variance (ANOVA) for continuous variables.^bChi-square values derived from chi-square test for shares.^cSD: standard deviation.^dScale ranges from 1 ("strongly disagree") to 5 ("strongly agree").^eHISB: health information-seeking behavior.^fScale ranges from 0 ("no" for all items of the scale) to 1 ("yes" for all items of the scale).^gN/A: not applicable.^hScale ranges from 1 ("Generally, I don't take care of my health") to 5 ("Generally, I take good care of my health").ⁱScale ranges from 1 ("very unsatisfied") to 5 ("very satisfied").

RQ1: Using the Internet to Search for Health Information

The results of the logistic regression models are depicted in Table 2 and the strength of the association between each predictor variable and the outcome is expressed in form of odds ratio (OR), which indicates the expected change in the odds to observe the outcome (ie, to be a health-onliner) when the respective predictor changes by one unit. There is no evidence for a main effect of gender (OR 1.21, 95% CI 0.90-1.61), but age and SES were significant predictors of being a health-onliner or health-offliner. However, a 1-year increase in age was associated with a decreased OR of being a health-onliner for women by the factor 0.97 (OR 0.97, 95% CI 0.96-0.99) and in the total sample (OR 0.99, 95% CI 0.98-0.997), but not for men. In contrast, a high SES was associated with significantly increased odds of going online for health-related information only for male respondents (OR 1.97, 95% CI 1.06-3.68) and in the combined model, that is, the whole sample of male and female respondents (OR 1.82, 95% CI 1.15-2.88), but not female respondents.

A higher frequency of general Internet use was associated with a nearly triple-increase in the odds of being a health-onliner (OR for “daily” use=2.91, 95% CI 1.92-4.41), with the slightly stronger effects for female (OR 3.23, 95% CI 1.86-5.59) than for male respondents (OR 2.50, 95% CI 1.30-4.78).

Persons who were chronically ill or severely affected by health problems were significantly more likely to be health-onliners, but only if they were women (OR 2.12, 95% CI 1.28-3.53). A similar relationship was found between perceived relevance of

understanding somatic processes and HISB, that is, for women, a one-point increase in the perceived importance of health literacy was associated with an OR 1.39 (95% CI 1.09-1.78) of being a health-onliner, whereas men had only a moderately heightened OR that did not reach significance. Although both male and female respondents appeared to be significantly influenced by having general interest in information on health topics, this impact was much stronger among female participants (OR_{women} 2.07, 95% CI 1.36-3.14; OR_{men} 1.70, 95% CI 1.09-2.63, respectively).

Degree of health-consciousness was associated with significantly increased OR for men (OR 1.46, 95% CI 1.10-1.94) and for the combined model (OR 1.33, 95% CI 1.10-1.61), but not for women alone. Higher satisfaction with one’s GP had a negative effect on the odds that men would seek health information online, that is, be health-onliners (OR 0.73, 95% CI 0.57-0.92).

The Hosmer-Lemeshow test inform on the proper classification of all cases included and gives a chi-square value of 11.5 (df=8; $P=.17$) for the total subsample. Both P values for the gender-specific models are also nonsignificant, what confirms no major differences between predicted and observed classification of cases [69]. Consistently, the overall model-fit is quite well, as indicated by the goodness of fit test comparing each full model with the empty model and yielding significant results in all three cases. The explained variance also indicated the existence of gender differences. Comparing men and women, the logit model better fits the data of female respondents: whereas Nagelkerke R^2 increased to 19.2% for women, it could only explain 13.7% of outcome variance (ie, whether the respondent was a health-onliner) for male respondents.

Table 2. Results of the logistic regression models predicting online health information-seeking behavior.

Determinants	Total (n=950) ^a		Male (n=463)		Female (n=487)	
	OR ^b (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
Age	0.99 (0.98-1.00)	.01	0.99 (0.98-1.01)	.46	0.97 (0.96-0.99)	.002
Gender (Ref: male)	1.21 (0.90-1.61)	.21	— ^c	—	—	—
Socioeconomic status (Ref: low)		.01		.02		.57
Medium	1.13 (0.77-1.66)	.54	1.07 (0.62-1.86)	.80	1.16 (0.67-2.01)	.59
High	1.82 (1.15-2.88)	.01	1.97 (1.06-3.68)	.03	1.46 (0.72-2.93)	.29
General Internet use (Ref: at least sometimes per month)		<.001		.005		<.001
Several times per week	1.57 (1.02-2.41)	.04	1.46 (0.72-2.94)	.29	1.70 (0.99-2.94)	.06
Daily	2.91 (1.92-4.41)	<.001	2.50 (1.30-4.78)	.006	3.23 (1.86-5.59)	<.001
Patient status: chronically or severely ill (Ref: mildly or not affected)	1.56 (1.11-2.19)	.01	1.22 (0.76-1.95)	.42	2.12 (1.28-3.53)	.004
Perceived relevance of understanding somatic processes ^d	1.27 (1.08-1.50)	.005	1.22 (0.97-1.53)	.10	1.39 (1.09-1.78)	.008
Strongly interested in information concerning health topics (Ref: weakly or not interested)	1.89 (1.40-2.54)	<.001	1.70 (1.09-2.63)	.02	2.07 (1.36-3.14)	.001
Health-consciousness ^e	1.33 (1.10-1.61)	.004	1.46 (1.10-1.94)	.008	1.24 (0.95-1.62)	.11
Satisfaction with general practitioner ^f	0.82 (0.70-0.96)	.02	0.73 (0.57-0.92)	.008	0.91 (0.72-1.14)	.40
Constant	0.31	.000	0.35	.006	0.35	.002
Hosmer-Lemeshow test (chi-square, df; P)	11.5, 8; .17		12.2, 8; .14		4.7, 8; .79	
Goodness of fit ^g (chi-square, df; P)	116.3, 11; <.001		50.2, 10; <.001		75.5, 10; <.001	
Nagelkerke R ²	.154		.137		.192	

^aThe difference between the number of total cases included in the descriptive section and in the logit models is due to the listwise exclusion of missing cases.

^bOR: odds ratio.

^cThe dash indicates the absence of the variable “gender” in both gender-specific models.

^d1 (“strongly disagree”) to 5 (“strongly agree”).

^e1 (Generally, I don’t take care of my health”) to 5 (“Generally, I take good care of my health”).

^f1 (“very unsatisfied”) to 5 (“very satisfied”).

^g(–2 Log L compared with –2 Log L of the empty model).

RQ2: Frequency of Online Health Information-Seeking Behavior (OHISB)

The results of the zero-truncated negative binomial regression models are shown as incidence rate ratio (IRR) in Table 3. The name of the measure has changed to IRR, because the outcome now reflects the number of incidences (Internet access events with the purpose to conduct OHISB) observed in the last year, but the interpretation remains analog to OR, as pointed out above. Focusing only on health-onliners, SES turned out not to be a relevant predictor for higher frequencies of information seeking. Analogous to the results for access to online health information (see Table 2), increasing age was also significantly associated with women’s OHISB frequency (IRR 0.99, 95% CI

0.975-0.996). Since the effect on this outcome is multiplicative, a 1-year increase thus leads to a predicted OHISB frequency, which is decreased by a factor of 0.99. There was no influence of gender on the frequency of OHISB for the total sample (IRR 1.00, 95% CI 0.99-1.01), but differences were again found in the patterns of influences on women versus men.

Although the effects of higher frequencies of general Internet use are similar in size and P values in the combined model, the gender-specific models revealed differences between men and women: whereas daily use of general Internet was only associated with a significant increase in OHISB frequency in males (IRR 2.49, 95% CI 1.43-4.35), using the Internet “several times per week” was significant only for female respondents (IRR 1.54, 95% CI 1.01-2.35).

Being a patient with a chronic or severe disease was a positive predictor of OHISB frequency (IRR 1.57, 95% CI 1.26-1.95) regardless of gender, as the estimated IRRs do not differ substantially between men and women. Respondents who reported being strongly interested in information concerning health topics were much more likely to seek out information online more frequently (IRR 1.45, 95% CI 1.19-1.77), and gender played a much weaker role than health status.

Self-reported health-consciousness in the zero-truncated negative binomial regression models was—as compared with the findings from the logit models predicting the utilization of the Internet for health information purposes—not associated with significant effects. In contrast, perceived relevance of understanding somatic processes had the opposite effect on OHISB frequency of that predicted by the logit models: In the zero-truncated negative binomial regression models, belief in health literacy

became a significant negative predictor, but for men only (IRR 0.81, 95% CI 0.68-0.96). Satisfaction with one's GP changed from being a significant factor only for males to being a significant factor only for females, the latter now with a strong negative effect (IRR 0.75, 95% CI 0.65-0.88), whereas men's frequency of OHISB seems to be statistically unrelated to their degree of satisfaction.

Some of the three sum indices representing different goals of OHISB showed strong explanatory potential: whereas *esteem support* seems to be an important motivational factor only for women (IRR 2.22, 95% CI 1.30-3.79), *informational support* was associated with a quadrupled OHISB frequency per point for women (IRR 4.03, 95% CI 2.17-7.49), and with a slightly weaker effect for men (IRR 2.56, 95% CI 1.34-4.90). The goal of *emotional support* had no influence on respondents' OHISB frequency.

Table 3. Results of the zero-truncated negative binomial regression models on the frequency of online health information-seeking behavior (OHISB).

Determinants	Total (n=510) ^a		Male (n=241)		Female (n=269)	
	IRR ^b (95% CI)	P	IRR (95% CI)	P	IRR (95% CI)	P
Age	0.99 (0.982-0.998)	.01	1.00 (0.99-1.01)	.58	0.99 (0.975-0.996)	.009
Gender (Ref: male)	0.99 (0.82-1.22)	.99	— ^c	—	—	—
Socioeconomic status (Ref: low)						
Medium	1.06 (0.79-1.41)	.71	1.01 (0.66-1.55)	.97	0.93 (0.63-1.36)	.70
High	1.25 (0.91-1.73)	.17	1.14 (0.72-1.79)	.57	1.14 (0.72-1.81)	.56
General Internet use (Ref: at least sometimes per month)						
Several times per week	1.60 (1.12-2.27)	.009	1.72 (0.94-3.16)	.08	1.54 (1.01-2.35)	.04
Daily	1.67 (1.19-2.33)	.003	2.49 (1.43-4.35)	.001	1.28 (0.84-1.96)	.25
Patient status: chronically or severely ill (Ref: mildly or not affected)	1.57 (1.26-1.95)	<.001	1.67 (1.22-2.29)	.001	1.43 (1.07-1.91)	.02
Perceived relevance of understanding somatic processes ^d	0.90 (0.80-1.01)	.06	0.81 (0.68-0.96)	.02	0.96 (0.82-1.13)	.65
Strongly interested in information concerning health topics (Ref: weekly or not interested)	1.45 (1.19-1.77)	<.001	1.46 (1.09-1.97)	.01	1.42 (1.10-1.83)	.01
Goals of OHISB^e						
Esteem support	1.91 (1.28-2.83)	.001	1.49 (0.82-2.72)	.19	2.22 (1.30-3.79)	.004
Emotional support	0.90 (0.57-1.43)	.65	1.13 (0.53-2.40)	.75	0.83 (0.46-1.49)	.52
Informational support	3.12 (1.97-4.96)	<.001	2.56 (1.34-4.90)	.004	4.03 (2.17-7.49)	<.001
Health-consciousness ^f	1.10 (0.97-1.26)	.14	1.70 (0.89-1.29)	.49	1.08 (0.91-1.30)	.38
Satisfaction with general practitioner ^g	0.86 (0.77-0.96)	.007	1.09 (0.93-1.28)	.29	0.75 (0.65-0.88)	<.001
Constant	0.47	.04	0.32	.36	0.65	.01

^aThe difference between the total number of cases included in the descriptive section and in the models depicted in this table is due to the listwise exclusion of missing cases.

^bIRR: incidence rate ratio.

^cThe — indicates the absence of the variable “gender” in both gender-specific models

^dFrom 1 (“strongly disagree”) to 5 (“strongly agree”).

^eOHISB: online health information-seeking behavior.

^fFrom 1 (“Generally, I don’t take care of my health”) to 5 (“Generally, I take good care of my health”).

^gFrom 1 (“very unsatisfied”) to 5 (“very satisfied”).

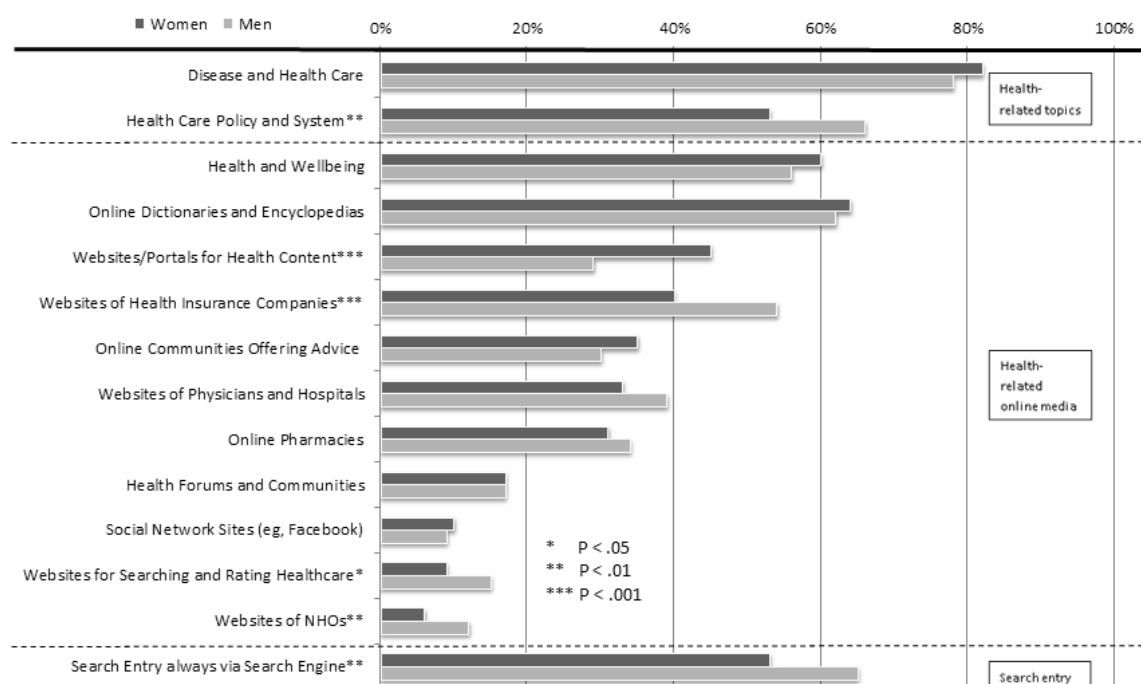
RQ3: Topics and Sources of Online Health Information-Seeking Behavior (OHISB)

Figure 1 illustrates which health topics women and men are interested in and which online sources they use when they search for health information. Diseases and health care are of particular interest for about 79.7% of the health-onliners ($n=510$ of 640 respondents), health and wellbeing are also relevant topics for both women and men. The Internet seems to serve as a source of specific information that is sought primarily using search engines and online dictionaries.

Online media offering the opportunity to share information and to interact with others, specifically, online health communities and social networking sites are not yet established as a means

of OHISB in the broad public, with an overall usage of 17.0% and 9.7% (corresponding to 109/640 and 62/640 respondents), respectively. In some cases, we can detect significant differences in issue- and channel-related preferences between women and men, that is, in general, men focus more on topics concerning health care policy and systems (66.4%, 202/304 males vs 53.0%, 178/336 females; $P=.001$), and visit the websites of health insurance companies (53.6%, 163/304 males vs 39.6%, 133/336 females; $P<.001$) and noncommercial health organizations (12.2%, 37/304 men vs 6.0%, 20/336 women; $P=.006$) more frequently than women do. Women reported significantly more usage of websites or portals for health content (44.6%, 150/336 female respondents vs 28.9%, 88/304 males; $P<.001$).

Figure 1. Gender differences in online health information-seeking behavior (OHISB) concerning topics and sources of online communication. NHOs: Noncommercial health organizations (total $n=640$).



Discussion

Principal Findings

Despite the fact that men and women reported equal access to online health information, our data indicate that OHISB should be explained using gender-specific models, to account for several significant gender differences among health-onliners. Dissatisfaction with primary care seems to more often trigger women to seek patient esteem support through online health information seeking; OHISB might therefore serve a compensatory function. These and additional

results—particularly regarding gender differences, implications, and methodical limitations—are discussed and compared with international data.

Our results indicate that SES and age remain relevant barriers to general access to health information on the Internet, but only for specific genders. We found increasing age to be significantly associated with both access to and frequency of OHISB for women only, thus enhancing understanding of the gender-specificity of the well-established negative correlation between age and OHISB [7]. In contrast, we found high SES to be associated with increased odds of conducting OHISB

among male respondents only; among females, SES had no significant effects on neither the use nor frequency of searching. This finding is partly consistent with former studies: whereas some researchers have reported no significant effect of higher educational levels on the frequency of OHISB [15], others found strong associations between both respondents' income and educational level on the likelihood of using the Internet to find out more about a medical condition [9].

Higher frequencies of general Internet use revealed to be consistently associated with more frequent OHISB [49]. We found a stronger effect of daily Internet use on the frequency of OHISB for men, whereas women's OHISB seems to be only slightly influenced by their general Internet use. This specific association lacks direct comparability to former results, although Renahy and colleagues [15] found a positive effect of more frequent Internet use on the frequency of OHISB, with no gender differences.

The effect of being severely or chronically ill affected OHISB differently for different groups. Only severely ill women, not men, were significantly more likely to be health-onliners, consistent with findings from a French study [15]. In contrast, the impact of patient status on the frequency of OHISB was slightly stronger for male respondents. We tentatively interpret these differences as supporting the 2-step data analysis strategy we chose. These results offer new insights into the relationship between patient status and OHISB when compared with existing nongender-differentiated findings [48,49].

The association between OHISB and related online activities and interests (eg, buying drugs and other health-related products online) that indicate a high level of interest in health information is neither surprising nor new, as this has been reported in both an analysis of cross-sectional data from 7 European countries [3] as well as in US data [49]. However, little evidence has been produced to date on the influence of perceived relevance of understanding somatic processes on OHISB. To our knowledge, only one study has investigated a similar variable: Bidmon and Terlutter [24] found that women reported a slightly higher personal disposition of being well-informed as a patient than men, but the difference was not significant and the association with details of OHISB was not explicitly assessed. Even less comparable evidence exists concerning health-consciousness as a predictor of OHISB. This means that our findings indicating that health-consciousness has a significant influence on male utilization of OHISB only may act as a benchmark for future studies.

We found that whereas women are inclined to engage in more frequent OHISB in light of their goals reflecting needs for esteem support and informational support, men tend to be driven more by purely informational motives. This is consistent with another recent finding that women were more likely than men to conduct OHISB for social motives and enjoyment [24]. Comparing the results on our two outcomes of access to and frequency of online health information seeking, the introduction of these indices may have absorbed some explanatory power from health-consciousness.

Our results further indicate that using the Internet can serve a compensatory function, but in different ways for women and

men. Whereas a lower satisfaction with one's GP motivates men to turn to the Internet for health-related purposes (raw usage, independent of the frequency), a lower satisfaction with one's GP is associated with an increased frequency of OHISB reported by women. These findings are in line with another study reporting that dissatisfied cancer patients seek health information from sources other than their physicians [70], and with a study that found that women engaged in more frequent OHISB when they suspected that their GP was not telling them everything about their health, or when they reported a general preference to wait before going to see a physician [15].

Our findings regarding sources employed for OHISB are partly consistent with a similar study: Females from our sample reported using health content-related websites significantly more often than males, which may reflect the stronger social supportive patterns detected among women [24]. The majority of respondents (n=374 respondents or 59.0%) reported always using a search engine when conducting OHISB; this is consistent with a finding from the United States, in which an even higher proportion of respondents (77%) reported following this strategy [9]. These large proportions indicate that the primary purpose of OHISB is to receive quick and easy access to online health information. Moreover, OHISB reflects a need—especially among men—for health information that is clearly explained and tailored to their specific needs.

Limitations

The first limitation is that the cross-sectional data used in our analysis do not allow for any causal attributions, even in cases that seem straightforward, such as the effects of age or health status on OHISB.

The second limitation is that outcome operationalization was somewhat explorative, as a well-established, validated scale for assessing access to and frequency of OHISB does not yet exist. Development of a validated measure to assess OHISB is the central precondition of conducting internationally comparable research on this behavior. Such a measure would also complement the valid and reliable measure for assessing eHealth literacy (ie, the ability to seek, find, understand, and appraise health information from electronic sources and apply it to addressing or solving a health problem) that has already been developed [71]. As this inquiry used data from a large-scale representative survey conducted regularly, our results are affected by the typical constraints of a secondary analysis of data that were not primarily collected for the analyzed purposes—the lack of a measure for eHealth literacy [71] is one obvious drawback. Other important characteristics—particularly health status—were measured using one-item self-reports, which offers only a superficial assessment. This criticism can also be applied to other measures of health-related online activities and online usage behavior, as self-reports of online activities often diverge from real behavior [72]. The validity of these findings should be enhanced in future studies by complementarily using objective measures of health-related online activities.

Further limitations include, third, that no differentiation is made between people who are searching for information for themselves and those who are searching for others ("surrogate seekers"). Finally, we only used a binary categorization of men

and women, which does not cover all facets of such a complex construct [73]. Our results might have also been influenced by individuals' gender-role orientation [74]. It might be the case, that women and men only searching for health information both scored rather high on femininity and therefore were more similar than people not seeking for health information. In addition, gender-roles orientations arising from differing social and cultural environments might differentially influence OHISB. Further research should therefore include measures of gender orientation such as the Bem Sex-Role Inventory [75] and samples more diverse in cultural background.

Conclusions

Our results provide promising and innovative insights into OHISB and indicate that a deeper understanding of OHISB requires differentiating between access to online health information (ie, differentiating between health-onliners and health-offliners) and the frequency of OHISB. This deeper understanding would be particularly valuable for the analysis of what are often subtle gender-based differences. Furthermore, sociodemographic, health-related, and motivational determinants of OHISB should be taken into account when explaining such complex behavior. This recommendation also applies to the associations between skills-related (ie, eHealth literacy) and

behavior-related (OHISB) concepts, whose interrelations have yet to be analyzed sufficiently [76].

Overall, although users indicate that the Internet is capable of providing a valuable source of informational support and esteem support, gender-specific, user-oriented sources and empowerment-strategies should be developed to increase the benefits of OHISB. This may include enlisting the support of health care providers to supply patients with health information sources that offer evidence-based, transparent, and credible information. To close the gap in OHISB due to age and SES, such resources might, for example, reduce the complexity of the language and enhance the understandability of the health information offered. Gender-specific determinants and patterns in information-seeking behavior should also be taken into account in theories of health information seeking and in the provision of online health information by offering information in accordance with male and female preferences regarding goals, sources, and topics. For example, men's technical affinity might be used as a pathway in health communication to raise their interest in health content about diseases and well-being [77], whereas women's need for emotional support might be met with communication in online communities via social network sites [20].

Acknowledgments

The authors thank the Bertelsmann Foundation for their cooperation and for providing the final dataset.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Items.

[PDF File (Adobe PDF File), 203KB - [jmir_v19i4e92_app1.pdf](#)]

References

1. 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](#)]
2. Nijman J, Hendriks M, Brabers A, de JJ, Rademakers J. Patient activation and health literacy as predictors of health information use in a general sample of Dutch health care consumers. *J Health Commun* 2014;19(8):955-969. [doi: [10.1080/10810730.2013.837561](#)] [Medline: [24397280](#)]
3. Santana S, Lausen B, Bujnowska-Fedak M, Chronaki CE, Prokosch H, Wynn R. Informed citizen and empowered citizen in health: results from an European survey. *BMC Fam Pract* 2011;12:20 [FREE Full text] [doi: [10.1186/1471-2296-12-20](#)] [Medline: [21496309](#)]
4. Jordan JE, Buchbinder R, Osborne RH. Conceptualising health literacy from the patient perspective. *Patient Educ Couns* 2010 Apr;79(1):36-42. [doi: [10.1016/j.pec.2009.10.001](#)] [Medline: [19896320](#)]
5. Kickbusch I. Health literacy: a search for new categories. *Health Promot Int* 2002;17(1):1-2. [doi: [10.1093/heapro/17.1.1](#)]
6. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int* 2000;15(3):259-267. [doi: [10.1093/heapro/15.3.259](#)]
7. Hallyburton A, Evarts LA. Gender and online health information seeking: a five survey meta-analysis. *J Consum Health Internet* 2014 Jun 04;18(2):128-142. [doi: [10.1080/15398285.2014.902268](#)]
8. Johnson J, Case D. Health information seeking. New York: Peter Lang; 2013.
9. Fox S, Duggan M. Pew Research Center. 2013. Health Online 2013 URL: http://pewinternet.org/~media/Files/Reports/PIP_HealthOnline.pdf [accessed 2016-05-13] [WebCite Cache ID 6hTGrIKHI]
10. Higgins O, Sixsmith J, Barry M, Domegan C. European Centre for Disease Prevention and Control. 2011. A literature review on health information-seeking behaviour on the web: a health consumer and health professional perspective URL:

- <http://www.ecdc.europa.eu/en/publications/Publications/Literature%20review%20on%20health%20information-seeking%20behaviour%20on%20the%20web.pdf> [accessed 2016-09-20] [WebCite Cache ID 6ketNGKCJ]
11. Abrahamson JA, Fisher KE, Turner AG, Durrance JC, Turner TC. Lay information mediary behavior uncovered: exploring how nonprofessionals seek health information for themselves and others online. *J Med Libr Assoc* 2008 Oct;96(4):310-323 [FREE Full text] [doi: [10.3163/1536-5050.96.4.006](https://doi.org/10.3163/1536-5050.96.4.006)] [Medline: [18974809](https://pubmed.ncbi.nlm.nih.gov/18974809/)]
 12. Cutrona SL, Mazor KM, Vieux SN, Luger TM, Volkman JE, Finney RL. Health information-seeking on behalf of others: characteristics of 'surrogate seekers'. *J Cancer Educ* 2015 Mar;30(1):12-19 [FREE Full text] [doi: [10.1007/s13187-014-0701-3](https://doi.org/10.1007/s13187-014-0701-3)] [Medline: [24989816](https://pubmed.ncbi.nlm.nih.gov/24989816/)]
 13. Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. *Am J Prev Med* 2007 May;32(5 Suppl):S131-S133 [FREE Full text] [doi: [10.1016/j.amepre.2007.02.012](https://doi.org/10.1016/j.amepre.2007.02.012)] [Medline: [17466818](https://pubmed.ncbi.nlm.nih.gov/17466818/)]
 14. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. *J Med Internet Res* 2012 Jan;14(1):e19 [FREE Full text] [doi: [10.2196/jmir.1619](https://doi.org/10.2196/jmir.1619)] [Medline: [22357448](https://pubmed.ncbi.nlm.nih.gov/22357448/)]
 15. Renahy E, Parizot I, Chauvin P. Determinants of the frequency of online health information seeking: results of a web-based survey conducted in France in 2007. *Inform Health Soc Care* 2010 Jan;35(1):25-39 [FREE Full text] [doi: [10.3109/17538150903358784](https://doi.org/10.3109/17538150903358784)] [Medline: [20302437](https://pubmed.ncbi.nlm.nih.gov/20302437/)]
 16. Feng Y, Xie W. Digital divide 2.0: the role of social networking sites in seeking health information online from a longitudinal perspective. *J Health Commun* 2015;20(1):60-68. [doi: [10.1080/10810730.2014.906522](https://doi.org/10.1080/10810730.2014.906522)] [Medline: [25119019](https://pubmed.ncbi.nlm.nih.gov/25119019/)]
 17. Kimbrough AM, Guadagno RE, Muscanell NL, Dill J. Gender differences in mediated communication: Women connect more than do men. *Comput Human Behav* 2013 May;29(3):896-900. [doi: [10.1016/j.chb.2012.12.005](https://doi.org/10.1016/j.chb.2012.12.005)]
 18. Rowley J, Johnson F, Sbaffi L. Gender as an influencer of online health information-seeking and evaluation behavior. *J Assn Inf Sci Tec* 2015 Aug 25:n/a-n/a. [doi: [10.1002/asi.23597](https://doi.org/10.1002/asi.23597)]
 19. Wellstead P. Inform Res. 2011. Information behaviour of Australian men experiencing stressful life events: the role of social networks and confidants URL: http://s3.amazonaws.com/academia.edu.documents/30618903/Information_behaviour_of_Australian_men_experiencing_stressful_life_events_.pdf [accessed 2015-09-08] [WebCite Cache ID 6kN4Z6xmi]
 20. Mo PK, Malik SH, Coulson NS. Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups. *Patient Educ Couns* 2009 Apr;75(1):16-24. [doi: [10.1016/j.pec.2008.08.029](https://doi.org/10.1016/j.pec.2008.08.029)] [Medline: [19233586](https://pubmed.ncbi.nlm.nih.gov/19233586/)]
 21. Wajcman J. Gender and the cultures of technology, work and management. In: Godfroy-Genin AS, editor. *Women in engineering and technology research*. Berlin: LIT Verlag; 2010:29-39.
 22. Zhang X, Guo X, Lai K, Guo F, Li C. Understanding gender differences in m-health adoption: a modified theory of reasoned action model. *Telemed J E Health* 2014 Jan;20(1):39-46. [doi: [10.1089/tmj.2013.0092](https://doi.org/10.1089/tmj.2013.0092)] [Medline: [24161004](https://pubmed.ncbi.nlm.nih.gov/24161004/)]
 23. Grellhesl M, Punyanunt-Carter NM. Using the uses and gratifications theory to understand gratifications sought through text messaging practices of male and female undergraduate students. *Comput Human Behav* 2012 Nov;28(6):2175-2181. [doi: [10.1016/j.chb.2012.06.024](https://doi.org/10.1016/j.chb.2012.06.024)]
 24. 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](https://doi.org/10.2196/jmir.4127)] [Medline: [26099325](https://pubmed.ncbi.nlm.nih.gov/26099325/)]
 25. 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](https://doi.org/10.1016/0749-5978(91)90020-T)]
 26. Afifi WA. Toward a theory of motivated information management. *Communication Theory* 2004 May 01;14(2):167-190. [doi: [10.1111/j.1468-2885.2004.tb00310.x](https://doi.org/10.1111/j.1468-2885.2004.tb00310.x)]
 27. Rimal RN. Perceived risk and efficacy beliefs as motivators of change: use of the risk perception attitude (rpa) framework to understand health behaviors. *Hum Commun Res* 2003 Jul 01;29(3):370-399. [doi: [10.1093/hcr/29.3.370](https://doi.org/10.1093/hcr/29.3.370)]
 28. Griffin RJ, Dunwoody S, Neuwirth K. Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. *Environ Res* 1999 Feb;80(2 Pt 2):S230-S245. [doi: [10.1006/enrs.1998.3940](https://doi.org/10.1006/enrs.1998.3940)] [Medline: [10092438](https://pubmed.ncbi.nlm.nih.gov/10092438/)]
 29. 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](https://doi.org/10.1108/00220411211225575)]
 30. Yang ZJ, Aloe AM, Feeley TH. Risk information seeking and processing model: a meta-analysis. *J Commun* 2014 Jan 07;64(1):20-41. [doi: [10.1111/jcom.12071](https://doi.org/10.1111/jcom.12071)]
 31. Manierre MJ. Gaps in knowledge: tracking and explaining gender differences in health information seeking. *Soc Sci Med* 2015 Mar;128:151-158. [doi: [10.1016/j.socscimed.2015.01.028](https://doi.org/10.1016/j.socscimed.2015.01.028)] [Medline: [25618604](https://pubmed.ncbi.nlm.nih.gov/25618604/)]
 32. Montano D, Kasprzyk D. Theory of reasoned action, theory of planned behavior, the integrated behavioral model. In: Glanz K, Rimer BK, Viswanath K, editors. *Health behavior and health education: Theory, research, and practice*. 4th edition. San Francisco, CA: Jossey-Bass; 2008:67-95.
 33. Eagly A. Sex differences in social behavior: A social-role interpretation. Hillsdale, NJ: Erlbaum; 1987.
 34. Simou E. Health information sources: trust and satisfaction. *IJH* 2015 Oct 25;2(1):38-43. [doi: [10.5430/ijh.v2n1p38](https://doi.org/10.5430/ijh.v2n1p38)]

35. Meyers-Levy J, Loken B. Revisiting gender differences: what we know and what lies ahead. *J Consum Psychol* 2015 Jan;25(1):129-149. [doi: [10.1016/j.jcps.2014.06.003](https://doi.org/10.1016/j.jcps.2014.06.003)]
36. World Bank. 2016. Internet users (per 100 people) URL: http://data.worldbank.org/indicator/IT.NET.USER.P2?cid=GPD_44 [accessed 2016-05-25] [WebCite Cache ID 6h1THHgc1]
37. Reisdorf BC, Groselj D. Internet (non-)use types and motivational access: Implications for digital inequalities research. *New Media Soc* 2015 Dec 27;1-46. [doi: [10.1177/1461444815621539](https://doi.org/10.1177/1461444815621539)]
38. 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](https://doi.org/10.1177/1461444813487959)]
39. Zillien N, Hargittai E. Digital distinction: status-specific types of Internet usage. *Soc Sci Quart* 2009;90(2):274-291. [doi: [10.1111/j.1540-6237.2009.00617.x](https://doi.org/10.1111/j.1540-6237.2009.00617.x)]
40. Pew Research Center. 2014. Internet user demographics URL: <http://www.pewinternet.org/data-trend/internet-use/latest-stats/> [accessed 2016-05-13] [WebCite Cache ID 6hTHm6aqq]
41. Välimäki M, Nenonen H, Koivunen M, Suhonen R. Patients' perceptions of Internet usage and their opportunity to obtain health information. *Med Inform Internet Med* 2007 Dec;32(4):305-314. [doi: [10.1080/14639230701819792](https://doi.org/10.1080/14639230701819792)] [Medline: [18072007](https://pubmed.ncbi.nlm.nih.gov/18072007/)]
42. International Telecommunication Union (ITU). 2015. Gender ICT statistics: individuals using the internet 2010-2014 URL: https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2015/Gender_2010-2014.xls [accessed 2016-05-13] [WebCite Cache ID 6hTHKdYyZ]
43. Frees V, Koch W. Internetnutzung: Frequenz und Vielfalt nehmen in allen Altersgruppen zu [internet use: frequency and diversity increase in all age groups.]. *Media Perspektiven* 2015;10:366-377.
44. Office for National Statistics. 2016. Statistical bulletin: Internet users: 2015 URL: <http://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2015#age> [accessed 2016-12-07] [WebCite Cache ID 6mZiQAvZs]
45. Muscanell NL, Guadagno RE. Make new friends or keep the old: Gender and personality differences in social networking use. *Comput Hum Behav* 2012 Jan;28(1):107-112. [doi: [10.1016/j.chb.2011.08.016](https://doi.org/10.1016/j.chb.2011.08.016)]
46. European Commission. 2014. Europeans becoming enthusiastic users of online health information URL: <https://ec.europa.eu/digital-agenda/en/news/europeans-becoming-enthusiastic-users-online-health-information> [accessed 2016-02-17] [WebCite Cache ID 6hIUBeFCp]
47. Gitlow S. The online community as a healthcare resource. In: Nash DB, Manfredi MP, Bozarth B, Howell S, editors. *Connecting with the new healthcare consumer: Defining your strategy*. New York: McGraw-Hill; 2000:113-133.
48. Lorence DP, Park H, Fox S. Assessing health consumerism on the Web: a demographic profile of information-seeking behaviors. *J Med Syst* 2006 Aug;30(4):251-258. [Medline: [16978004](https://pubmed.ncbi.nlm.nih.gov/16978004/)]
49. 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](https://doi.org/10.1016/j.ijmedinf.2005.07.032)] [Medline: [16125453](https://pubmed.ncbi.nlm.nih.gov/16125453/)]
50. 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](https://doi.org/10.1016/j.pec.2005.10.006)] [Medline: [16406474](https://pubmed.ncbi.nlm.nih.gov/16406474/)]
51. Oh HJ, Lee B. The effect of computer-mediated social support in online communities on patient empowerment and doctor-patient communication. *Health Commun* 2012;27(1):30-41. [doi: [10.1080/10410236.2011.567449](https://doi.org/10.1080/10410236.2011.567449)] [Medline: [21797714](https://pubmed.ncbi.nlm.nih.gov/21797714/)]
52. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. *Health Educ Behav* 2015 Apr;42(2):202-209 [FREE Full text] [doi: [10.1177/1090198114547815](https://doi.org/10.1177/1090198114547815)] [Medline: [25156311](https://pubmed.ncbi.nlm.nih.gov/25156311/)]
53. 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](https://doi.org/10.2196/jmir.3117)] [Medline: [25048379](https://pubmed.ncbi.nlm.nih.gov/25048379/)]
54. Courtenay W. Key determinants of the health and well-being of men and boys. *Int J Mens Health* 2003 Jan 1;2(1):1-30. [doi: [10.3149/jmh.0201.1](https://doi.org/10.3149/jmh.0201.1)]
55. Hatchell AC, Bassett-Gunter RL, Clarke M, Kimura S, Latimer-Cheung AE. Messages for men: the efficacy of EPPM-based messages targeting men's physical activity. *Health Psychol* 2013;32(1):24-32. [doi: [10.1037/a0030108](https://doi.org/10.1037/a0030108)]
56. Ek S. Gender differences in health information behaviour: a Finnish population-based survey. *Health Promot Int* 2015 Sep;30(3):736-745 [FREE Full text] [doi: [10.1093/heapro/dat063](https://doi.org/10.1093/heapro/dat063)] [Medline: [23985248](https://pubmed.ncbi.nlm.nih.gov/23985248/)]
57. Plotnikoff RC, Lippke S, Johnson ST, Hugo K, Rodgers W, Spence JC. Awareness of Canada's Physical Activity Guide to Healthy Active Living in a Large Community Sample. *Am J Health Promot* 2011 May;25(5):294-297. [doi: [10.4278/ajhp.090211-ARB-60](https://doi.org/10.4278/ajhp.090211-ARB-60)]
58. AOK Bundesverband. 2014. Prävention bei Männern muss digital sein [Prevention has to be digital for men] URL: http://www.aok-bv.de/presse/pressemitteilungen/2013/index_09735.html [accessed 2014-02-11] [WebCite Cache ID 6kOKWoyQa]
59. Gesundheitsmonitor. Gesundheitsmonitor Bertelsmann Healthcare Monitor URL: <http://www.gesundheitsmonitor.de/> [accessed 2017-03-11] [WebCite Cache ID 6osnvLR8Y]

60. Rückert I, Böcken J, Mielck A. Are German patients burdened by the practice charge for physician visits ('Praxisgebuehr')? a cross sectional analysis of socio-economic and health related factors. *BMC Health Serv Res* 2008 Nov;8:232 [FREE Full text] [doi: [10.1186/1472-6963-8-232](https://doi.org/10.1186/1472-6963-8-232)] [Medline: [19014476](https://pubmed.ncbi.nlm.nih.gov/19014476/)]
61. Böcken J, Braun B, Meierjürgen R. Gesundheitsmonitor 2015 [Healthcare Monitor 2015] Bürgerorientierung im Gesundheitswesen - Kooperationsprojekt der Bertelsmann Stiftung und der BARMER GEK. Gütersloh: Bertelsmann Stiftung; 2015.
62. GfK Healthcare. 2014. Gesundheitsmonitor- feld- und methodenbericht - welle 22 [healthcare monitor - methods - wave 22] URL: http://gesundheitsmonitor.de/uploads/tx_itao_download/Bericht_Feld_Method_Frabo_Welle_22_final.pdf [accessed 2016-11-18] [WebCite Cache ID 6mZj03gsx]
63. Holt-Lunstead J, Uchino B. Social support and health. In: Glanz K, Rimer BK, Viswanath K, editors. *Health behavior and health education: Theory, research, and practice*. 5th edition. San Francisco: Jossey-Bass; 2015:183-204.
64. Gardner W, Mulvey EP, Shaw EC. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychol Bull* 1995 Nov;118(3):392-404. [Medline: [7501743](https://pubmed.ncbi.nlm.nih.gov/7501743/)]
65. Kim H, Kriebel D. Regression models for public health surveillance data: a simulation study. *Occup Environ Med* 2009 Nov;66(11):733-739. [doi: [10.1136/oem.2008.042887](https://doi.org/10.1136/oem.2008.042887)] [Medline: [19687020](https://pubmed.ncbi.nlm.nih.gov/19687020/)]
66. Afifi AA, Kotlerman JB, Ettner SL, Cowan M. Methods for improving regression analysis for skewed continuous or counted responses. *Annu Rev Public Health* 2007;28:95-111. [doi: [10.1146/annurev.publhealth.28.082206.094100](https://doi.org/10.1146/annurev.publhealth.28.082206.094100)] [Medline: [17112339](https://pubmed.ncbi.nlm.nih.gov/17112339/)]
67. Hayat MJ, Higgins M. Understanding poisson regression. *J Nurs Educ* 2014 Apr;53(4):207-215. [doi: [10.3928/01484834-20140325-04](https://doi.org/10.3928/01484834-20140325-04)] [Medline: [24654593](https://pubmed.ncbi.nlm.nih.gov/24654593/)]
68. Hilbe JM. *Negative binomial regression*. 2nd edition. New York: Cambridge University Press; 2011.
69. Hosmer D, Lemeshow S, Sturdivant R. *Applied logistic regression*. 3rd edition. Hoboken, NJ: Wiley; 2011.
70. Tustin N. The role of patient satisfaction in online health information seeking. *J Health Commun* 2010 Jan;15(1):3-17. [doi: [10.1080/10810730903465491](https://doi.org/10.1080/10810730903465491)] [Medline: [20390974](https://pubmed.ncbi.nlm.nih.gov/20390974/)]
71. 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](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
72. Scharkow M. The accuracy of self-reported internet use-a validation study using client log data. *Commun Methods Meas* 2016 Mar 24;10(1):13-27. [doi: [10.1080/19312458.2015.1118446](https://doi.org/10.1080/19312458.2015.1118446)]
73. Archer J, Lloyd B. *Sex and gender*. Cambridge: University Press; 2002.
74. Clément-Guillotin C, Chalabaev A, Fontayne P. Is sport still a masculine domain. A psychological glance. *Inter J Sport Psychol* 2012;42:67-78.
75. Bem S. *Bem sex-role inventory*. Palo Alto, CA: Consulting Psychologists Press; 1981.
76. Diviani N, van den Putte B, Giani S, van Weert JC. Low health literacy and evaluation of online health information: a systematic review of the literature. *J Med Internet Res* 2015;17(5):e112 [FREE Full text] [doi: [10.2196/jmir.4018](https://doi.org/10.2196/jmir.4018)] [Medline: [25953147](https://pubmed.ncbi.nlm.nih.gov/25953147/)]
77. Klenk S, Reifegerste D, Renatus R. Zielgruppenspezifische Gesundheitskommunikation fuer Maenner? Apps als Praxisoption [targeted health communication for men - apps an option for practise]. In: Camerini AL, Ludolph R, Rothenfluh FB, editors. *Gesundheitskommunikation im Spannungsfeld zwischen Theorie und Praxis*. Baden-Baden: Nomos; 2016:286-297.

Abbreviations

GP: general practitioner
HINTS: Health Information National Trends Survey
HISB: health information-seeking behavior
IRR: incidence rate ratio
OHISB: online health information-seeking behavior
OR: odds ratio
SES: socioeconomic status
TPB: theory of planned behavior

Edited by CL Parra-Calderón; submitted 20.09.16; peer-reviewed by E Neter, S Bidmon, R Terlutter; comments to author 08.11.16; revised version received 07.12.16; accepted 14.02.17; published 04.04.17.

Please cite as:

Baumann E, Czerwinski F, Reifegerste D

Gender-Specific Determinants and Patterns of Online Health Information Seeking: Results From a Representative German Health Survey

J Med Internet Res 2017;19(4):e92

URL: <http://www.jmir.org/2017/4/e92/>

doi: [10.2196/jmir.6668](https://doi.org/10.2196/jmir.6668)

PMID: [28377367](https://pubmed.ncbi.nlm.nih.gov/28377367/)

©Eva Baumann, Fabian Czerwinski, Doreen Reifegerste. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 04.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey

Clemens Ernsting¹; Stephan U Dombrowski², PhD; Monika Oedekoven¹; Julie L O'Sullivan¹; Melanie Kanzler³; Adelheid Kuhlmei¹, PhD; Paul Gellert¹, PhD

¹Institute of Medical Sociology, Charité - Universitätsmedizin Berlin, Berlin, Germany

²Faculty of Natural Sciences, University of Stirling, Stirling, United Kingdom

³Pfizer Deutschland GmbH, Berlin, Germany

Corresponding Author:

Clemens Ernsting

Institute of Medical Sociology

Charité - Universitätsmedizin Berlin

Chariteplatz 1

Berlin,

Germany

Phone: 49 30450529215

Fax: 49 30450529972

Email: clemens.ernsting@charite.de

Abstract

Background: Chronic conditions are an increasing challenge for individuals and the health care system. Smartphones and health apps are potentially promising tools to change health-related behaviors and manage chronic conditions.

Objective: The aim of this study was to explore (1) the extent of smartphone and health app use, (2) sociodemographic, medical, and behavioral correlates of smartphone and health app use, and (3) associations of the use of apps and app characteristics with actual health behaviors.

Methods: A population-based survey (N=4144) among Germans, aged 35 years and older, was conducted. Sociodemographics, presence of chronic conditions, health behaviors, quality of life, and health literacy, as well as the use of the Internet, smartphone, and health apps were assessed by questionnaire at home visit. Binary logistic regression models were applied.

Results: It was found that 61.25% (2538/4144) of participants used a smartphone. Compared with nonusers, smartphone users were younger, did more research on the Internet, were more likely to work full-time and more likely to have a university degree, engaged more in physical activity, and less in low fat diet, and had a higher health-related quality of life and health literacy. Among smartphone users, 20.53% (521/2538) used health apps. App users were younger, less likely to be native German speakers, did more research on the Internet, were more likely to report chronic conditions, engaged more in physical activity, and low fat diet, and were more health literate compared with nonusers who had a smartphone. Health apps focused on smoking cessation (232/521, 44.5%), healthy diet (201/521, 38.6%), and weight loss (121/521, 23.2%). The most common app characteristics were planning (264/521, 50.7%), reminding (188/521, 36.1%), prompting motivation (179/521 34.4%), and the provision of information (175/521, 33.6%). Significant associations were found between planning and the health behavior physical activity, between feedback or monitoring and physical activity, and between feedback or monitoring and adherence to doctor's advice.

Conclusions: Although there were many smartphone and health app users, a substantial proportion of the population was not engaged. Findings suggest age-related, socioeconomic-related, literacy-related, and health-related disparities in the use of mobile technologies. Health app use may reflect a user's motivation to change or maintain health behaviors. App developers and researchers should take account of the needs of older people, people with low health literacy, and chronic conditions.

(*J Med Internet Res* 2017;19(4):e101) doi:[10.2196/jmir.6838](https://doi.org/10.2196/jmir.6838)

KEYWORDS

telemedicine; eHealth; mHealth; smartphone; mobile apps; health promotion; chronic disease; health literacy; quality of life

Introduction

Mobile Health Apps

An increasing number of people take advantage of smartphones for health issues. Mobile health apps have risen in popularity, providing new opportunities to change health-related behaviors and to manage chronic conditions [1]. Typical health apps provide immediate access to health information, medication reminders, or help track progress in physical exercise regime. However, the factors related with smartphone and health app use are not yet fully understood. This study investigated the associations between demographic and health-related factors, common chronic conditions, and health behaviors with smartphone and health app use.

Research has shown associations between (1) health app use and behavior, and (2) app use and management of chronic conditions [2-6]. Krebs and Duncan [7] examined the health app use among mobile phone users in the United States. Their findings suggest an association between health app use and sociodemographic factors; that is, app users were younger, had a higher income, were better educated, were more likely to be Hispanic, and had a higher body mass index (BMI). A systematic review of mobile health interventions based primarily on randomized controlled trials by Riley et al [2], who investigated the effectiveness and theoretical background of mobile interventions such as apps for smartphones or tablets. Results demonstrated that intervention participants were more successful in changing a variety of health behaviors and behavior-related outcomes, including physical activity, smoking cessation, healthy diet, weight loss, medication adherence, improvement of blood pressure control, and improvement of blood sugar control [2].

Factors Related to the Use of Health Apps

In a cross-sectional survey, Cho et al [8] examined factors related to the use of health apps and found a significant correlation between app use and health consciousness as well as an indirect link between health literacy and health-app use efficacy on health app use. Bailey et al [9] found widespread age and health literacy-related disparities in technology access, with older and less literate individuals being less likely to own smartphones and use the Internet, especially for health reasons. In a recent longitudinal survey, Levine et al [10] found that older people (mean age 75 years) used digital health at low rates but there were modest increases from 2011 to 2014. Despite a growing body of research on health apps, there is still a lack of evidence concerning the associations of sociodemographic, medical, and behavioral factors such as health-related quality of life, health behaviors, Internet use, multiple chronic conditions, and health app use.

Health App Characteristics

In addition to individual factors, health app characteristics such as monitoring progress or reminding may be related to the actual health behavior of the users. Findings from a systematic review that analyzed characteristics of Internet-based interventions revealed that theory-based approaches were associated with increases in effect sizes [11]. In a recent content analysis,

Morrissey et al [12] examined the extent to which certain established behavior change techniques were used in apps designed to promote medication adherence. They found a range from 0 to 7 behavior change techniques implemented in these apps. Planning, prompts or cues, monitoring, and feedback on behavior were included most frequently.

Aims of the Study

The aims of this study were to (1) investigate the prevalence of smartphone and health app use, (2) identify sociodemographic, medical, and behavioral correlates of smartphone and health app use, and (3) explore the correlations between behaviors targeted by the apps and actual health behavior, and health app characteristics and actual health behavior.

Methods

Sample and Procedure

A population-based sample of 4144 individuals from Germany participated in this cross-sectional survey. Data were collected in July 2015. An external agency was authorized to run the study. The agency employed interviewers to recruit participants on their own responsibility. To enhance the representativeness of the study, interviewers got specification concerning the composition of the sample, that is, the sample was stratified for sex, age, German federal state, and education to increase representativeness. Participants had to meet the following inclusion criteria: (1) German resident, (2) sufficient German language skills, (3) aged 35 years and older. There were no exclusion criteria. After a first contact, appointments for home visits were made with a response rate of 55%. Computer-assisted personal interviews (CAPI) were conducted by external, trained interviewers at home visits. Of the interviewed individuals, 7% refused to finish the survey and their data were subsequently deleted. The mean time participants need to finish the survey was 29 min. This study was conducted in compliance with the Declaration of Helsinki; written informed consent was obtained from participants [13].

Measures

Sociodemographics

Sex, age, height, weight, education (International Standard Classification of Education, ISCED) [14], occupation, income, and first language were assessed by standard survey items. Post-tax household income by month was categorized: low <€100; moderate €100-3600; high >€3600 (1 Euro=1.1 US dollar [October 6, 2016]).

Chronic Conditions

Chronic conditions were assessed by asking participants: "Do you suffer from one or more of the following chronic conditions: (1) cardiovascular disease, (2) cancer, (3) respiratory diseases, (4) diseases of the musculoskeletal system, (5) major depression, (6) chronic pain, (7) diabetes, (8) hypertension, and (9) other chronic condition." Furthermore, BMI was calculated by using self-reported weight and height ($BMI = \text{weight (in kg)} / \text{height (in squared meter)}$). A BMI of 30 or above was considered obese, in accordance with the World Health Organization (WHO) definition [15]. Chronic conditions were summed up to the

variable *multiple chronic conditions* ranging from “none,” “one,” “two,” to “three or more.”

Health Behaviors

Health behaviors were assessed by providing a list of common health-related behaviors (ie, smoking, physical activity, balanced diet, low fat diet, and adherence to doctor’s advice). These items were derived from the German Aging Survey (2013) and were adjusted for our survey. *Adherence to doctor’s advice* was taken from the 16-item short-form of the European Health Literacy Survey Questionnaire (HLS-EU-Q) instrument [16]. For *smoking*, participants were asked: “Do you smoke on a daily basis?” To assess *physical activity*, participants were asked: “Are you regularly physically active (following WHO recommendation, ie, 30 min of moderate activity at least 5 times per week or 30 min of intensive activity at least 3 times per week [17])?” *Balanced diet* was measured by asking participants: “Do you follow a balanced diet, that is, eat fruits and vegetables with every meal and including many wholegrain products?” *Low fat diet* was assessed by asking participants: “Do you follow a low fat diet, that is, eat few animal fats, peanuts, potato chips, and convenience food?” *Adherence to doctor’s advice* was assessed by asking participants: “On a scale from very easy to very difficult, how easy would you say it is to follow physician’s and pharmacist’s instructions?” The answer had a 4-point response format on a Likert scale. The answers “very easy” and “easy” were defined as adherence (coded 1=present) and the answers “difficult” and “very difficult” were defined as problematic adherence (coded 0=absent).

Health-Related Quality of Life and Perceived Health Literacy

Health-related quality of life was assessed by the European Health Interview Survey-Quality of Life (EUROHIS-QOL) 8-item index [18] with a Cronbach alpha of .90 in this analysis. Example items included: “How would you rate your quality of life?” and “How satisfied are you with your health?” All answers were given on a 5-point Likert scale. *Health literacy* was assessed by the 16-item short-form of the HLS-EU-Q instrument with a Cronbach alpha of .87 and a range from 0 to 50 indicating the perceived capability of an individual to acquire, understand, and act on health information [19]. An example item was: “On a scale from very easy to very difficult, how easy would you say it is to find information on treatments of illnesses that concern you?” Answers had a 4-point response format on a Likert scale.

Internet Use

Internet use was assessed by asking participants: “Do you use the Internet for obtaining information about health conditions or other health-related issues?” (coded as 1=present, 0=absent).

Smartphone and App Use

Smartphone use was assessed by asking participants: “Do you own a smartphone, that is, an Internet-compatible cell phone?” (coded 1=present, 0=absent).

Health app use and *behaviors targeted by the apps* were measured by asking participants: “Think about the last 12

months. Did you use smartphone apps to improve one of the following behaviors?: (1) to quit smoking, (2) to be regularly physically active, (3) to maintain a healthy diet, (4) to reduce weight, (5) to take medications regularly, (6) to improve blood pressure control, (7) to improve blood sugar control, and (8) I do not use smartphone apps to improve behaviors.” The health behaviors and behavior-related outcomes targeted by the apps were based on findings from a systematic review of health apps [2]. People who chose one or more of these behaviors were classified as health app users (coded 1=present) and people who chose “I do not use smartphone apps to improve behaviors” were classified as nonusers (coded 0=absent). *Behaviors targeted by the apps* were coded 1=present, 0=absent. There was no option to name alternative behaviors targeted by the apps.

App characteristics were assessed by asking participants: “Did the health apps you used contain one of the following characteristics: (1) provision of information on the target behavior, (2) motivational messages, (3) goal setting and action planning, (4) reminder, (5) information on the current status and individual progress, and (6) I do not use any of these app characteristics.” App characteristics were chosen in accordance to the behavior change technique taxonomy by Abraham and Michie [20], that is, *providing information, planning, reminding, providing feedback or monitoring*. Furthermore, the additional app characteristic *prompting motivation* was derived from a systematic review of mobile health interventions conducted by Riley et al [2]. It refers to the mobile intervention *messages containing motivational support* reviewed by Riley and encompasses the behavior change techniques such as the *provision of general encouragement or provision of contingent rewards* as outlined by Abraham. Participants were asked to choose one or more of the characteristics, coded 1=present and 0=absent. There was no option to name alternative app characteristics.

Statistical Analyses

We analyzed the entire sample as well as specific subgroups. Beyond smartphone users and nonusers, we compared health app users and non-health app users among the smartphone users. Binary logistic regressions with smartphone use (n=4144) and health app use (n=2538; subsample of smartphone users) as outcomes were conducted. Covariates were sex, age, number of chronic conditions, perceived health literacy, health-related quality of life, and Internet use. Other covariates were the health behaviors smoking, physical activity, balanced diet, and low fat diet.

Furthermore, we explored correlations between behaviors targeted by the apps and app characteristics with reported behavior by applying binary logistic regression models. We ran separate analyses for each behavior, such as for smoking, physical activity, balanced diet, low fat diet, and adherence to doctor’s advice. Health behaviors and behavior-related outcomes targeted by the apps were regular physical activity, smoking cessation, healthy diet, weight loss, medication adherence, improving blood pressure control, and improving blood sugar control. App characteristics comprised providing information, prompting motivation, planning, reminding, and feedback or

monitoring. Sex, age, behaviors targeted by the apps, and app characteristics were included as covariates into the analyses.

Additionally, we applied Hochberg's multistep procedure which is a slightly less restrictive alternative to the Bonferroni approach to correct for multiplicity (Hochberg [21]; for a discussion, see Streiner [22]).

Results

Characterization of the Sample

A total of 4144 individuals completed the population-based nationwide survey (see Table 1). The mean age was 57 years (SD 13.5) and 50.96% (2112/4144) were women. It was found that 68.97% (2858/4144) had a vocational qualification, 18.15% (752/4144) had a university degree, and 12.89% (534/4144) had no or basic qualification. Most participants were working full-time (2224/4144, 53.67%) and had a medium household income (1930/4144, 46.57%).

The majority of the sample (2231/4144, 53.84%) reported no chronic conditions, whereas 30.84% (1278/4144) reported having one, 11.25% (466/4144) two, and 4.08% (169/4144) three or more chronic conditions. The most common chronic conditions were hypertension (763/4144, 18.41%), musculoskeletal conditions (385/4144, 9.29%), and cardiovascular diseases (376/4144, 9.07%). The mean BMI was 24.9 (SD 3.5) and 6.49% (269/4144) of the participants were classified as obese ($BMI \geq 30$).

In total, 87.02% (3606/4144) of participants reported adhering to doctor's advice, 60.85% (2522/4144) were on a balanced diet, and 48.82% (2023/4144) were on a low fat diet. Furthermore, 38.95% (1614/4144) engaged in physical activity on a regular basis, and 28.50% (1181/4144) smoked daily. Average health-related quality of life was 3.9 out of 5 (SD 0.6) and mean health literacy was 33.5 out of 50 (SD 7.4).

The majority of the sample (2538/4144, 61.25%) reported using a smartphone and 20.53% (521/2538) of smartphone users also reported using health apps.

Table 1. Sample characteristics by smartphone and health app use.

Item	Total sample	Total sample		Smartphone users	
		No Smartphone ^a use	Smartphone use	No health app use	Health app use
	100.00% (4144/4144)	38.75% (1606/4144)	61.25% (2538/4144)	79.47% (2017/2538)	20.53% (521/2538)
	N=4144	n=1606	n=2538	n=2017	n=521
Female (vs male), n (%)	2112 (50.97)	859 (53.49)	1253 (49.37)	984 (48.79)	269 (51.6)
Age in years (SD^b)	57 (14)	68 (11)	50 (10)	51 (10)	47.9 (10)
Educational level (ISCED^c), n (%)					
No or basic qualification	534 (12.89)	229 (14.26)	305 (12.02)	244 (12.10)	61 (11.7)
Vocational qualification	2858 (68.97)	1144 (71.23)	1714 (67.53)	1373 (68.07)	341 (65.5)
University degree	752 (18.15)	233 (16.06)	519 (20.45)	400 (19.83)	119 (22.8)
Occupational status, n (%)					
Working full-time	2224 (53.67)	309 (19.24)	1915 (75.45)	1503 (74.52)	412 (79.1)
Working part-time	434 (10.47)	135 (8.41)	299 (11.78)	245 (12.17)	54 (10.4)
Not working	198 (4.78)	81 (5.04)	117 (4.61)	92 (4.56)	25 (4.8)
Retired	1287 (31.06)	1081 (67.31)	206 (8.12)	177 (9)	29 (5.6)
In school	1 (0.02)	0 (0.00)	1 (0.04)	0 (0.00)	1 (0.20)
Monthly post-tax household income^d, n (%)					
Low	1342 (32.38)	801 (49.88)	541 (21.32)	423 (20.97)	118 (22.6)
Medium	1930 (46.57)	523 (32.57)	1407 (55.44)	1125 (55.78)	282 (54.1)
High	290 (7.00)	40 (2.49)	250 (9.85)	201 (9.97)	49 (9.4)
No answer	582 (14.04)	242 (15.07)	340 (13.40)	268 (13.29)	72 (13.8)
First language, n (%)					
German	3773 (91.05)	1499 (93.34)	2274 (89.60)	1826 (90.53)	448 (86.0)
Other	371 (8.95)	107 (6.66)	264 (10.40)	191 (9.47)	73 (14.0)
Chronic conditions, n (%)					
Cardiovascular disease	376 (9.07)	275 (17.12)	101 (3.98)	81 (4.02)	20 (3.8)
Cancer	79 (1.91)	49 (3.05)	30 (1.18)	24 (1.19)	6 (1.2)
Respiratory diseases	232 (5.60)	117 (7.29)	115 (4.53)	94 (4.66)	21 (4.0)
Musculoskeletal system conditions	385 (9.29)	277 (17.25)	108 (4.26)	84 (4.16)	24 (4.6)
Depression	128 (3.09)	43 (2.68)	85 (3.35)	62 (3.07)	23 (4.4)
Chronic pain	310 (7.48)	180 (11.21)	130 (5.12)	103 (5.11)	27 (5.2)
Diabetes	361 (8.71)	201 (12.52)	160 (6.30)	117 (5.80)	43 (8.3)
Hypertension	763 (18.41)	416 (25.90)	347 (13.67)	266 (13.19)	81 (15.5)
Obesity	269 (6.49)	115 (7.16)	154 (6.07)	111 (5.50)	43 (8.3)
Multiple chronic conditions, n (%)					
None	2231 (53.84)	553 (34.43)	1577 (62.14)	1270 (62.96)	307 (58.9)
One	1278 (30.84)	599 (37.30)	684 (26.95)	546 (27.07)	138 (26.5)
Two	466 (11.25)	300 (18.68)	213 (8.39)	153 (7.59)	60 (11.5)
Three or more	169 (4.08)	154 (9.59)	64 (2.52)	48 (2.38)	16 (3.1)
Health behaviors, n (%)					
Smoking	1181 (28.50)	364 (22.67)	817 (32.19)	658 (32.62)	159 (30.5)

Item	Total sample	Total sample		Smartphone users	
		No Smartphone ^a use	Smartphone use	No health app use	Health app use
	100.00% (4144/4144)	38.75% (1606/4144)	61.25% (2538/4144)	79.47% (2017/2538)	20.53% (521/2538)
	N=4144	n=1606	n=2538	n=2017	n=521
Physical activity	1614 (38.95)	468 (29.14)	1146 (45.15)	863 (42.79)	283 (54.3)
Balanced diet	2522 (60.86)	992 (61.77)	1530 (60.28)	1189 (58.95)	341 (65.5)
Low fat diet	2023 (48.82)	846 (52.68)	1177 (46.38)	895 (44.37)	282 (54.1)
Adherence to doctor's advice	3606 (87.01)	1338 (83.31)	2268 (89.36)	1803 (89.39)	465 (89.3)
BMI ^c (SD)	24.9 (3.5)	25.1 (3.7)	24.9 (3.5)	24.8 (3.3)	25.0 (3.9)
Health-related quality of life (SD)	3.9 (0.6)	3.8 (0.6)	4.1 (0.6)	4.1 (0.6)	4.0 (0.6)
Health literacy (SD)	33.5 (7.4)	33.1 (7.8)	33.0 (6.7)	34.9 (6.7)	35.7 (6.8)
Behaviors targeted by the apps, n (%)					
Smoking cessation	232 (5.60)	0 (0.00)	232 (9.14)	0 (0.00)	232 (44.5)
Regular physical activity	89 (2.15)	0 (0.00)	89 (3.51)	0 (0.00)	89 (17.1)
Healthy diet	201 (4.85)	0 (0.00)	201 (7.92)	0 (0.00)	201 (38.6) (201/521)
Weight loss	121 (2.92)	0 (0.00)	121 (4.77)	0 (0.00)	121 (23.2)
Medication adherence	49 (1.18)	0 (0.00)	49 (1.93)	0 (0.00)	49 (9.4)
Blood pressure control	36 (0.87)	0 (0.00)	36 (1.42)	0 (0.00)	36 (6.9)
Blood sugar control	13 (0.31)	0 (0.00)	13 (0.51)	0 (0.00)	13 (2.5)
App characteristics, n (%)					
Providing information	175 (4.22)	0 (0.00)	175 (6.90)	0 (0.00)	175 (33.6)
Prompting motivation	179 (4.32)	0 (0.00)	179 (7.05)	0 (0.00)	179 (34.4)
Planning	264 (6.37)	0 (0.00)	264 (10.40)	0 (0.00)	264 (50.7)
Reminding	188 (4.54)	0 (0.00)	188 (7.41)	0 (0.00)	188 (36.1)
Feedback or monitoring	94 (2.27)	0 (0.00)	94 (3.70)	0 (0.00)	94 (18.0)

^aSmartphone refers to all kinds of Internet-compatible cell phones.

^bSD: standard deviation.

^cISCED: International Standard Classification of Education.

^dPost-tax household income: Low <€100, moderate €100-3600, high >€3600 (1 Euro=1.1 US dollar; October 6, 2016).

^eBMI: body mass index.

What Factors Are Associated With Smartphone Use?

Results from a binary logistic regression revealed that smartphone users were significantly younger (OR 0.92, $P<.001$) and were more likely to have a university degree (OR 1.69, $P=.003$). Furthermore, they were more likely to use the Internet for research about health issues (OR 3.24, $P<.001$) compared with non-smartphone users (see Table 2). Employment was associated with smartphone use, that is, individuals reporting to work part-time (OR 0.60, $P<.001$), not to work (OR 0.48, $P<.001$), and to be retired (OR 0.23, $P<.001$) were less likely to use smartphones compared with full-time workers.

Participants who had a university degree were more likely to own a smartphone compared with people without or with basic qualification (OR 1.69, $P=.003$). Smartphone users were also more engaged in physical activity (OR 1.26, $P=.02$) but they were less likely to follow a low fat diet (OR 0.67, $P<.001$). Higher health-related quality of life was associated with smartphone use (OR 1.24, $P=.03$) and there was an association between health literacy (OR 1.05, $P<.001$) and smartphone use. Smartphone use was not associated with sex; mother tongue; chronic conditions; or the health behaviors smoking, balanced diet, and adherence.

Table 2. Multivariate associations with smartphone and health app use.

Item	Smartphone ^a use (N=4144)			Health app use (n=2538)		
	Odds ratio	95% CI	P	Odds ratio	95% CI	P
Intercept	21.74 ^b		<.001	0.47		0.19
Female (vs male)	0.92	0.76-1.1	0.37	1.15	0.94-1.42	0.18
Age	0.92 ^b	0.91-0.93	<.001	0.97 ^b	0.96-0.98	<.001
Mother tongue (other vs German)	1.18	0.85-1.64	0.32	1.5	1.11-2.02	0.008
Internet use	3.24 ^b	2.6-4.03	<.001	1.68 ^b	1.37-2.06	<.001
Employment						
Working full-time	Reference ^c			Reference		
Working part-time	0.60 ^b	0.46-0.79	<.001	0.82	0.89-1.15	0.25
Not working	0.48 ^b	0.33-0.7	<.001	0.99	0.61-1.62	0.98
Retired	0.23 ^b	0.17-0.3	<.001	1.1	0.67-1.81	0.71
Education						
No or basic qualification	Reference			Reference		
Vocational qualification	1.04	0.8-1.35	0.79	0.89	0.65-1.23	0.49
University degree	1.69	1.2-2.37	0.003	1.04	0.72-1.51	0.84
Multiple chronic conditions						
None	Reference			Reference		
One	1.02	0.83-1.26	0.85	1.34	1.04-1.72	0.02
Two	0.97	0.72-1.3	0.83	2.36 ^b	1.63-3.42	<.001
Three or more	0.78	0.49-1.25	0.3	2	1.07-3.75	0.03
Health behaviors						
Smoking	1.11	0.9-1.35	0.35	0.98	0.78-1.22	0.84
Physical activity	1.26	1.03-1.52	0.02	1.38	1.11-1.72	0.003
Balanced diet	1.03	0.85-1.26	0.75	1.06	0.84-1.34	0.63
Low fat diet	0.67 ^b	0.55-0.81	<.001	1.33	1.06-1.66	0.01
Adherence (doctor's advice)	0.96	0.72-1.27	0.76	0.81	0.57-1.14	0.23
Health-related quality of life	1.24	1.02-1.5	0.03	0.94	0.76-1.16	0.55
Health literacy	1.05 ^b	1.04-1.07	<.001	1.02	1-1.03	0.04

^aSmartphone refers to all kinds of Internet-compatible cell phones.^bStill significant after correction for multiplicity.^cReference: reference category set to 1.

Which Factors Are Associated With Health App Use?

Out of all smartphone users, 20.53% (n=521/2538) reported using health apps. Health app users were significantly younger (OR 0.97, $P<.001$) and used the Internet for research about health issues (OR 1.68, $P<.001$) more often than those not using health apps (see Table 2). Participants who had another language than German as mother tongue were more likely to use health apps (OR 1.50, $P=.008$). Individuals reporting chronic conditions were more likely to use health apps: individuals with one (OR 1.34, $P=.02$), two (OR 2.36, $P<.001$), and three or more chronic

conditions (OR 2.00, $P=.03$) were more likely to use a health app compared with individuals without chronic conditions. Regarding health behaviors, a significant association with health app use was only found for physical activity (OR 1.38, $P=.003$) and low fat diet (OR 1.33, $P=.01$). Finally, health app users were more health literate than nonusers (OR 1.02, $P=.04$). Health app use was not associated with sex, employment, education, smoking, balanced diet, and adherence to doctor's advice; and health-related quality of life.

Which App Characteristics Do the Different Health Apps Include?

Health app users reported using apps to support changes in smoking cessation (232/521, 44.5%), healthy diet (201/521, 38.6%), weight loss (121/521, 23.2%), physical activity (89/521, 17.08%), medication adherence (49/521, 9.4%), blood pressure control (36/521, 6.9%), and blood sugar control (13/521, 2.5%; see [Table 1](#)). The most common app characteristic was planning (264/521, 50.7%), followed by reminding (188/521, 36.08%), prompting motivation (179/521, 34.4%), providing information (175/521, 33.6%), and feedback or monitoring (94/521, 18.0%).

Which Behaviors Targeted by the Apps Are Associated With Actual Behavior?

[Multimedia Appendix 1](#) shows associations of behaviors targeted by the apps and actual behavior. There are five separate analyses. All app users were included in each analysis ($n=521$). Covariates per the model are noted down the column on the left side, which is sex, age, behaviors targeted by the apps, and app characteristics.

Participants who were smokers were more likely to use health apps for physical activity (OR 14.44, $P=.001$; see [Multimedia Appendix 1](#)). Conversely, participants who reported to engage in regular physical activity were less likely to use apps targeting physical activity (OR 0.54, $P=.03$) and weight loss (OR 0.48, $P=.003$) and were more likely to use apps targeting smoking cessation (OR 3.01, $P<.001$) and healthy diet (OR 1.65, $P=.02$) compared with inactive participants. Participants who were engaged in physical activity were also significantly younger (OR 0.97, $P=.001$) than inactive ones. Individuals following a low fat diet were more likely to use apps on smoking cessation (OR 2.34, $P<.001$) and healthy diet (OR 1.83, $P=.007$), but were less likely to use apps targeting weight loss (OR 0.31, $P<.001$). Individuals engaging in a balanced diet were using apps for smoking cessation (OR 2.16, $P=.001$) and healthy diet (OR 1.97, $P=.003$) more often, but used apps for weight loss (OR 0.31, $P<.001$) less often. Participants who reported adherence to doctor's advice were more likely to use apps on smoking cessation (OR 3.07, $P=.005$). There were no other associations between behaviors targeted by the apps and actual behavior beyond the reported findings.

Which App Characteristics Are Associated With Actual Behavior?

[Multimedia Appendix 1](#) shows associations of behaviors targeted by the apps and actual behavior. There are five separate analyses. All app users were included in each analysis ($n=521$). Covariates per the model are noted down the column on the left side, that is sex, age, behaviors targeted by the apps, and app characteristics.

Participants using health apps including the app characteristics, planning (OR 1.51, $P=.04$) and feedback or monitoring (OR 2.15, $P=.007$), were more likely to engage in physical activity (see [Multimedia Appendix 1](#)). Moreover, people who reported to use apps including the app characteristic feedback or monitoring were more likely to adhere to doctor's advice (OR 5.94, $P=.02$). There were no other associations between app characteristics and actual behavior beyond the reported findings.

Discussion

Principal Findings

Two-thirds of the participants owned a smartphone, but only one in five reported using apps for health-related purposes. Those using smartphones tended to differ in their demographic profile compared with nonusers. They tended to be younger, more health literate, had a higher socioeconomic status, and felt better about their life. Importantly, those who used health apps appeared to be different from smartphone owners who did not use health apps. Health app users were younger, less likely to be native German speakers, more likely to conduct health-related Internet searches, more likely to suffer from chronic conditions, and more likely to follow health behaviors. Although using a health app may reveal an interest in health and certain health behaviors, the apps that people were using were not necessarily reflective of the health behaviors they were performing. Hence, health app use may reflect a user's motivation to change health behaviors. Finally, we found weak associations of health behaviors and health app characteristics, with planning and self-monitoring being the only significant exceptions.

Smartphone Use

The extent of smartphone use found in our study was comparable with the findings of the 2015 US Pew Research Center Survey [23], where 64% of the participants were using a smartphone, compared with 61.25% (2538/4144) in our sample. A 2012 Pew Research Center Survey found that 45% of the US adults owned a smartphone [24]. In a recent longitudinal survey, Levine et al [10] found that older people with a mean age of 75 years used digital health at low rates. There were modest increases from 2011 through 2014. This underlines the rising importance of mobile technology in people's lives. The 2015 Pew Survey also identified age as a correlates for smartphone use, which was confirmed in this study: younger people were more engaged [23]. Furthermore, our findings contribute to previous findings on literacy-related disparities in access to mobile technologies by revealing an association between smartphone use and health literacy [9].

We contributed to the evidence of correlates of smartphone use with health-related quality of life, Internet research behavior, physical activity, and low fat diet. The relationship between health-related quality of life and smartphone use may be explained by socioeconomic variables. People with higher education and higher incomes score higher in quality of life [25]. Indeed, we found a university degree and a full-time employment to be associated with smartphone use. Further research on this interplay is needed. Smartphones offer direct and easy access to the Internet, which may be a possible explanation for the correlation between Internet research behavior and smartphone use. This finding underlines the importance of the Internet and mobile technologies for health issues.

Health App Use

Regarding health app use, the 2012 Pew Survey found that 19% of the mobile phone users took advantage of health apps, consistent with our finding of 20.53% (521/2538) [24]. In both

studies, 4 out of 5 participants were not using health apps. Results from a recent study by Krebs and Duncan [7] revealed a health app use of nearly 60%. Krebs and Duncan investigated a younger population with a mean age of 40 years (mean age of our sample was 57 years) which may be a possible explanation for the difference. Nonetheless, this underlines the potential of health apps in the future, as they seem to be very popular among young people; yet also older people are increasingly engaged in mobile technologies [10,23].

We found native German speakers to be less likely to use health apps compared with nonnative speakers. This finding is consistent with Krebs and Duncan, who found in their survey among a US sample that hispanics were more likely to use health apps. Greater use by minorities in each country could be due to difficulties accessing the health care system as well as greater use of Internet on smartphones versus home connections.

We found that the association of health app use, age, and health-literacy in our survey was in accordance with results from Krebs and Duncan [7] and Bailey et al [9], who found age and literacy-related disparities in the use of mobile health apps. The role of age in the use of health apps highlights that the relevance of new potential ways of supporting health topics is growing in the future. However, app developers should not forget about older people, especially because health issues become increasingly important in later years [26]. Tailored theory-based health app interventions may be a way to reach people from all ages, and app developers should take the needs of older people and people with low health literacy into account to further decrease the “digital gap” between users and nonusers [2].

There is a need for mobile solutions for disease management. Our findings indicate that people suffering from chronic conditions were more likely to use health apps. We found this correlation not only when we were looking at smartphone users ($n=2538$) but also when we were investigating the whole sample ($N=4144$). Another finding from this study is the correlation between physical activity and health app use, that is, people who exercise are more likely to use health apps. Both findings indicate that people use mobile technologies that help them to manage diseases and certain health behaviors.

Looking at the behaviors targeted by the apps, the 2012 Pew Survey identified exercise, diet, and weight loss apps as the most popular behaviors [24]. In this study, we found smoking cessation, healthy diet, and weight loss to be the most popular behaviors targeted by the apps. Considering that cardiovascular diseases are the most common cause of death in western countries, this is an encouraging finding [27].

Behaviors Targeted by the Apps and Characteristics of the Apps

These analyses were conducted to explore the correlations between health behaviors targeted by the apps, and health app characteristics and actual health behavior, rather than to strictly test hypotheses. Future research should establish specific hypotheses based on our results. In this study, health app use was seen as motivation to change or maintain health behaviors using mobile technologies. One-fifth of the smartphone users

reported use of health apps, which may reflect an interest in health and health behavior change. Conversely, using health apps was not necessarily related to an active engagement in the respective health behavior. It appears that the apps which people have on their smartphones indicate what they want to change, not their actual behavior. Concerning correlations of health app characteristics with actual behavior there were associations between planning and physical activity, between feedback or monitoring and physical activity, and between feedback or monitoring and adherence to doctor's advice. These findings relate to other studies that found self-monitoring to be associated with health behaviors—although our findings were limited to certain health behaviors [28].

Strengths and Limitations

A strength of this study was the large nationwide sample. It was a stratified sample by sex, age, German federal state, and education to increase the representativeness. Moreover, we provided an overview about the extent of smartphone and app use. Finally, we supplied novel information, contributing to the research field, by including health-related quality of life, health literacy, multiple chronic conditions, and health behaviors as well as information about health app characteristics aimed at changing health behaviors.

Our research was limited by the design of the survey. Due to the cross-sectional character of the survey, changes could not be examined, including associations of behaviors targeted by the apps and app characteristics with actual behavior. As the findings suggest, smartphones and app use were more common among younger respondents; however, the survey did not include people aged below 35 years, as our focus was individuals with and without chronic conditions. Income could not be used in multivariate predictions because a critical number of participants (582/4144, 14.04%) did not answer that question. Previous studies found app use to be associated with higher income [7]. Another aspect is that the nature of the survey questions limited the conclusions that could be drawn from the results. For example, BMI was calculated based on self-reported weight and height which is a possible source of bias. Smoking was defined as daily smoking which overlooks the complexity of smoking behavior, for example, social smoking. This limits the conclusions that can be drawn from the finding that smoking status was not associated with the use of apps for smoking cessation.

Furthermore, a large number of analyses were conducted increasing the probability of type I errors (ie, stating an effect when none was present). However, our study has an explorative character. Nonetheless, in addition to the uncorrected results, we decided to report the multiplicity corrected results following a recommendation by Streiner [22], who discussed arguments for and against a correction of multiplicity.

Namely, our ambition is to provide a huge basis that can be further examined in future research.

Implications for Further Research, Policy, and Practice

Our findings have shown that two-thirds of the participants aged above 35 years used a smartphone, but of these only 1 in 5 participants has been using health apps. Thus, when apps are

designed for public health, many people cannot be reached because they either do not have a smartphone or they do not use it for health reasons. Furthermore, lower levels of health literacy and age appropriateness need to be taken into account when designing health apps. Special regard for people with multiple chronic conditions is needed. Tailoring the training regime to the special needs of chronically ill people have not yet been implemented in many health and lifestyle apps. Health apps may play a major role for future health participation of individuals as well as for their autonomy and health literacy.

We gained insight into the relationship between app characteristics and actual behavior. For future research, the effects of health app use on health behaviors should be analyzed by applying longitudinal or experimental research designs. An important point, which should be investigated, is the association between the presence or absence of chronic conditions and the use of specific health apps or app characteristics. Future research should investigate additional behaviors targeted by apps such as sleep tracking or managing doctoral appointments as our questions constrained the survey results to certain behaviors based on the review by Riley et al [2].

The influence of health app use on health-related quality of life, health behaviors, and disease management should be

investigated in greater detail. More research on effective app characteristics is needed. Information on the effectiveness of health apps based on evidence from rigorous research designs should be provided for users. This would increase the transparency in the market, given the huge number of apps available. Not only governmental guidelines and regulations, but also WHO recommendations could help people choose effective and appropriate apps [29]. A first attempt in this direction is the MEDDEV Guideline 2.1/6 for the European market and a guideline from the Food and Drug Administration (FDA) for the US market [30,31]. Although these are not legally binding, they offer an orientation for developers and consumers of mobile health apps.

We found age- and literacy-related disparities in the use of mobile technologies. Thus, app developers and researchers should take the needs of older people with multiple chronic conditions and people with low health literacy, for example, by providing tailored health apps tested in intervention studies, into account [7,9]. Likewise, campaigns should be launched aimed at training older adults in mobile technologies and enhancing the health literacy of the population to decrease inequalities resulting from technological progress.

Acknowledgments

This material is based on research conceptually and financially supported by the Pfizer Deutschland GmbH. Any findings and conclusions expressed in this publication are those of the authors and do not necessarily reflect the views of the Pfizer Deutschland GmbH.

Conflicts of Interest

None declared.

Multimedia Appendix 1

[PDF File (Adobe PDF File), 65KB - [jmir_v19i4e101_app1.pdf](#)]

References

1. Smith A. Pew Research Center. 2016. Smartphone ownership 2013 URL: <http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013/> [accessed 2016-10-19] [WebCite Cache ID 6INPjGXQV]
2. 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](https://doi.org/10.1007/s13142-011-0021-7)] [Medline: [21796270](#)]
3. Wang J, Wang Y, Wei C, Yao NA, Yuan A, Shan Y, et al. Smartphone interventions for long-term health management of chronic diseases: an integrative review. *Telemed J E Health* 2014 Jun;20(6):570-583. [doi: [10.1089/tmj.2013.0243](https://doi.org/10.1089/tmj.2013.0243)] [Medline: [24787747](#)]
4. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. *J Med Internet Res* 2012 Nov 21;14(6):e161 [FREE Full text] [doi: [10.2196/jmir.2171](https://doi.org/10.2196/jmir.2171)] [Medline: [23171838](#)]
5. McLean G, Band R, Saunderson K, Hanlon P, Murray E, Little P, et al. Digital interventions to promote self-management in adults with hypertension systematic review and meta-analysis. *J Hypertens* 2016 Apr;34(4):600-612 [FREE Full text] [doi: [10.1097/HJH.0000000000000859](https://doi.org/10.1097/HJH.0000000000000859)] [Medline: [26845284](#)]
6. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. *Behav Modif* 2015 Nov;39(6):785-804 [FREE Full text] [doi: [10.1177/0145445515595198](https://doi.org/10.1177/0145445515595198)] [Medline: [26187164](#)]
7. Krebs P, Duncan DT. Health app use among us mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](#)]

8. Cho J, Park D, Lee HE. Cognitive factors of using health apps: systematic analysis of relationships among health consciousness, health information orientation, eHealth literacy, and health app use efficacy. *J Med Internet Res* 2014 May 09;16(5):e125 [FREE Full text] [doi: [10.2196/jmir.3283](https://doi.org/10.2196/jmir.3283)] [Medline: [24824062](https://pubmed.ncbi.nlm.nih.gov/24824062/)]
9. Bailey SC, O'Connor R, Bojarski EA, Mullen R, Patzer RE, Vicencio D, et al. Literacy disparities in patient access and health-related use of Internet and mobile technologies. *Health Expect* 2015 Dec;18(6):3079-3087 [FREE Full text] [doi: [10.1111/hex.12294](https://doi.org/10.1111/hex.12294)] [Medline: [25363660](https://pubmed.ncbi.nlm.nih.gov/25363660/)]
10. Levine DM, Lipsitz SR, Linder JA. Trends in seniors' use of digital health technology in the United States, 2011-2014. *J Am Med Assoc* 2016 Aug 02;316(5):538-540. [doi: [10.1001/jama.2016.9124](https://doi.org/10.1001/jama.2016.9124)] [Medline: [27483069](https://pubmed.ncbi.nlm.nih.gov/27483069/)]
11. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010 Feb 17;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
12. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior change techniques in apps for medication adherence: a content analysis. *Am J Prev Med* 2016 May;50(5):e143-e146. [doi: [10.1016/j.amepre.2015.09.034](https://doi.org/10.1016/j.amepre.2015.09.034)] [Medline: [26597504](https://pubmed.ncbi.nlm.nih.gov/26597504/)]
13. World Medical Association. WMA declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Med Assoc* 2013 Nov 27;310(20):2191-2194. [doi: [10.1001/jama.2013.281053](https://doi.org/10.1001/jama.2013.281053)] [Medline: [24141714](https://pubmed.ncbi.nlm.nih.gov/24141714/)]
14. Uis.unesco. International standard classification of education (ISCED) 2011 URL: <http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf> [accessed 2016-10-19] [WebCite Cache ID 6lNSHFkc3]
15. World Health Organization. WHO. Obesity and overweight URL: <http://www.who.int/mediacentre/factsheets/fs311/en/> [accessed 2016-10-19] [WebCite Cache ID 6lNSZn4fq]
16. Deutsches Zentrum für Altersfragen. DZA. 2013. Deutscher Alterssurvey (DEAS): Instrumente der DEAS-Erhebung 2011 URL: https://www.dza.de/fileadmin/dza/pdf/fdz/DEAS2011_Instrumentenband.pdf [accessed 2017-03-17] [WebCite Cache ID 6p1wS0GfC]
17. WHO. Global Recommendations on Physical Activity for Health URL: http://apps.who.int/iris/bitstream/10665/44399/1/9789241599979_eng.pdf [accessed 2016-10-19] [WebCite Cache ID 6lNSqH5Da]
18. da Rocha NS, Power MJ, Bushnell DM, Fleck MP. The EUROHIS-QOL 8-item index: comparative psychometric properties to its parent WHOQOL-BREF. *Value Health* 2012 May;15(3):449-457 [FREE Full text] [doi: [10.1016/j.jval.2011.11.035](https://doi.org/10.1016/j.jval.2011.11.035)] [Medline: [22583455](https://pubmed.ncbi.nlm.nih.gov/22583455/)]
19. Sørensen K, Van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, HLS-EU Consortium. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC Public Health* 2013 Oct 10;13:948 [FREE Full text] [doi: [10.1186/1471-2458-13-948](https://doi.org/10.1186/1471-2458-13-948)] [Medline: [24112855](https://pubmed.ncbi.nlm.nih.gov/24112855/)]
20. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol* 2008 May;27(3):379-387. [doi: [10.1037/0278-6133.27.3.379](https://doi.org/10.1037/0278-6133.27.3.379)] [Medline: [18624603](https://pubmed.ncbi.nlm.nih.gov/18624603/)]
21. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 1988;75(4):800-802 [FREE Full text] [doi: [10.1093/biomet/75.4.800](https://doi.org/10.1093/biomet/75.4.800)]
22. Streiner DL. Best (but oft-forgotten) practices: the multiple problems of multiplicity-whether and how to correct for many statistical tests. *Am J Clin Nutr* 2015 Oct;102(4):721-728 [FREE Full text] [doi: [10.3945/ajcn.115.113548](https://doi.org/10.3945/ajcn.115.113548)] [Medline: [26245806](https://pubmed.ncbi.nlm.nih.gov/26245806/)]
23. Smith A. Pew Research Center. 2015. U.S. Smartphone Use in 2015 URL: <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/> [accessed 2016-10-19] [WebCite Cache ID 6lNTNdY5H]
24. Fox S, Duggan M. Pew Research Center. 2012. Mobile health 2012 URL: <http://www.pewinternet.org/2012/11/08/mobile-health-2012/> [accessed 2016-10-19] [WebCite Cache ID 6lNTehQED]
25. Kim J, Park E. Impact of socioeconomic status and subjective social class on overall and health-related quality of life. *BMC Public Health* 2015 Aug 15;15:783 [FREE Full text] [doi: [10.1186/s12889-015-2014-9](https://doi.org/10.1186/s12889-015-2014-9)] [Medline: [26275823](https://pubmed.ncbi.nlm.nih.gov/26275823/)]
26. Kruse CS, Mileski M, Moreno J. Mobile health solutions for the aging population: a systematic narrative analysis. *J Telemed Telecare* 2016 Jun 01:1-13. [doi: [10.1177/1357633X16649790](https://doi.org/10.1177/1357633X16649790)] [Medline: [27255207](https://pubmed.ncbi.nlm.nih.gov/27255207/)]
27. Beard JR, Officer A, de Carvalho I, Sadana R, Pot AM, Michel J, et al. The World report on ageing and health: a policy framework for healthy ageing. *Lancet* 2016 May 21;387(10033):2145-2154. [doi: [10.1016/S0140-6736\(15\)00516-4](https://doi.org/10.1016/S0140-6736(15)00516-4)] [Medline: [26520231](https://pubmed.ncbi.nlm.nih.gov/26520231/)]
28. Harkin B, Webb TL, Chang BP, Prestwich A, Conner M, Kellar I, et al. Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychol Bull* 2016 Feb;142(2):198-229. [doi: [10.1037/bul0000025](https://doi.org/10.1037/bul0000025)] [Medline: [26479070](https://pubmed.ncbi.nlm.nih.gov/26479070/)]
29. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. *JMIR Mhealth Uhealth* 2014 May 16;2(2):e24 [FREE Full text] [doi: [10.2196/mhealth.3328](https://doi.org/10.2196/mhealth.3328)] [Medline: [25099752](https://pubmed.ncbi.nlm.nih.gov/25099752/)]
30. U.S. Department of HealthHuman Services. FDA. Mobile medical applications - guidance for industry and food and drug administration staff URL: <http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf> [accessed 2016-10-19] [WebCite Cache ID 6lNUY0ecS]

31. European Commission. Meddev. Guidelines on the qualification and classification of stand alone software used in healthcare within the regulatory framework of medical devices URL: http://meddev.info/wp-content/uploads/2016/08/MEDDEV_2_1_6.pdf [accessed 2016-10-19] [[WebCite Cache ID 6INVY8o6R](#)]

Abbreviations

BMI: body mass index

CAPI: computer-assisted personal interviews

EUROHIS-QOL: European Health Interview Survey-Quality of Life

FDA: Food and Drug Administration

HLS-EU-Q: European Health Literacy Survey Questionnaire

ISCED: International Standard Classification of Education

WHO: World Health Organization

Edited by G Eysenbach; submitted 31.10.16; peer-reviewed by P Krebs, L Lim, P Hanlon; comments to author 01.12.16; revised version received 27.12.16; accepted 26.02.17; published 05.04.17.

Please cite as:

Ernsting C, Dombrowski SU, Oedekoven M, O'Sullivan JL, Kanzler M, Kuhlmei A, Gellert P

Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey

J Med Internet Res 2017;19(4):e101

URL: <http://www.jmir.org/2017/4/e101/>

doi: [10.2196/jmir.6838](https://doi.org/10.2196/jmir.6838)

PMID: [28381394](https://pubmed.ncbi.nlm.nih.gov/28381394/)

©Clemens Ernsting, Stephan U Dombrowski, Monika Oedekoven, Julie L O'Sullivan, Melanie Kanzler, Adelheid Kuhlmei, Paul Gellert. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 05.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Social Health Inequalities and eHealth: A Literature Review With Qualitative Synthesis of Theoretical and Empirical Studies

Karine Latulippe¹, MSc, Ph.D. Candidate; Christine Hamel¹, PhD; Dominique Giroux^{2,3}, PhD

¹Department of Studies of Teaching and Learning, Laval University, Québec, QC, Canada

²Faculté de médecine, Département de réadaptation, Laval University, Québec, QC, Canada

³Centre d'Excellence du Vieillissement de Québec, Chu de Québec, Québec, QC, Canada

Corresponding Author:

Karine Latulippe, MSc, Ph.D. Candidate

Department of Studies of Teaching and Learning

Laval University

2320, rue des Bibliothèques

Québec, QC, G1V 0A6

Canada

Phone: 1 418 457 3542

Email: karine.latulippe.3@ulaval.ca

Abstract

Background: eHealth is developing rapidly and brings with it a promise to reduce social health inequalities (SHIs). Yet, it appears that it also has the potential to increase them.

Objectives: The general objective of this review was to set out how to ensure that eHealth contributes to reducing SHIs rather than exacerbating them. This review has three objectives: (1) identifying characteristics of people at risk of experiencing social inequality in health; (2) determining the possibilities of developing eHealth tools that avoid increasing SHI; and (3) modeling the process of using an eHealth tool by people vulnerable to SHI.

Methods: Following the EPPI approach (Evidence for Policy and Practice of Information of the Institute of Education at the University of London), two databases were searched for the terms SHIs and eHealth and their derivatives in titles and abstracts. Qualitative, quantitative, and mixed articles were included and evaluated. The software NVivo (QSR International) was employed to extract the data and allow for a metasynthesis of the data.

Results: Of the 73 articles retained, 10 were theoretical, 7 were from reviews, and 56 were based on empirical studies. Of the latter, 40 used a quantitative approach, 8 used a qualitative approach, 4 used mixed methods approach, and only 4 were based on participatory research-action approach. The digital divide in eHealth is a serious barrier and contributes greatly to SHI. Ethnicity and low income are the most commonly used characteristics to identify people at risk of SHI. The most promising actions for reducing SHI via eHealth are to aim for universal access to the tool of eHealth, become aware of users' literacy level, create eHealth tools that respect the cultural attributes of future users, and encourage the participation of people at risk of SHI.

Conclusions: eHealth has the potential to widen the gulf between those at risk of SHI and the rest of the population. The widespread expansion of eHealth technologies calls for rigorous consideration of interventions, which are not likely to exacerbate SHI.

(*J Med Internet Res* 2017;19(4):e136) doi:[10.2196/jmir.6731](https://doi.org/10.2196/jmir.6731)

KEYWORDS

Internet; social media; telemedicine; healthcare disparities

Introduction

Background

A number of studies have demonstrated that eHealth is effective in preventing and treating illness for the entire population [1-6]. eHealth is the way to improve health care locally, regionally,

and worldwide by using information and communication technology [7]. At a political level, the American Recovery Reinvestment Act authorized the government to spend US \$38 billion over 10 years to develop eHealth in order to increase accessibility to care [8]. Australia's National eHealth Strategy predicts that eHealth will transform the manner in which

consumers interact with the health care system and will lead to a reduction in costs and demands on the system [9]. According to Health Canada [10], eHealth is an essential element in the renewal of health care and its application to the Canadian system, thanks to improvements in accessibility, and the quality and efficiency of the system; this is beneficial to Canadians. However, eHealth also has the potential to increase social health inequalities (SHIs) [6,11-13]. SHIs, such as the difference in the prevalence of illness and of illness repercussions, the mortality rate, and the burden of illness and other health conditions for specific population groups are caused by unjust and modifiable social factors [14]. This term includes inequalities and inequities in the environment, access, utilization and the quality of services, health status, and the results of interventions [15]. Indeed, eHealth is effective to the extent that individuals are in a position to use it well. Yet, this is not the case for everyone; in fact, this creates a gap between users and nonusers in terms of the improvement of health services.

The reduction of SHI is a key challenge for health systems worldwide, including in Canada, and eHealth is an economical and political means to that end. Yet, since it also has the potential to increase them, it is essential to focus on developing eHealth tools that, in fact, contribute to the reduction of SHI and not their exacerbation. This leads to the following question: how do we ensure that eHealth contributes to reducing social inequalities rather than exacerbating them? In responding to the research question, we have three objectives: (1) identifying characteristics of people at risk of experiencing social inequality in health; (2) determining the possibilities of developing eHealth tools that avoid increasing SHI; and (3) modeling the process of using an eHealth tool by people vulnerable to SHI. To answer this question, a review of the literature is required.

eHealth Tool and SHIs

Previous literature reviews on the relationship between SHIs and eHealth are summarized here to both incorporate already existing knowledge on the subject and to demonstrate the relevance and contribution of this review. Seven reviews on the relationship between SHI and eHealth were identified over the last decade. First, Gibbons et al [16] reviewed some design principles based on solid data to improve the facility with which people at risk of SHI handle eHealth tools. They identified 5 principles to consider when developing an inclusive eHealth tool:

1. Use a design based on experimentation with the tool allowing us to identify the nature of possible errors and the strategies to employ.
2. Create a tool for people with limited resources in order to ensure that all users are readily able to use it.
3. Whenever possible, avoid authentication procedures with the tool (if this aspect is indispensable, considering the personal data that utilization of the tool requires, ensure that technical assistance is provided to users).
4. Minimize the potential of having harmful information inadvertently available.
5. Evaluate the tool with representative users.

For their part, the objective of Dorstyn et al [17] was to synthesize quantitative evidence related to the efficacy of adult

telecounseling for a racial minority. They demonstrated the efficacy of this type of eHealth tool in comparison to monitoring alone, but this has yet to be proven in comparison to face-to-face encounters.

Next, Montague and Perchonok [18], in their review of the literature, examined how technology is used by historically disadvantaged populations to reduce health inequalities. Thus, they addressed four research questions: (1) What types of technologies were used to improve health results of historically disadvantaged populations? They discovered that videos, the Internet (including access via mobile phones), computers, and radios were the most studied technologies. (2) On what health problems is technology focused? The five most studied problems are cancer, diabetes, human immunodeficiency virus (HIV), nutrition, physical activity, and sexually transmitted infections. (3) For which historically disadvantaged groups have eHealth interventions been designed? 19 groups were identified in the literature, including Americans of African origin, Hispanics, indigenous people, and Americans of Asian origin. (4) How were the impacts of the use of such technologies evaluated? Self-evaluation measures are the most common, followed by physiological changes.

For their part, Huxley et al [19] attempted to understand the effects of interventions linked to digital communication in specific contexts (marginalized groups vs the general population). They revealed a number of barriers to the use of communication in general for marginalized groups including, notably, difficulties of access, and stigmatizing reactions from both health professionals and other patients. Nonetheless, digital communication has the potential to reduce these barriers by providing anonymity and offering advantages for those needing an interpreter. This form of communication is liable to function well when there is a preexisting relationship with the practitioner.

For their part, Chou and colleagues [20] explored the evidence concerning the use of Web 2.0 and social media and their impact on the promotion of health. From this, they concluded that the lack of empirical research meant that further investigation was required, especially concerning the design of tools accessible to vulnerable populations.

McInnes et al [21] studied access to and utilization of information technologies among the homeless. They found that use varied from 24% to 84%, depending on the technology (cellular, computer, or access to a public computer), and suggest that this technology could contribute to improving the health of this population.

Finally, Piette et al [22] conducted a scoping review to identify data on the effects of eHealth on health outcomes and costs. They conclude that “Although large programs for eHealth implementation and research are being conducted in many low- and middle-income countries, more information on the impacts of eHealth on outcomes and costs in these settings is still needed.”

Thus, although these reviews of the literature make a major contribution to the body of knowledge on the relationship between SHI and eHealth, they only partially address the

question and research objectives. This review of the literature is intended to complement the reviews cited previously.

Methods

Approach

The EPPI (Evidence for Policy and Practice of Information of the Institute of Education at the University of London) approach was used in this procedure [23]. This approach suggests an iterative process with an explanation and a justification of the choices made. The EPPI approach offers an armory of tools and strategies for conducting research reviews on “how” to use eHealth tools to reduce SHI. The EPPI approach was chosen for its openness to integrating different types of studies and their variety of methodologies. It aims at the understanding of a phenomenon, to which every study, regardless of design, has the potential to contribute [24].

Criteria of Inclusion and Exclusion

In order to respond adequately to the research questions, criteria of inclusion and exclusion were established. The criteria of inclusion were (1) articles published within the last decade (2006-2016); (2) in a peer-reviewed academic journal; (3) in English or French; (4) with an abstract available for screening by title and abstract; and (5) related to the research subject. For this last criterion, it was established that the article must concern eHealth and SHI; eHealth and the populations at risk of SHI (related to poverty, ethnicity, gender, mental health, age, low levels of literacy, HIV, low levels of numeracy, sexual orientation, rural residence, or tobacco addiction); or eHealth in the general population, but demonstrate inequality through a differentiated sociodemographic analysis. The four first criteria of inclusion were applied through research filters available from the databases. In the context of this review, the eHealth tools examined are those concerning education of the entire population or of individuals and do not include technological tools related to the management of the health care system, the monitoring of the health of the entire population, education for professionals, and the exchange of information between organizations.

Articles were excluded if the study focused on health or educational professionals, if the eHealth tool was exclusively a method to collect data for research, or if the article was not available.

Research Strategy

There are countless knowledge transfer platforms related to eHealth and SHI. Nevertheless, to make this review as replicable as possible, it was decided that the references needed to be tracked by database. From April to July 2016, articles were identified from two databases related to the research subject, Medline (PubMed) and Cumulative Index to Nursing and Allied Health Literature (CINAHL). For each database, terms corresponding to key concepts as well as those associated with the thesaurus of each database were used and searched for in titles and abstracts. The terms corresponding to key concepts were identified from leading articles on eHealth and SHI, with the help of a specialist in documentation from the University of Laval.

The first chain of terms related to eHealth included: eHealth, Web-based, Internet, interactive health communication*, health communication*, computer communication network*, computer-assisted therapy, computer assisted, software, communication* media, telecommunication*, multimedia, medical information technolog*, computing, consumer* health information technolog*, World Wide Web, computer-assisted instruction*, interactive technolog* application*, hypermedia*, video game*, Virtual realit*, online learning, social media*, new media*, participatory media*, user-generated content, Facebook, MySpace, Twitter, YouTube, Second Life, wiki*, blog*, Web 2.0, online social network, social networking, health application*+thesaurus: (PubMed) Internet, social media; (CINHAL) and information science with the Boolean operator “OR” between each term.

The second chain of terms stemming from SHIs included underprivileged, health inequalit*, inequalit* in health, poverty, inequalit*, social inequalit*, socioeconomic inequalit*, health for all, health-related exclusion, health disparit*, health equit*, equit*, in health, vulnerable group*, inequalit*, disparit* in health+thesaurus: (PubMed) Health Care Disparities+socioeconomic factors+poverty; (CINAHL) Health Care Disparities+health status disparities+poverty, also with the Boolean operator “OR” between each term. Then, the two chains of concepts were interconnected with the Boolean operator “AND.”

The articles identified were exported to software for bibliographical references (Zotero) to facilitate the classification, importation, and exportation of documents, as well as the removal of duplicates. References were then imported using Covidence [25] (a Cochrane technology platform) to select articles respecting the criteria of inclusion and exclusion, first by titles and abstracts, and then by the complete article. Covidence was specifically designed to support systematic reviews.

Evaluation of the Quality of Articles

To evaluate the quality of quantitative studies, the *Quality Assessment Tool for Quantitative Studies* [26] was used. This tool was judged to be excellent in evaluating the quality of studies in public health [27]. The qualitative studies were themselves evaluated based on an adaptation of quality standards from qualitative studies of Letts et al [28], including Guba and Lincoln [29] and Howe and Eisenhart [30]. There are no generally accepted norms by which to assess the methodological quality of mixed methods [31]. Nonetheless, we chose to use the criteria of Schifferdecker and Reed [32] to produce more precise guidelines. The reviews were evaluated by assessment of multiple systematic reviews (AMSTAR) [33], a valid and reliable instrument for evaluating the methodological quality of systematic reviews [27,34].

In order to meet the objectives of this review of the literature and to ensure that we were not eliminating data relevant to the research, the quality of articles was not evaluated with the goal of excluding articles but rather to consider their limitations in the course of the analysis and synthesis of knowledge [24].

Analysis

The analysis was completed in two stages. First, thematic analysis [35], with the use of Nvivo software (QSR International) allowed for the classification of themes related to the research goals emerging from the articles retained. Thematic analysis permits us to identify all the relevant themes for our research [35]. The data were extracted from all the articles selected and organized, with an inductive approach, by theme, according in conjunction with the objectives of the review. The final themes selected are (1) characteristics of those studied who are at risk of SHI; (2) potential obstacles to the use of eHealth tools; (3) interventions in eHealth that could potentially contribute to the diminution of SHI; (4) eHealth interventions that could potentially contribute to the increase of SHI; and (5) the types of technology. Next, a metasynthesis was performed to enhance understanding of the creation of SHI in eHealth [24]. Metasynthesis serves to comprehend a phenomenon [36]. The analysis, with the help of conceptualizing categories, forms the basis of this metasynthesis [35]. A conceptual map was created with MindMaple Lite (MindMaple Inc) according to the interpretation of the articles in order to model the process of using an eHealth tool by people at risk of experiencing SHI. This map was then designed to facilitate

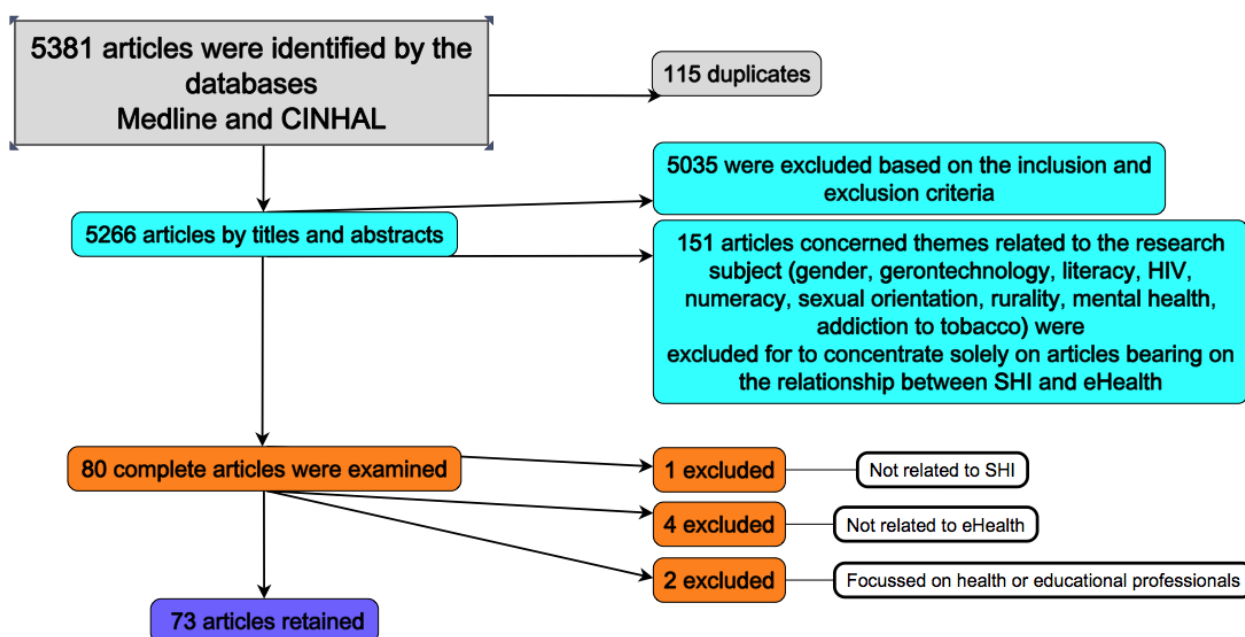
comprehension. Each stage of the process, as well as the resulting choices and justifications, were documented in a logbook and supervised by the director (CH) of the principal author (KL). A general outline of the studies done (descriptive mapping) will first be presented.

Results

Articles Selected

A total of 5381 articles were identified by the databases. Of this number, 115 duplicates were eliminated. Thus, the titles and abstracts of 5266 articles were first examined. It was found that 5035 were excluded based on the inclusion and exclusion criteria. Of the remaining number, 151 articles concerned themes related to the research subject (gender, gerontechnology, literacy, HIV, numeracy, sexual orientation, rurality, mental health, addiction to tobacco). Although these articles could contribute to the exploration of some principal themes, it was decided to concentrate solely on articles bearing on the relationship between SHI and eHealth. Thus, 80 complete articles were examined and, of these 7 were ultimately excluded, bringing the final number to 73 articles retained for this analysis (Figure 1).

Figure 1. PRISMA flow diagram.



Description of Included Studies

Of the 73 articles retained, 10 were theoretical, 7 were reviews of the literature previously referenced, and 56 were empirical studies. Of the latter, 40 adopted a quantitative approach, 8 employed a qualitative approach, 4 used mixed methods, and 4 were based on participatory research action (see [Multimedia Appendix 1](#)). The majority of empirical studies were American (44), and the rest were British (4), Australian (4), Dutch (2), and German (1).

The quality of studies varied ([Multimedia Appendix 2](#)). However, the conceptual categories retained in the metasynthesis all recurred in the articles, which diminished the effect of the weaker studies.

A large proportion of the studies that focused on the utilization of technology in daily life among people at risk of SHI (15) attempt to comprehend how people at risk of SHI seek health information on the Internet (11) or how a specific eHealth tool is used by people at risk of SHI (10). Only three studies examined document the development of an eHealth tool for those in a situation of SHI. The other studies concern the

evaluation of attrition in the use of an eHealth tool, education about the utilization of the technology, acceptance of a technology by this clientele, differences in the type of communication, accessibility, and confidentiality among people at risk of SHI. Another study focused on the development of a measure of SHI in eHealth. Finally, certain studies (3) evaluate the frequency of use and the rate of attrition of a technology for the population as a whole through a differentiated analysis of sociodemographic data (see [Multimedia Appendix 1](#)).

Digital Divide and Social Health Inequalities

Unequal access to the Internet, the primary digital divide, has an effect on the utilization of eHealth [37]. The term “digital divide” first sends us back to the separation between those who have access to technologies such as computers, mobile phones, or the Internet and those who do not have access, often people with low income [6,11-13,38-42].

Although the digital divide still exists, it has diminished every year [43,44] with the use of mobile phones and other mobile devices [45], the reduced cost of technology and the Internet, as well as the spread of places where the Internet is free [40]. However, even though access is a crucial element in the utilization of a technology, this is not sufficient [12]. Thus, certain researchers have determined that knowledge related to the utilization of the Internet also has an impact. This knowledge gap between users is called the secondary digital divide [12,40,44,46-48]. Indeed, it is possible to have the capacity to connect to the Internet, but to lack sufficient knowledge to use it adequately. This highlights the need to develop new users’ skills, along with interventions to increase access [11,12,44].

Today, some authors are identifying other barriers to the utilization of technologies, referred to as the tertiary digital divide. Much more widespread, this tertiary digital divide refers to the concept of significant (or universal) access encompassing equipment, Internet connections, the development of skills, technical assistance, and appropriate content, that is, that health information be comprehensible and useful for disadvantaged populations [37,39-41,47]. In particular, this includes geographical location, literacy, attitudes and behavior with respect to the search for information, confidence and concerns about private life and institutional policies, and content, including the lack of local information, language, incapacities, and the lack of cultural sensitivity [39].

It is important to mention that the digital divide is also influenced by the choices of managers of medical services [49]. Innovations chosen earlier were not necessarily developed with a consideration of people at risk of SHI and may pose problems for these individuals [49]. Managers also have the role of evaluating a potential eHealth tool with respect to its universal (or significant) access.

The digital divide may also be accentuated in the stage before the utilization of an eHealth tool, that is, the search for assistance, information, or services (help-seeking). Indeed, people who have less of a tendency to seek information and to use services are those most at risk of SHI [41,46,47,50]. However, a number of them will still seek information in their local community [51] or when they are particularly interested in a subject [52].

Finally, it appears that the digital divide is more a continuum than a dichotomous concept [53]. The consequences of the digital divide on the health of individuals have been recognized by the United States since the turn of the century. It is now a matter of justice in health [42], since the digital divide in eHealth is a significant barrier that serves to accentuate SHI [49].

Characteristics of People at Risk of Experiencing a Situation of Social Inequality in Health

SHI and the digital divide generally affect the same individuals [37,54]. eHealth tools are primarily developed for people with good digital skills and Internet access [39,49]. Meanwhile, most nonusers of the Internet are older people or those with low income. Thus, inequalities are accentuated for these groups [42]. An effective design of an eHealth tool for one group could bring about negative and unforeseen consequences for another group with different characteristics (physical, cognitive, or cultural) [16]. Ethnicity (48) and low income (47) are the most common characteristics. Next comes a low level of education (34), age (26), a low literacy level (18), gender (14), rurality (11), incapacities (8), psychological distress (1), homelessness (1), and sexual orientation (1). According to Feng’s study [44], groups identified as particularly disadvantaged in the utilization of social networks are low-income individuals, those with little education or literacy problems, the unemployed, the aged, the handicapped, women, and the people of ethnic origin. However, since Feng [44] used the correction of Bonferroni in his analysis and, thus, chose a more conservative stance, it is possible that certain links were not brought to light (see [Table 1](#)).

Table 1. Characteristics of people at risk of experiencing a situation of social inequality in health.

This review (n)	Feng [44]
Ethnicity (48)	Ethnicity
Low income (47)	Low income
Low level of education (34)	Low level of education
Age (26)	Age
Low literacy level (18)	Low literacy level
Gender (14)	Gender
Rurality (11)	
Incapacities (8)	Incapacities
Psychological distress (1)	
Homelessness (1)	
Sexual orientation (1)	

In Canada, it appears that income is the factor with the greatest impact on Internet access, more than other factors such as education level, geographic location, gender, and age [55]. However, research on racial and ethnic health inequalities has demonstrated that SHI may persist despite the inclusion of measures related to socioeconomic status [56]. Thus, people with average family income could be at risk of SHI if they belong to another ethnic community.

Although older individuals are the group for whom the use of the Internet is growing most rapidly, this is still a group that uses it the least [57]. Also, within this group, certain disparities exist. Seniors from a minority ethnic group, with little education and literacy, aged 75 years older, or with low income are much less likely to use the Internet [57]. The presence of a number of cognitive and psychomotor barriers related to age may make it difficult to use digital technology, and the effort required to master a new technology can then be perceived as greater than any possible benefits [57]. In general, for older people, and even more for those with low income and incapacities, literacy and Internet access are important factors in explaining the digital divide [58]. Finally, in terms of gender, although women have a tendency to use technology less, they still use eHealth more [59,60].

What can be done to lessen the digital divide (primary, secondary, and tertiary)? Four promising strategies for the development of the eHealth tool are highlighted in the analysis of different studies examined in the context of this review.

Promising Strategies for Development of the eHealth Tool to Reduce Social Health Inequalities

Ensuring Universal Access to the eHealth Tool

To guarantee universal access and reduce the digital divide, it is important to clearly understand the systemic barriers, which potential users may confront [42]. An approach centered on the user is recommended, placing the person's needs, preferences, capacities, values, and goals in the forefront, in particular, when this concerns people at risk of SHI.

To reduce problems of physical access to a computer, the strategies generally proposed are increasing the quantity of

computers available in public spaces such as libraries and community centers or providing more personal computers in people's homes [47,58]. The former solution has the advantage of being more economical. However, it may also constitute a violation of confidentiality, particularly when content related to health is involved [47]. Even if computers are available in libraries, problems with transportation may limit this solution. As for personal computers, Ryan [61] illustrates some very concrete difficulties related to the utilization of an eHealth tool at home. For example, some participants have transportation problems and were unable to get their equipment repaired by the provider who did not make home visits. Another participant burned the motherboard of the personal computer (PC) because the latter was plugged into the same electrical outlet as a kitchen appliance; there were a limited number of outlets in the house. Other participants could not pay for sufficient bandwidth to use the tool. Thus, the importance of access to quality Internet bandwidth should not be underestimated. eHealth tools employ modalities that require a certain performance (in terms of graphics, software, and interfaces), including the downloading of documents for later use or the participation in forums to share with peers [40]. It has also been suggested that hospitals and health services offer free Internet access to their patients [55]. The utilization of mobile devices (tablets and mobile phones) is higher among people at risk of experiencing SHI [18]. This type of technology needs more research, but is still promising in terms of increasing access to eHealth [45,62,63]. Furthermore, a combination of online and offline tools may prove necessary, along with more traditional technologies such as the telephone, printed material, digital versatile disks (DVDs), and printed mail [49,53,57,64].

However, it is not sufficient to provide a tool. It appears that it is necessary to ensure that the future user has the knowledge required for an optimal utilization of the tool on offer. Thus, training and technical assistance are crucial, according to the authors [39,47,57,58,64,65]. In that regard, it is possible to create a support network to bolster users' skills. For example, volunteers could help older people to learn to use the Internet [40]. Beyond the usage of an eHealth tool, it is also crucial that users be trained to be able to evaluate the quality of sources on

the Internet [66]. However, Chu's study [66] suggests that the attrition rate for this type of training is problematic. Motivation then becomes a critical factor [67].

Respecting Users' Level of Literacy

In designing the tool, the patient's literacy level and principal language, as well as access and facility of use, must be considered [68]. Thus, audio accompaniment, available in a variety of languages, could compensate for literacy difficulties [66,68]. Certain authors have tested the utilization of modalities of providing health information that demand less in terms of reading skills. They suggest more educational entertainment, using animation or multimedia narrative tools (television programs, video capsules, and so on) [41,47,66].

Creating eHealth Tools That Respect the Cultural Characteristics of Future Users

Bacigalupe [49] and McAuley [12] stress that the cultural component in the development of eHealth tools is critical for populations at risk of SHI and, thus, they suggest using targeted strategies (tools specifically designed for these populations), rather than universal strategies (intended for everyone). A failure to consider beliefs, values, and habits of populations or individuals targeted can lessen the value of the tool developed for these individuals [47]. The utilization of photographs representing populations at risk of SHI and a variety of testimonies, the availability of the tool in a number of languages, and focusing on specific needs of this clientele are concrete examples of strategies favoring the consideration of the cultural dimension in eHealth [49,50,63,69].

Inviting the Participation of People at Risk of SHI in Developing eHealth Technologies

The active participation of future users and, in particular, people at risk of SHI, in the development of eHealth tools has the potential to reduce inequalities [49]. Involving future users with diverse perspectives, circumstances, capacities, and experiences in the design process increases the chances that the tool will ensure significant (universal) access [42]. Future users have the skills to evaluate, choose, and use eHealth tools and to gain from the experience [42]. Nonetheless, the involvement of low income or poorly educated people, various ethnic groups, as well as those with low literacy levels, still requires specific abilities on the part of the designer to encourage their active participation in designing an eHealth tool [42].

Discussion

Principal Findings

This review of the literature had three objectives: (1) identifying characteristics of people at risk of experiencing a situation of SHI; (2) determining the possibilities for action in the development of eHealth tools that avoid increasing SHI; and (3) modeling the process of using an eHealth tool by people at risk of experiencing a situation of SHI.

For the first objective, we saw that a number of sociodemographic characteristics were brought up in various studies to identify or characterize individuals at risk of SHIs (ethnicity, low income, low level of education, age, low literacy

level, gender, rurality, incapacities, psychological distress, homelessness, and sexual orientation). Now, these characteristics should be analyzed with due caution. On one hand, they could contribute to supporting a discourse based on differences, but they also fail to consider the heterogeneity that one finds within a single population group [70]. Thus, it seems essential to ensure a range of characteristics when recruiting participants for studies on SHIs and eHealth.

For the second and third objectives, the results obtained from this review of the literature show that the digital divide, in its primary, secondary, and tertiary forms is the principal cause of the exacerbation of SHI by eHealth and that it affects those people already at risk of SHI [42,47,49]. Alternative ways of modeling the link between eHealth and SHI exist. Among others, the integrative model of eHealth use suggests that disparities in social structures (eg, the demographic data) are linked to SHI through health literacy, motivation to use eHealth, and the person's capacity to use this technology. In this model, existing SHI are exacerbated by technologies that require a certain level of literacy, sustained motivation, and digital capacities [41]. Also, the Structural Influence model identifies the importance of communication in the relationship between social determinants and results linked to health [40,71,72]. It suggests that the differences among social groups (including ethnic minorities) in the utilization of channels of communication result in an exacerbation of SHI [72]. These are highly interesting models. However, the goal of these models is not the development of eHealth tools, and certain key elements, such as the cultural component and the importance of involving future users, are absent. Thus, descriptive metasynthesis allows us to respond to the second objective. Individuals characteristics linked to SHI will encounter difficulties during the process of using an eHealth tool. First, it is possible that they will be less inclined to seek health-related information or to use an eHealth tool to improve their health [41,46,50]. In the case where these people do initiate a process of looking for help, they will need physical access to digital technology (a computer, electronic tablet, or mobile phone) and sufficient bandwidth [66,61,73]. Then, they need to draw upon their capacities to use the technology. Probably they will lack confidence in their abilities or in the technology and will interrupt the process [11,12,73]. However, if they persevere, they will require a level of health literacy sufficient to understand what the eHealth tool is able to offer them and a capacity to integrate and make use of what has been learned [40,73,74]. Individuals with sufficient income, a high level of education, and adequate digital health literacy will be better able to complete the process and improve their state of health. Thus, it is possible that there are gaps between these groups of individuals in the effective utilization of eHealth tools and, therefore, in the improvement of their health, which will contribute to increasing SHI [72].

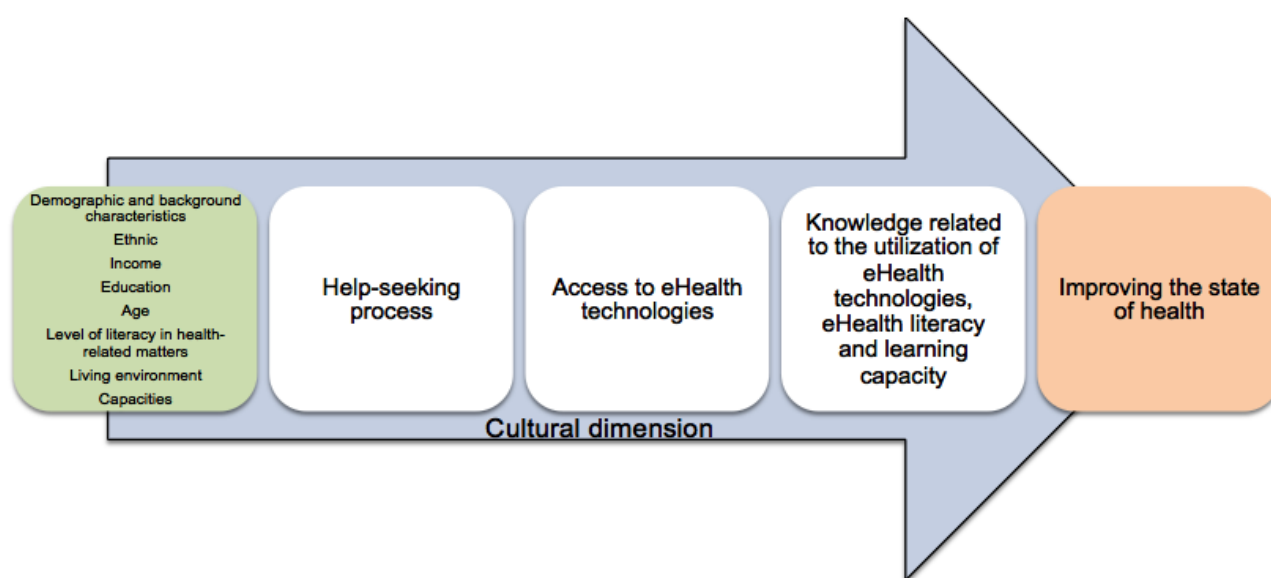
Nonetheless, if in the designing or adaptation of the tool, the developers consider the future user as a person at risk of SHI [49], design or adapt the tool to respond to the needs of such a user at each stage [42], and integrate the cultural dimension in the process of development [49], it might be possible to reduce the digital divide present in eHealth (Figure 2).

The current increase of technologies in eHealth justifies a reexamination of interventions unlikely to worsen SHI [42,47]. Among other suggestions, it is proposed to target interventions for populations at risk of SHI. Yet, developing an eHealth tool is an undertaking requiring time, energy, and funds. Realistically, developers hope to reach the greatest possible number, and targeted interventions are likely to be rarer. Little participatory research action has been done despite the promising nature of participation of people at risk of SHI in developing eHealth tools to reduce these inequalities. Can we consider developing eHealth tools with the end goal of a universal strategy, but designed to take into account people at risk of SHI and even to involve them in the process? Could we, in developing the tool, question ourselves and question the people at risk of SHI at each stage of the process of using an

eHealth tool (Figure 2) and reduce the barriers liable to interrupt the process? Each stage of the process (Figure 2) or conceptualizing category refers to its own field of research. It is difficult, indeed impossible, in the context of this article, to showcase the wealth of knowledge available for each of these concepts. However, the relation between these concepts, more iterative than linear, allows us to envisage a process of coherent codesign, the effect of which might be to reduce SHI.

Although research often raises the potential of eHealth to reduce SHI and offers promising solutions for reducing the digital divide, we agree with Chou [20] that, to date, there are still insufficient empirical studies to prove this definitively, as demonstrated in this review of the literature. Indeed, only three studies examined document the development of an eHealth tool with individuals in a situation of SHI.

Figure 2. Process of using an eHealth tool.



Limitations

The strengths and weaknesses of this study were assessed with the help of AMSTAR [27,33,34]. Although we have attempted to provide a rigorous review of the literature, including a metasynthesis, this review has its limitations. First, in concentrating on a research strategy supported by databases, the gray literature and nontraditional sites of knowledge transfer (eg, the Web) were not included. In addition, we have only used two databases. A limited search can generate a set of studies that are not representative, incomplete results, inadequate selection, and reduced generalization [23]. We have prioritized this choice to ensure greater transparency and reproducibility for this review of the literature. AMSTAR mentions that it is necessary to utilize least two different databases [33]. To avoid biases in the publications, it is recommended not to exclude articles on the basis of year of publication or language. Now, considering that the Internet, social networks, and new technologies have considerably modified the eHealth environment, it was judged sufficient to focus on articles published in the last decade. Furthermore, for reasons of feasibility, the translation of articles was not possible, and free

translation software still leaves much to be desired. In accordance with the suggestion of the EPPI group, it was decided to look for articles in all languages initially but, for greater transparency, to exclude articles that are not in languages in which we are fluent [23]. Another limitation of this study is the presence of only a single analyst, which could trigger selection bias. To counter this aspect, often linked to student reality, two supervisors provided support for the writing of this text, and a biostatistician examined the articles from a quantitative perspective. Finally, since the analysis was not based on the quality of studies, the results must be interpreted as possibilities, rather than generalizable facts based on solid data. The rigor of this review stems from the fact that it is systematic (undertaken according to a fixed plan or system or method) and that it is explicit and justified [23]. Nonetheless, because this review does not adopt the same high standards in terms of protection against bias and the quality assessment for the selection of primary research" [75], we called it a "literature review" and not a "systematic review" [75].

Conclusions

The synthesis of knowledge allowed for (1) a modeling of the process of using an eHealth tool, (2) identifying the actions in eHealth that do not help to reduce SHI, but (3) determining the possibilities for action in the development of tools of eHealth that avoid increasing SHI as well. The massive expansion of

technologies in eHealth justifies the study of interventions less likely to exacerbate SHI through the usage of eHealth, and few current empirical studies reveal concrete and effective solutions. Furthermore, very few studies involve future users at risk of SHI. Research is still necessary for eHealth to fulfill its promise to reduce SHI.

Acknowledgments

The author is extremely grateful to Dr Christine Hamel, director, and Dr Dominique Giroux, codirector of her PhD project; the graduate student learning and practice community (CAPES) of the Faculty of Education at Laval University for the coconstruction of knowledge related to the production of a systematic literature review; as well as Mr Pierre-Hugues Carmichael, biostatistician at the Centre d'Excellence sur le vieillissement de Québec for his support. A special thanks also to AGE-WELL, which supported the author (KL) financially in her thesis project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Included articles.

[PDF File (Adobe PDF File), 246KB - [jmir_v19i4e136_app1.pdf](#)]

Multimedia Appendix 2

Quality of included articles.

[PDF File (Adobe PDF File), 194KB - [jmir_v19i4e136_app2.pdf](#)]

References

1. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. *J Med Internet Res* 2013;15(7):e146 [FREE Full text] [doi: [10.2196/jmir.2665](#)] [Medline: [23859884](#)]
2. Ammann R, Vandelanotte C, de Vries H, Mummery WK. Can a website-delivered computer-tailored physical activity intervention be acceptable, usable, and effective for older people? *Health Educ Behav* 2013 Apr;40(2):160-170. [doi: [10.1177/1090198112461791](#)] [Medline: [23077157](#)]
3. Hamel LM, Robbins LB. Computer- and web-based interventions to promote healthy eating among children and adolescents: a systematic review. *J Adv Nurs* 2013 Jan;69(1):16-30. [doi: [10.1111/j.1365-2648.2012.06086.x](#)] [Medline: [22757605](#)]
4. McClure JB, Shortreed SM, Bogart A, Derry H, Riggs K, St JJ, et al. The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial. *J Med Internet Res* 2013;15(3):e69 [FREE Full text] [doi: [10.2196/jmir.2508](#)] [Medline: [23529377](#)]
5. de Josselin de Jong S, Candel M, Segaar D, Cremers H, de Vries H. Efficacy of a Web-based computer-tailored smoking prevention intervention for Dutch adolescents: randomized controlled trial. *J Med Internet Res* 2014;16(3):e82 [FREE Full text] [doi: [10.2196/jmir.2469](#)] [Medline: [24657434](#)]
6. Reinwand DA, Schulz DN, Crutzen R, Kremers SP, de VH. Who follows eHealth interventions as recommended? a study of participants' personal characteristics from the experimental arm of a randomized controlled trial. *J Med Internet Res* 2015;17(5):e115 [FREE Full text] [doi: [10.2196/jmir.3932](#)] [Medline: [25963607](#)]
7. Eysenbach G. What is e-Health? *J Med Internet Res* 2001 Jun;3(2):E20. [doi: [10.2196/jmir.3.2.e20](#)] [Medline: [11720962](#)]
8. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. *Health Informatics J* 2011 Sep;17(3):224-243. [doi: [10.1177/1460458211414843](#)] [Medline: [21937464](#)]
9. Australian Health Ministers Conference. Health. National e-Health strategy: summary URL: [http://www.health.gov.au/internet/main/publishing.nsf/content/69B9E01747B836DCCA257BF0001DC5CC/\\$File/Summary%20National%20E-Health%20Strategy%20final.pdf](http://www.health.gov.au/internet/main/publishing.nsf/content/69B9E01747B836DCCA257BF0001DC5CC/$File/Summary%20National%20E-Health%20Strategy%20final.pdf) [accessed 2017-04-02] [WebCite Cache ID 6kwTTHZps]
10. Santé Canada. Cybersanté URL: <http://www.hc-sc.gc.ca/hcs-sss/ehealth-esante/index-fra.php> [accessed 2016-10-01] [WebCite Cache ID 6kwTcstqW]
11. Viswanath K, McCloud R, Minsky S, Puleo E, Kontos E, Bigman-Galimore C, et al. Internet use, browsing, and the urban poor: implications for cancer control. *J Natl Cancer Inst Monogr* 2013 Dec;2013(47):199-205 [FREE Full text] [doi: [10.1093/jncimonographs/igt029](#)] [Medline: [24395992](#)]

12. McAuley A. Digital health interventions: widening access or widening inequalities? Public Health 2014 Dec;128(12):1118-1120. [doi: [10.1016/j.puhe.2014.10.008](https://doi.org/10.1016/j.puhe.2014.10.008)] [Medline: [25458115](#)]
13. 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](https://doi.org/10.2196/jmir.3117)] [Medline: [25048379](#)]
14. Hyppolite S. Ciusss Capitale Nationale. 2012. Comprendre et agir autrement pour viser l'équité en santé dans la région de la Capitale-Nationale, Rapport du directeur régional de santé publique sur les inégalités sociales de santé URL: http://www.ciusss-capitenationale.gouv.qc.ca/sites/default/files/rapportiss_versionintegrale.pdf [accessed 2017-04-02] [WebCite Cache ID 6pQLhpf3F]
15. National Institutes of Health. NIMHD. 2002. Strategic research plan and budget to reduce and ultimately eliminate health disparities volume I fiscal years 2002-2006 URL: http://www.nimhd.nih.gov/docs/2002_2006_vol1_031003ed_rev.pdf [accessed 2016-10-01] [WebCite Cache ID 6kwTmWRC8]
16. Gibbons MC, Lowry SZ, Patterson ES. Applying human factors principles to mitigate usability issues related to embedded assumptions in health information technology design. JMIR Hum Factors 2014;1(1):e3 [FREE Full text] [doi: [10.2196/humanfactors.3524](https://doi.org/10.2196/humanfactors.3524)] [Medline: [27025349](#)]
17. Dorstyn DS, Saniotis A, Sobhanian F. A systematic review of telecounseling and its effectiveness in managing depression amongst minority ethnic communities. J Telemed Telecare 2013 Sep;19(6):338-346. [doi: [10.1177/1357633X13501767](https://doi.org/10.1177/1357633X13501767)] [Medline: [24163298](#)]
18. Montague E, Perchonok J. Health and wellness technology use by historically underserved health consumers: systematic review. J Med Internet Res 2012;14(3):e78 [FREE Full text] [doi: [10.2196/jmir.2095](https://doi.org/10.2196/jmir.2095)] [Medline: [22652979](#)]
19. Huxley CJ, Atherton H, Watkins JA, Griffiths F. Digital communication between clinician and patient and the impact on marginalised groups: a realist review in general practice. Br J Gen Pract 2015 Dec;65(641):e813-e821. [doi: [10.3399/bjgp15X687853](https://doi.org/10.3399/bjgp15X687853)] [Medline: [26622034](#)]
20. Chou WS, Prestin A, Lyons C, Wen K. 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](https://doi.org/10.2105/AJPH.2012.301071)] [Medline: [23153164](#)]
21. McInnes DK, Li AE, Hogan TP. Opportunities for engaging low-income, vulnerable populations in health care: a systematic review of homeless persons' access to and use of information technologies. Am J Public Health 2013 Dec;103 Suppl 2:e11-e24. [doi: [10.2105/AJPH.2013.301623](https://doi.org/10.2105/AJPH.2013.301623)] [Medline: [24148036](#)]
22. Piette JD, Lun KC, Moura LA, Fraser HS, Michael PN, Powell J, et al. Impacts of e-health on the outcomes of care in low- and middle-income countries: where do we go from here? Bull World Health Organ 2012 May 1;90(5):365-372 [FREE Full text] [doi: [10.2471/BLT.11.099069](https://doi.org/10.2471/BLT.11.099069)] [Medline: [22589570](#)]
23. Gough D, Oliver S, Thomas J. An Introduction to Systematic Reviews. London: Sage Publications Ltd; 2012.
24. Beaucher V, Jutras F. Recherche Qualitative. 2007. Étude comparative de la métasynthèse et de la méta-analyse qualitative URL: [http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero27\(2\)/beaucher27\(2\).pdf](http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero27(2)/beaucher27(2).pdf) [accessed 2017-04-02] [WebCite Cache ID 6pQM35CPK]
25. Covidence. URL: <https://www.covidence.org/> [accessed 2017-04-02] [WebCite Cache ID 6pQMAP6Es]
26. Thomas BH, Ciliska D, Dobbins M, Micucci S. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. Worldviews Evid Based Nurs 2004;1(3):176-184. [doi: [10.1111/j.1524-475X.2004.04006.x](https://doi.org/10.1111/j.1524-475X.2004.04006.x)] [Medline: [17163895](#)]
27. NCCDH. 2012. Évaluation de l'incidence et de l'efficacité de l'action intersectorielle exercée sur les déterminants sociaux de la santé et l'équité en santé: une revue systématique accélérée URL: http://nccdh.ca/images/uploads/ISA_Report_FR1.pdf [accessed 2016-10-01] [WebCite Cache ID 6kwUQdg0k]
28. Letts L, Wilkins S, Law M, Stewart D, Bosch J, Westmorland M. McMaster. Quality assessment tool for qualitative studies URL: <http://srs-mcmaster.ca/wp-content/uploads/2015/05/Guidelines-for-Critical-Review-Form-Qualitative-Studies.pdf> [accessed 2016-10-01] [WebCite Cache ID 6kwUcKzaB]
29. Guba E, Lincoln Y. Fourth Generation Evaluation. Newbury Park, CA: Sage Publications; 1989.
30. Howe K, Eisenhart M. Standards for qualitative (and quantitative) research: a prolegomenon. Educational Researcher 1990 May 01;19(4):2-9. [doi: [10.3102/0013189X019004002](https://doi.org/10.3102/0013189X019004002)]
31. Pluye P, Nadeau L, Gagnon MP, Grad R, Johnson-Lafleur J, Griffiths F. Les méthodes mixtes. In: Approches et pratiques en évaluation de programme (French Edition). Montréal: Presses de l'Université de Montréal; 2009.
32. Schifferdecker KE, Reed VA. Using mixed methods research in medical education: basic guidelines for researchers. Med Educ 2009 Jul;43(7):637-644. [doi: [10.1111/j.1365-2923.2009.03386.x](https://doi.org/10.1111/j.1365-2923.2009.03386.x)] [Medline: [19573186](#)]
33. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10 [FREE Full text] [doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10)] [Medline: [17302989](#)]
34. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009 Oct;62(10):1013-1020. [doi: [10.1016/j.jclinepi.2008.10.009](https://doi.org/10.1016/j.jclinepi.2008.10.009)] [Medline: [19230606](#)]
35. Paillé P, Mucchielli A. L'analyse qualitative en sciences humaines et sociales. Paris: A. Colin; 2012:1-423.

36. Jensen LA, Allen MN. Meta-synthesis of qualitative findings. *Qual Health Res* 1996 Nov 01;6(4):553-560. [doi: [10.1177/104973239600600407](https://doi.org/10.1177/104973239600600407)]
37. Lindsay S, Bellaby P, Smith S, Baker R. Enabling healthy choices: is ICT the highway to health improvement? *Health (London)* 2008 Jul;12(3):313-331. [doi: [10.1177/1363459308090051](https://doi.org/10.1177/1363459308090051)] [Medline: [18579630](https://pubmed.ncbi.nlm.nih.gov/18579630/)]
38. Neuenschwander LM, Abbott A, Mobley AR. Comparison of a web-based vs in-person nutrition education program for low-income adults. *J Acad Nutr Diet* 2013 Jan;113(1):120-126. [doi: [10.1016/j.jand.2012.07.034](https://doi.org/10.1016/j.jand.2012.07.034)] [Medline: [23092741](https://pubmed.ncbi.nlm.nih.gov/23092741/)]
39. Morey OT. Digital disparities. *J Consum Health Internet* 2007 Dec;11(4):23-41. [doi: [10.1300/J381v11n04_03](https://doi.org/10.1300/J381v11n04_03)]
40. Kontos EZ, Bennett GG, Viswanath K. Barriers and facilitators to home computer and internet use among urban novice computer users of low socioeconomic position. *J Med Internet Res* 2007;9(4):e31 [FREE Full text] [doi: [10.2196/jmir.9.4.e31](https://doi.org/10.2196/jmir.9.4.e31)] [Medline: [17951215](https://pubmed.ncbi.nlm.nih.gov/17951215/)]
41. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: eHealth literacy, health disparities, and the digital divide. *Health Mark Q* 2008 Jan;25(1-2):175-203. [doi: [10.1080/07359680802126301](https://doi.org/10.1080/07359680802126301)] [Medline: [18935884](https://pubmed.ncbi.nlm.nih.gov/18935884/)]
42. Baur C. An analysis of factors underlying e-health disparities. *Camb Q Healthc Ethics* 2008;17(4):417-428. [doi: [10.1017/S0963180108080547](https://doi.org/10.1017/S0963180108080547)] [Medline: [18724881](https://pubmed.ncbi.nlm.nih.gov/18724881/)]
43. Hoffmann R, Eikemo TA, Kulhánová I, Dahl E, Deboosere P, Dzúrová D, et al. The potential impact of a social redistribution of specific risk factors on socioeconomic inequalities in mortality: illustration of a method based on population attributable fractions. *J Epidemiol Community Health* 2013 Jan;67(1):56-62. [doi: [10.1136/jech-2011-200886](https://doi.org/10.1136/jech-2011-200886)] [Medline: [22760220](https://pubmed.ncbi.nlm.nih.gov/22760220/)]
44. Feng Y, Xie W. Digital divide 2.0: the role of social networking sites in seeking health information online from a longitudinal perspective. *J Health Commun* 2015;20(1):60-68. [doi: [10.1080/10810730.2014.906522](https://doi.org/10.1080/10810730.2014.906522)] [Medline: [25119019](https://pubmed.ncbi.nlm.nih.gov/25119019/)]
45. Breitenstein SM, Gross D. Web-based delivery of a preventive parent training intervention: a feasibility study. *J Child Adolesc Psychiatr Nurs* 2013 May;26(2):149-157. [doi: [10.1111/jcap.12031](https://doi.org/10.1111/jcap.12031)] [Medline: [23607827](https://pubmed.ncbi.nlm.nih.gov/23607827/)]
46. Bell J. Harmful or helpful? the role of the internet in self-harming and suicidal behaviour in young people. *Mental Health Review Journal* 2014 Mar 05;19(1):61-71. [doi: [10.1108/MHRJ-05-2013-0019](https://doi.org/10.1108/MHRJ-05-2013-0019)]
47. Beacom AM, Newman SJ. Communicating health information to disadvantaged populations. *Fam Community Health* 2010;33(2):152-162. [doi: [10.1097/FCH.0b013e3181d59344](https://doi.org/10.1097/FCH.0b013e3181d59344)] [Medline: [20216358](https://pubmed.ncbi.nlm.nih.gov/20216358/)]
48. Zhao S. Parental education and children's online health information seeking: beyond the digital divide debate. *Soc Sci Med* 2009 Nov;69(10):1501-1505. [doi: [10.1016/j.socscimed.2009.08.039](https://doi.org/10.1016/j.socscimed.2009.08.039)] [Medline: [19765874](https://pubmed.ncbi.nlm.nih.gov/19765874/)]
49. Bacigalupe G, Askari SF. E-Health innovations, collaboration, and healthcare disparities: developing criteria for culturally competent evaluation. *Fam Syst Health* 2013 Sep;31(3):248-263. [doi: [10.1037/a0033386](https://doi.org/10.1037/a0033386)] [Medline: [24059273](https://pubmed.ncbi.nlm.nih.gov/24059273/)]
50. Bhandari N, Shi Y, Jung K. Seeking health information online: does limited healthcare access matter? *J Am Med Inform Assoc* 2014;21(6):1113-1117 [FREE Full text] [doi: [10.1136/amiajnl-2013-002350](https://doi.org/10.1136/amiajnl-2013-002350)] [Medline: [24948558](https://pubmed.ncbi.nlm.nih.gov/24948558/)]
51. Zach L, Dalrymple PW, Rogers ML, Williver-Farr H. Assessing internet access and use in a medically underserved population: implications for providing enhanced health information services. *Health Info Libr J* 2012 Mar;29(1):61-71. [doi: [10.1111/j.1471-1842.2011.00971.x](https://doi.org/10.1111/j.1471-1842.2011.00971.x)] [Medline: [22335290](https://pubmed.ncbi.nlm.nih.gov/22335290/)]
52. Lee C, Ramírez AS, Lewis N, Gray SW, Hornik RC. Looking beyond the Internet: examining socioeconomic inequalities in cancer information seeking among cancer patients. *Health Commun* 2012;27(8):806-817 [FREE Full text] [doi: [10.1080/10410236.2011.647621](https://doi.org/10.1080/10410236.2011.647621)] [Medline: [22356137](https://pubmed.ncbi.nlm.nih.gov/22356137/)]
53. Massey PM. Where do U.S. adults who do not use the internet get health information? examining digital health information disparities from 2008 to 2013. *J Health Commun* 2016;21(1):118-124. [doi: [10.1080/10810730.2015.1058444](https://doi.org/10.1080/10810730.2015.1058444)] [Medline: [26166484](https://pubmed.ncbi.nlm.nih.gov/26166484/)]
54. Carlson BA, Neal D, Magwood G, Jenkins C, King MG, Hossler CL. A community-based participatory health information needs assessment to help eliminate diabetes information disparities. *Health Promot Pract* 2006 Jul;7(3 Suppl):213S-222S. [doi: [10.1177/1524839906288694](https://doi.org/10.1177/1524839906288694)] [Medline: [16760247](https://pubmed.ncbi.nlm.nih.gov/16760247/)]
55. Gilmour JA. Reducing disparities in the access and use of Internet health information. a discussion paper. *Int J Nurs Stud* 2007 Sep;44(7):1270-1278. [doi: [10.1016/j.ijnurstu.2006.05.007](https://doi.org/10.1016/j.ijnurstu.2006.05.007)] [Medline: [16828775](https://pubmed.ncbi.nlm.nih.gov/16828775/)]
56. Anthony DL, Campos-Castillo C. Do health care users think electronic health records are important for themselves and their providers?: exploring group differences in a national survey. *AMIA Annu Symp Proc* 2013;2013:42-49 [FREE Full text] [Medline: [24551321](https://pubmed.ncbi.nlm.nih.gov/24551321/)]
57. Gordon M, Henderson R, Holmes JH, Wolters MK, Bennett IM, SPIRIT (Stress in Pregnancy: Improving Results with Interactive Technology) Group. Participatory design of ehealth solutions for women from vulnerable populations with perinatal depression. *J Am Med Inform Assoc* 2016 Jan;23(1):105-109. [doi: [10.1093/jamia/ocv109](https://doi.org/10.1093/jamia/ocv109)] [Medline: [26342219](https://pubmed.ncbi.nlm.nih.gov/26342219/)]
58. 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](https://doi.org/10.2196/jmir.2645)] [Medline: [23639979](https://pubmed.ncbi.nlm.nih.gov/23639979/)]
59. Nölke L, Mensing M, Krämer A, Hornberg C. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. *BMC Public Health* 2015;15:31 [FREE Full text] [doi: [10.1186/s12889-015-1423-0](https://doi.org/10.1186/s12889-015-1423-0)] [Medline: [25631456](https://pubmed.ncbi.nlm.nih.gov/25631456/)]
60. Miller EA, West DM. Where's the revolution? Digital technology and health care in the internet age. *J Health Polit Policy Law* 2009 Apr;34(2):261-284. [doi: [10.1215/03616878-2008-046](https://doi.org/10.1215/03616878-2008-046)] [Medline: [19276318](https://pubmed.ncbi.nlm.nih.gov/19276318/)]

61. Ryan JG, Schwartz R, Jennings T, Fedders M, Vittoria I. Feasibility of an internet-based intervention for improving diabetes outcomes among low-income patients with a high risk for poor diabetes outcomes followed in a community clinic. *Diabetes Educ* 2013;39(3):365-375. [doi: [10.1177/0145721713484594](https://doi.org/10.1177/0145721713484594)] [Medline: [23610181](https://pubmed.ncbi.nlm.nih.gov/23610181/)]
62. Bennett GG. Connecting eHealth with 2-1-1 to reduce health disparities. *Am J Prev Med* 2012 Dec;43(6 Suppl 5):S509-S511. [doi: [10.1016/j.amepre.2012.09.022](https://doi.org/10.1016/j.amepre.2012.09.022)] [Medline: [23157773](https://pubmed.ncbi.nlm.nih.gov/23157773/)]
63. Chilukuri N, West M, Henderson JL, Lawson S, Ehsanipoor R, Costigan K, et al. Information and communication technology use among low-income pregnant and postpartum women by race and ethnicity: a cross-sectional study. *J Med Internet Res* 2015;17(7):e163 [FREE Full text] [doi: [10.2196/jmir.3916](https://doi.org/10.2196/jmir.3916)] [Medline: [26142162](https://pubmed.ncbi.nlm.nih.gov/26142162/)]
64. Paul CL, Clinton-McHarg T, Lynagh M, Sanson-Fisher RW, Tzelepis F. On-line information and support for supporters and carers of haematological cancer patients: is access an issue? *Support Care Cancer* 2012 Nov;20(11):2687-2695 [FREE Full text] [doi: [10.1007/s00520-012-1388-9](https://doi.org/10.1007/s00520-012-1388-9)] [Medline: [22350593](https://pubmed.ncbi.nlm.nih.gov/22350593/)]
65. Atkinson NL, Desmond SM, Saperstein SL, Billing AS, Gold RS, Tournas-Hardt A. Assets, challenges, and the potential of technology for nutrition education in rural communities. *J Nutr Educ Behav* 2010;42(6):410-416. [doi: [10.1016/j.jneb.2009.09.004](https://doi.org/10.1016/j.jneb.2009.09.004)] [Medline: [20813589](https://pubmed.ncbi.nlm.nih.gov/20813589/)]
66. Chu A, Huber J, Mastel-Smith B, Cesario S. "Partnering with Seniors for Better Health": computer use and Internet health information retrieval among older adults in a low socioeconomic community. *J Med Libr Assoc* 2009 Jan;97(1):12-20 [FREE Full text] [doi: [10.3163/1536-5050.97.1.003](https://doi.org/10.3163/1536-5050.97.1.003)] [Medline: [19159002](https://pubmed.ncbi.nlm.nih.gov/19159002/)]
67. Geraghty AW, Torres LD, Leykin Y, Pérez-Stable EJ, Muñoz RF. Understanding attrition from international internet health interventions: a step towards global eHealth. *Health Promot Int* 2013 Sep;28(3):442-452 [FREE Full text] [doi: [10.1093/heapro/das029](https://doi.org/10.1093/heapro/das029)] [Medline: [22786673](https://pubmed.ncbi.nlm.nih.gov/22786673/)]
68. Apter AJ. Can patient portals reduce health disparities? a perspective from asthma. *Ann Am Thorac Soc* 2014 May;11(4):608-612. [doi: [10.1513/AnnalsATS.201401-032PS](https://doi.org/10.1513/AnnalsATS.201401-032PS)] [Medline: [24640983](https://pubmed.ncbi.nlm.nih.gov/24640983/)]
69. Eddens KS, Kreuter MW, Morgan JC, Beatty KE, Jasim SA, Garibay L, et al. Disparities by race and ethnicity in cancer survivor stories available on the web. *J Med Internet Res* 2009;11(4):e50 [FREE Full text] [doi: [10.2196/jmir.1163](https://doi.org/10.2196/jmir.1163)] [Medline: [19945948](https://pubmed.ncbi.nlm.nih.gov/19945948/)]
70. Cloos P. Pouvoirs, différence et stéréotypes. Québec: Presses de l'Université Laval; 2015:1-208.
71. Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. *Am J Prev Med* 2007 May;32(5 Suppl):S131-S133 [FREE Full text] [doi: [10.1016/j.amepre.2007.02.012](https://doi.org/10.1016/j.amepre.2007.02.012)] [Medline: [17466818](https://pubmed.ncbi.nlm.nih.gov/17466818/)]
72. Kontos EZ, Emmons KM, Puleo E, Viswanath K. Communication inequalities and public health implications of adult social networking site use in the United States. *J Health Commun* 2010;15 Suppl 3:216-235 [FREE Full text] [doi: [10.1080/10810730.2010.522689](https://doi.org/10.1080/10810730.2010.522689)] [Medline: [21154095](https://pubmed.ncbi.nlm.nih.gov/21154095/)]
73. Jones R. Development of a questionnaire and cross-sectional survey of patient eHealth readiness and eHealth inequalities. *Med* 2013 Sep;2(2):e9 [FREE Full text] [doi: [10.2196/med20.2559](https://doi.org/10.2196/med20.2559)] [Medline: [25075244](https://pubmed.ncbi.nlm.nih.gov/25075244/)]
74. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. *J Am Med Inform Assoc* 2011 May 1;18(3):318-321 [FREE Full text] [doi: [10.1136/jamia.2010.006015](https://doi.org/10.1136/jamia.2010.006015)] [Medline: [21262921](https://pubmed.ncbi.nlm.nih.gov/21262921/)]
75. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005 Feb;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
76. Atkinson NL, Desmond SM, Saperstein SL, Billing AS, Gold RS, Tournas-Hardt A. Assets, challenges, and the potential of technology for nutrition education in rural communities. *J Nutr Educ Behav* 2010 Nov;42(6):410-416. [doi: [10.1016/j.jneb.2009.09.004](https://doi.org/10.1016/j.jneb.2009.09.004)] [Medline: [20813589](https://pubmed.ncbi.nlm.nih.gov/20813589/)]
77. Bell AV. "think about Oprah": social class differences in sources of health information. *Qual Health Res* 2014 Apr;24(4):506-516. [doi: [10.1177/1049732314524637](https://doi.org/10.1177/1049732314524637)] [Medline: [24623661](https://pubmed.ncbi.nlm.nih.gov/24623661/)]
78. Brouwer W, Oenema A, Raat H, Crutzen R, de NJ, de Vries Nanne K, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. *Health Educ Res* 2010 Aug;25(4):585-595. [doi: [10.1093/her/cyp063](https://doi.org/10.1093/her/cyp063)] [Medline: [19897515](https://pubmed.ncbi.nlm.nih.gov/19897515/)]
79. 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 Nov;11(4):e48 [FREE Full text] [doi: [10.2196/jmir.1249](https://doi.org/10.2196/jmir.1249)] [Medline: [19945947](https://pubmed.ncbi.nlm.nih.gov/19945947/)]
80. Dart J. The internet as a source of health information in three disparate communities. *Aust Health Rev* 2008 Aug;32(3):559-569. [Medline: [18666885](https://pubmed.ncbi.nlm.nih.gov/18666885/)]
81. Dart JM, Gallois C. Community desires for an online health information strategy. *Aust Health Review* 2010;34(4):467-476. [doi: [10.1071/ah08719](https://doi.org/10.1071/ah08719)]
82. Dudas RA, Crocetti M. Pediatric caregiver attitudes toward email communication: survey in an urban primary care setting. *J Med Internet Res* 2013 Oct 23;15(10):e228 [FREE Full text] [doi: [10.2196/jmir.2738](https://doi.org/10.2196/jmir.2738)] [Medline: [24152542](https://pubmed.ncbi.nlm.nih.gov/24152542/)]
83. Eysenbach G. Poverty, human development, and the role of eHealth. *J Med Internet Res* 2007 Oct 22;9(4):e34 [FREE Full text] [doi: [10.2196/jmir.9.4.e34](https://doi.org/10.2196/jmir.9.4.e34)] [Medline: [17951217](https://pubmed.ncbi.nlm.nih.gov/17951217/)]
84. Fowles ER, Gentry B. The feasibility of personal digital assistants (PDAs) to collect dietary intake data in low-income pregnant women. *J Nutr Educ Behav* 2008 Nov;40(6):374-377. [doi: [10.1016/j.jneb.2007.08.015](https://doi.org/10.1016/j.jneb.2007.08.015)] [Medline: [18984494](https://pubmed.ncbi.nlm.nih.gov/18984494/)]

85. Gibbons MC, Casale CR. Reducing disparities in health care quality: the role of health IT in underresourced settings. *Med Care Res Rev* 2010 Oct;67(5 Suppl):155S-162S. [doi: [10.1177/1077558710376202](https://doi.org/10.1177/1077558710376202)] [Medline: [20829238](https://pubmed.ncbi.nlm.nih.gov/20829238/)]
86. Gordon NP, Hornbrook MC. Differences in access to and preferences for using patient portals and other eHealth technologies based on race, ethnicity, and age: a database and survey study of seniors in a large health plan. *J Med Internet Res* 2016 Mar 04;18(3):e50 [FREE Full text] [doi: [10.2196/jmir.5105](https://doi.org/10.2196/jmir.5105)] [Medline: [26944212](https://pubmed.ncbi.nlm.nih.gov/26944212/)]
87. Kerr C, Murray E, Noble L, Morris R, Bottomley C, Stevenson F, et al. The potential of Web-based interventions for heart disease self-management: a mixed methods investigation. *J Med Internet Res* 2010 Dec 02;12(4):e56 [FREE Full text] [doi: [10.2196/jmir.1438](https://doi.org/10.2196/jmir.1438)] [Medline: [21156471](https://pubmed.ncbi.nlm.nih.gov/21156471/)]
88. Kukafka R, Khan SA, Hutchinson C, McFarlane DJ, Li J, Ancker JS, et al. Digital partnerships for health: steps to develop a community-specific health portal aimed at promoting health and well-being. *AMIA Annu Symp Proc* 2007 Oct 11:428-432. [Medline: [18693872](https://pubmed.ncbi.nlm.nih.gov/18693872/)]
89. Liszka HA, Steyer TE, Hueston WJ. Virtual medical care: how are our patients using online health information? *J Community Health* 2006 Jun 3;31(5):368-378. [doi: [10.1007/s10900-006-9019-3](https://doi.org/10.1007/s10900-006-9019-3)]
90. Lorence D, Park H. Study of education disparities and health information seeking behavior. *Cyberpsychol Behav* 2007 Feb;10(1):149-151. [doi: [10.1089/cpb.2006.9977](https://doi.org/10.1089/cpb.2006.9977)] [Medline: [17305464](https://pubmed.ncbi.nlm.nih.gov/17305464/)]
91. Malone M, While A, Roberts J. Parental health information seeking and re-exploration of the 'digital divide'. *Prim Health Care Res Dev* 2014 Apr;15(2):202-212. [doi: [10.1017/S1463423613000194](https://doi.org/10.1017/S1463423613000194)] [Medline: [23676618](https://pubmed.ncbi.nlm.nih.gov/23676618/)]
92. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. *J Med Internet Res* 2010 Dec;12(5):e60 [FREE Full text] [doi: [10.2196/jmir.1463](https://doi.org/10.2196/jmir.1463)] [Medline: [21169162](https://pubmed.ncbi.nlm.nih.gov/21169162/)]
93. Neuenschwander LM, Abbott A, Mobley AR. Assessment of low-income adults' access to technology: implications for nutrition education. *J Nutr Educ Behav* 2012 Jan;44(1):60-65. [doi: [10.1016/j.jneb.2011.01.004](https://doi.org/10.1016/j.jneb.2011.01.004)] [Medline: [21924959](https://pubmed.ncbi.nlm.nih.gov/21924959/)]
94. Newman L, Biedrzycki K, Baum F. Digital technology use among disadvantaged Australians: implications for equitable consumer participation in digitally-mediated communication and information exchange with health services. *Aust Health Rev* 2012 May;36(2):125-129. [doi: [10.1071/AH11042](https://doi.org/10.1071/AH11042)] [Medline: [22624630](https://pubmed.ncbi.nlm.nih.gov/22624630/)]
95. Nölke L, Mensing M, Krämer A, Hornberg C. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. *BMC Public Health* 2015;15:31 [FREE Full text] [doi: [10.1186/s12889-015-1423-0](https://doi.org/10.1186/s12889-015-1423-0)] [Medline: [25631456](https://pubmed.ncbi.nlm.nih.gov/25631456/)]
96. Sadah SA, Shahbazi M, Wiley MT, Hristidis V. A study of the demographics of web-based health-related social media users. *J Med Internet Res* 2015 Aug 06;17(8):e194 [FREE Full text] [doi: [10.2196/jmir.4308](https://doi.org/10.2196/jmir.4308)] [Medline: [26250986](https://pubmed.ncbi.nlm.nih.gov/26250986/)]
97. Schmeida M, McNeal RS. The telehealth divide: disparities in searching public health information online. *J Health Care Poor Underserved* 2007 Aug;18(3):637-647. [doi: [10.1353/hpu.2007.0068](https://doi.org/10.1353/hpu.2007.0068)] [Medline: [17675719](https://pubmed.ncbi.nlm.nih.gov/17675719/)]
98. Vanderbilt AA, Wright MS, Brewer AE, Murithi LK, Coney P. Increasing knowledge and health literacy about preterm births in underserved communities: an approach to decrease health disparities, a pilot study. *Glob J Health Sci* 2015 May 17;8(1):83-89 [FREE Full text] [doi: [10.5539/gjhs.v8n1p83](https://doi.org/10.5539/gjhs.v8n1p83)] [Medline: [26234992](https://pubmed.ncbi.nlm.nih.gov/26234992/)]
99. Viswanath K, Ackerson LK. Race, ethnicity, language, social class, and health communication inequalities: a nationally-representative cross-sectional study. *PLoS One* 2011 Jan 18;6(1):e14550 [FREE Full text] [doi: [10.1371/journal.pone.0014550](https://doi.org/10.1371/journal.pone.0014550)] [Medline: [21267450](https://pubmed.ncbi.nlm.nih.gov/21267450/)]
100. Wang J, Bennett K, Probst J. Subdividing the digital divide: differences in internet access and use among rural residents with medical limitations. *J Med Internet Res* 2011 Mar 03;13(1):e25 [FREE Full text] [doi: [10.2196/jmir.1534](https://doi.org/10.2196/jmir.1534)] [Medline: [21371989](https://pubmed.ncbi.nlm.nih.gov/21371989/)]
101. West DM, Miller EA. The digital divide in public e-health: barriers to accessibility and privacy in state health department websites. *J Health Care Poor Underserved* 2006 Aug;17(3):652-667. [doi: [10.1353/hpu.2006.0115](https://doi.org/10.1353/hpu.2006.0115)] [Medline: [16960328](https://pubmed.ncbi.nlm.nih.gov/16960328/)]
102. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. *Arch Intern Med* 2011 Mar 28;171(6):568-574. [doi: [10.1001/archinternmed.2011.34](https://doi.org/10.1001/archinternmed.2011.34)] [Medline: [21444847](https://pubmed.ncbi.nlm.nih.gov/21444847/)]

Abbreviations

AMSTAR: assessment of multiple systematic reviews

CINAHL: Cumulative Index to Nursing and Allied Health Literature

DVD: digital versatile disk

EPPI: Evidence for Policy and Practice of Information of the Institute of Education at the University of London

HIV: human immunodeficiency virus

PC: personal computer

SHI: social health inequalities

Edited by G Eysenbach; submitted 01.10.16; peer-reviewed by P Pluye, J Apolinário-Hagen, I Mircheva, TR Soron, A Paglialonga; comments to author 29.12.16; revised version received 09.02.17; accepted 02.03.17; published 27.04.17.

Please cite as:

Latulippe K, Hamel C, Giroux D

Social Health Inequalities and eHealth: A Literature Review With Qualitative Synthesis of Theoretical and Empirical Studies

J Med Internet Res 2017;19(4):e136

URL: <http://www.jmir.org/2017/4/e136/>

doi: [10.2196/jmir.6731](https://doi.org/10.2196/jmir.6731)

PMID: [28450271](https://pubmed.ncbi.nlm.nih.gov/28450271/)

©Karine Latulippe, Christine Hamel, Dominique Giroux. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Influence of Telecommunication Modality, Internet Transmission Quality, and Accessories on Speech Perception in Cochlear Implant Users

Georgios Mantokoudis¹, MD; Roger Koller¹, B Med; Jérémie Guignard¹, MSc, PhD; Marco Caversaccio¹, MD; Martin Kompis¹, MD, PhD; Pascal Senn^{1,2}, MD

¹Inselspital, Bern University Hospital, Department of Otorhinolaryngology, Head and Neck Surgery, University of Bern, Bern, Switzerland

²Department of Clinical Neurosciences, Service ORL and CCF, University Hospital of Geneva, Geneva, Switzerland

Corresponding Author:

Georgios Mantokoudis, MD

Inselspital, Bern University Hospital

Department of Otorhinolaryngology, Head and Neck Surgery

University of Bern

Freiburgstrasse

Bern, 3010

Switzerland

Phone: 41 316323321

Fax: 41 316324900

Email: georgios.mantokoudis@insel.ch

Abstract

Background: Telecommunication is limited or even impossible for more than one-thirds of all cochlear implant (CI) users.

Objective: We sought therefore to study the impact of voice quality on speech perception with voice over Internet protocol (VoIP) under real and adverse network conditions.

Methods: Telephone speech perception was assessed in 19 CI users (15-69 years, average 42 years), using the German HSM (Hochmair-Schulz-Moser) sentence test comparing Skype and conventional telephone (public switched telephone networks, PSTN) transmission using a personal computer (PC) and a digital enhanced cordless telecommunications (DECT) telephone dual device. Five different Internet transmission quality modes and four accessories (PC speakers, headphones, 3.5 mm jack audio cable, and induction loop) were compared. As a secondary outcome, the subjective perceived voice quality was assessed using the mean opinion score (MOS).

Results: Speech telephone perception was significantly better (median 91.6%, $P<.001$) with Skype compared with PSTN (median 42.5%) under optimal conditions. Skype calls under adverse network conditions (data packet loss > 15%) were not superior to conventional telephony. In addition, there were no significant differences between the tested accessories ($P>.05$) using a PC. Coupling a Skype DECT phone device with an audio cable to the CI, however, resulted in higher speech perception (median 65%) and subjective MOS scores (3.2) than using PSTN (median 7.5%, $P<.001$).

Conclusions: Skype calls significantly improve speech perception for CI users compared with conventional telephony under real network conditions. Listening accessories do not further improve listening experience. Current Skype DECT telephone devices do not fully offer technical advantages in voice quality.

(*J Med Internet Res* 2017;19(4):e135) doi:[10.2196/jmir.6954](https://doi.org/10.2196/jmir.6954)

KEYWORDS

communication aids for disabled; telecommunications devices for the deaf; cochlear implants; speech discrimination tests; hearing loss; telephone

Introduction

A cochlear implant (CI) is an electronic device which allows an auditory stimulation in patients with severe or profound hearing loss [1,2]. It is the most successful neural prosthesis developed till date [1,3]. The implant consists of two parts, an external speech processor connected with a transmitting coil worn behind the ear and the implant itself placed under the skin behind the ear. The internal part of the implant receives the signals from the transmitting coil and sends electrical impulses to 12-22 electrodes, which are placed into the cochlea. The auditory nerve is hereby directly stimulated with high frequencies at the base of the cochlea and low frequencies at the apex [4]. CIs offer an improved hearing and quality of life [5]. Telecommunication, however, is limited or even impossible for more than one-third of CI users [6-10]. Assisting listening devices may improve speech perception performance, however, communication abilities still remain limited because of restricted frequency bandwidth (300-3400 Hz) and digital compression of voice data applied in conventional telephony (public switched telephone network, PSTN) [11,12]. Telephone speech perception might be additionally impaired by the coupling mode with a hearing aid or CI [13,14]. Ability to use a telephone is important for maintaining social contacts or in emergency situations. In addition, cognitive decline and dementia in older adults are often associated with hearing loss and reduced communication abilities [15-18]. Rehabilitation of hearing communication, however, improves cognitive function [19], quality of life and social participation, and any solution to improve telecommunication in CI patients should be pursued.

Recent laboratory studies showed advantages of Internet telephony (voice over Internet protocol, VOIP) with improved voice quality caused by a wider frequency bandwidth (200-8500Hz) than the conventional telephone [20,21]; however,

all these studies were performed under laboratory conditions. There is currently no study showing this advantage under real network conditions. Software solutions such as Skype or Google Talk among others are supporting video telephony, which improves speech perception by adding visual cues [22]. In addition, VoIP software offers a wider range of transmitted frequencies (200Hz-12kHz, Figure 1) and should—in theory—offer a better voice quality and speech perception performance; however, limited Internet connection speed might deteriorate speech signals and voice quality [20,23,24]. The final transmitted voice quality depends on data transmission network capacities, delays, and the extension of data packet loss (PL) [25].

The aim of this study was to test telephone speech perception in CI users comparing the conventional telephone (PSTN) with VoIP (Skype) under real network conditions. Voice quality, coupling mode, accessories, and a Skype telephone digital enhanced cordless telecommunications (DECT) device were assessed to observe any improvement in distant communication.

Methods

Test Subjects

Nineteen adult CI users aged between 15 and 69 years (average 42 years) participated in the study. We included CI users with at least 6-month unilateral implant experience and either a Cochlear Nucleus Freedom or a CP810 Sound Processor fitted with a frequency allocation table reaching higher than 5 kHz. Each test subject had a minimal speech perception score of 50% for German monosyllabic words at 60 dB sound pressure level (SPL), 3 months after implantation. Table 1 shows patient characteristics of the included subjects. The study protocol was fully approved by the local institutional review board. All patients gave written informed consent.

Figure 1. Test setup and VoIP. The standardized version of the HSM sentence test (for clinical testing) was sent from a CD player connected via an audio mixing console (XENYX 502 Behringer, Willich, Germany) to either a voice modem (A; Way2Call, Hi-Phone Desktop lite, Way2Call Communications Inc, Newmark, NJ, USA) or a laptop (B). Small voice data packets were sent from Skype PC software (B) through a router (C) over a transmission control protocol and Internet protocol network to the receiver. The receiver devices (D) or (E) branched to a router (C) collected all incoming data packets. The installed Skype PC software (E) or Skype App (D) was decoding the voice signal. A software controlled the number of lost data packets (0-20%), to induce different network scenarios. Alternatively, a conventional telephone line (PSTN connection) was used to transmit the audio signal from a voice modem (A) to a DECT telephone (D).

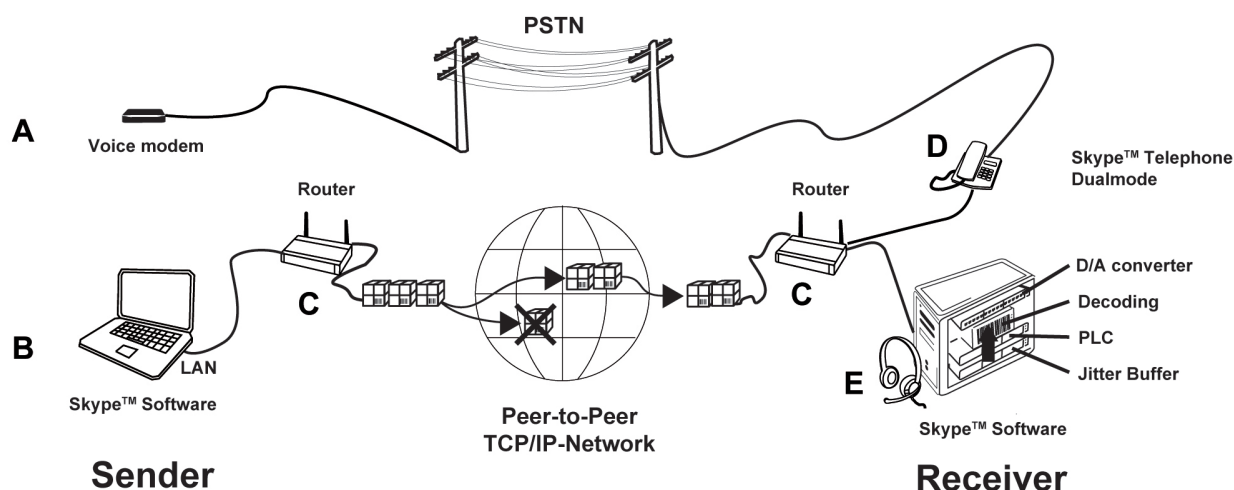


Table 1. Clinical data of cochlear implant (CI) users.

ID	Age at measurement (years)	Gender	Cochlear implant Model	Speech processor	Age at implantation (years)	Years since implantation
1	51	Female	CI512	CP 810	50	1
2	24	Female	CI 512	CP 810	23	1
3	61	Male	CI 512	CP 810	59	2
4	40	Male	CI 24RE(CA)	Freedom	34	6
5	18	Male	CI 24RE(CA)	Freedom	12	6
6	15	Male	CI 24RE(CA)	Freedom	9	6
7	17	Male	CI 24RE(ST)	Freedom	12	5
8	63	Male	CI 512	CP 810	62	1
9	67	Female	Hybrid L24	Freedom	65	2
10	54	Female	CI 24RE(CA)	Freedom	48	6
11	25	Female	CI 24RE(CA)	Freedom	22	3
12	15	Female	CI 24RE(ST)	Freedom	12	3
13	31	Female	CI 24RE(CA)	Freedom	23	8
14	68	Male	Hybrid L24	Freedom	66	2
15	27	Male	CI 512	CP 810	26	1
16	55	Male	CI 512	CP 810	54	1
17	69	Male	CI 24R(CA)	CP 810	61	8
18	58	Female	CI 24(ST)	CP 810	48	10
19	40	Male	CI 422	CP 810	39	1

Telephone Transmission Mode and Devices

We compared Internet telephony (Skype, local area network connection) against the conventional telephone (PSTN, landline connection) using two devices, a personal computer (PC; Latitude E6510, Dell, Round Rock, TX, USA) and a cordless DECT telephone device (Philips VoIP855, Royal Philips Electronics, Amsterdam, The Netherlands) which has dual transmission functions (Skype app and PSTN). In addition, we compared four accessories (PC speakers Z320; Logitech headphones, Behringer HPS 500; 3.5 mm jack audio cable, cochlear Ltd; and induction loop) and five different Internet transmission speeds (controlled with a connection emulator, Perfect Soft Research, Version 1.3.2 Brisbane, Australia). We used Internet connections with random data PL in 5% steps ranging from 0% PL (perfect), 5% PL (mild), 10% PL (medium), 15% (severe) to 20% PL (very severe). [Figure 1](#) shows the test setup with the two transmission modes (PSTN versus Skype) connected with two devices (PC, [Figure 1 E](#) and cordless DECT telephone, [Figure 1 D](#)).

The frequency response was measured for all devices using an audio analyzer (UPV, Rhode & Schwarz, Munich, Germany) and a head and torso simulator (KEMAR Manikin Type 45BA, Brüel & Kjaer, Naerum, Denmark). For acoustic measurements, the Manikin's ear simulator (Type 4158) was used to simulate the situation of a telephone held on the ear. The ear simulator is composed of a silicon external ear (or pinna) coupled to an ear canal terminated by a half-inch condenser microphone and pre-amplifier. A sweep of 50 logarithmically spread, pure sinus tones was generated and the output was filtered with a 1% bandwidth filter locked on the stimulus frequency. The root mean square amplitude of the output was calculated to create a frequency-domain graph. Objective voice quality was tested using the audio analyzer's built-in algorithm for Perceptual Evaluation of Speech Quality (PESQ) based on the guidelines of the ITU (International Telecommunication Union) in accordance with the ITU-T P.862 protocol. Test subjects rated the subjective perceived voice quality from 1-5 using the mean opinion score (MOS; [Table 2](#), according to specifications ITU-T Rec. P.862.1 and P.862.2) [26].

Table 2. Mean Opinion Score (MOS) for subjective voice quality assessment.

Score	Quality	Listening effort scale
5	Excellent	No effort required
4	Good	No appreciable effort required
3	Fair	Moderate effort required
2	Poor	Considerable effort required
1	Bad	No meaning understood with reasonable effort

Speech Perception Test Protocol

The standardized German “HSM” sentence test [27] was used for open set monosyllable speech perception testing in noise (60 dB SPL) at a constant signal sound level of 70 dB SPL (free field at 1-m distance). The HSM test consists of 30 lists with 20 short sentences containing 106 monosyllable words. The order of the sentence was changed at random to avoid learning effects. The subject had to repeat the presented sentences, and received 1 point for each correctly reproduced word. The percent of speech perception from a total of 106 words was calculated for each condition (one list). Test subjects were tested monaurally in a sound treated room (Type 402A, Industrial Acoustics Company, Niederkrüchten, Germany), with an average reverberation time of 0.10 s (125-10000 Hz). The contralateral ear canal was closed by an earplug (USA EARlink 3C, EAR Corporation, Indianapolis) if there was a residual hearing. Other hearing aids or a second CI had to be turned off. Patients kept the everyday settings of their speech processor. The cordless DECT telephone (Figure 1 D) was coupled to the implant either with an audio cable (Personal Audio Cable, Cochlear Limited, Sydney, Australia) branched to the 3.5-mm jack socket of the handset or by holding it directly to the ear. We chose this experimental setting with speech signal presented in noise not only to simulate an everyday listening situation but also to avoid ceiling effects.

Statistics

Robust nonparametric analyses were performed to assess the potentially non-normally distributed speech perception scores from this small study population. A two-tailed Wilcoxon

matched pairs signed-rank test was used to compare Skype with PSTN. For the ideal condition with no Internet data PL (condition 0% PL), a one-tailed test was applied because of the expected superiority of Skype under this condition [21,23]. A P value < .05 was considered significant after applying a Bonferroni correction for multiple testing.

Results

Telephone Transmission Mode and Voice Quality

The differences of the transmitted frequencies of PSTN versus Skype (LAN connection) are shown in Figure 2. There was a better frequency response of the audio signal derived from the headphone jack compared with the telephone handset (Figure 2). A broader frequency range (50–5000 Hz) was transmitted via Skype than PSTN (200–3000 Hz) independently of whether the built-in loudspeaker or the headphone jack was used.

Speech perception with an optimal Skype connection (0% PL; median 91.6%, $n=18$, range 48.1–99.1%; interquartile range, IQR, 15.6) which was significantly better ($P<.001$) than the telephone speech perception (PSTN median 42.5%, $n=18$, range 11.3–85.8%, IQR 15.6); however, there was statistically no advantage in speech understanding using Skype at lower Internet quality connections (PL 15% and 20%; Figure 3). Quality measurements of the transmitted sound (PESQ measurement) showed that all Skype quality connections (PL 0–20%) offered a significantly better voice quality ($P<.001$) compared to PSTN (Figure 3). Skype voice quality was maintained even with a medium quality Internet connection (PL 5% and 10%; Figure 3)

Figure 2. Frequency response for Skype and telephone. It shows the frequency response of the conventional telephone (PSTN) and Skype (LAN connection). The transmitted audio signal was tested using an audio analyzer and a head and torso simulator. The x-axis shows the logarithmic frequency scale, and the y-axis, the recorded sound pressure level or electrical output from either the handset loudspeaker (panel A) or the handset 3.5-mm headphone jack (panel B) of the telephone.

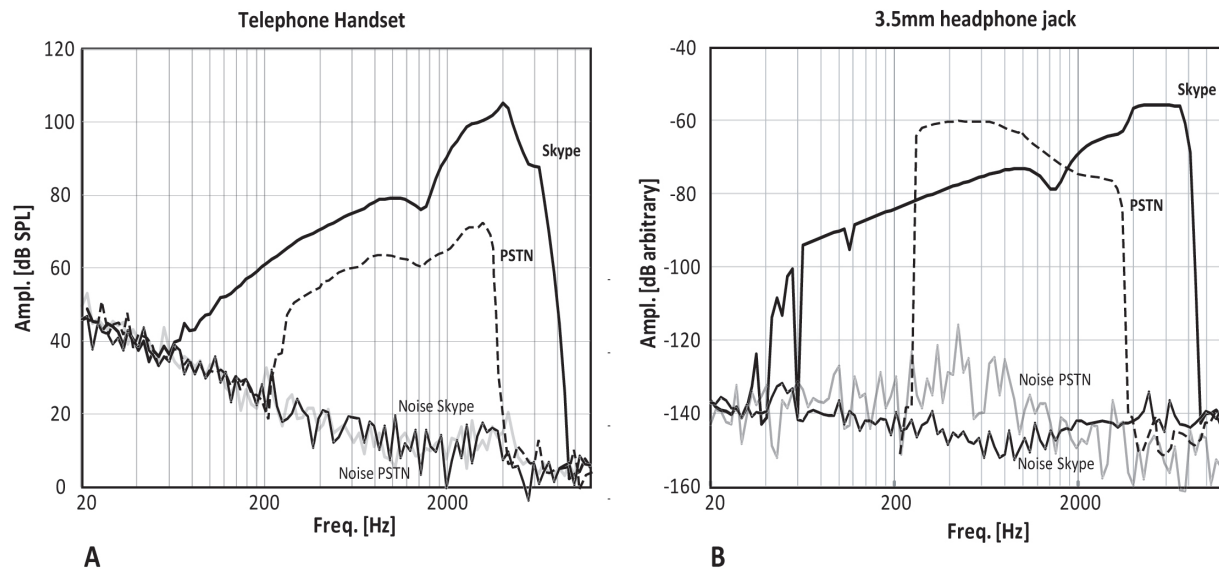
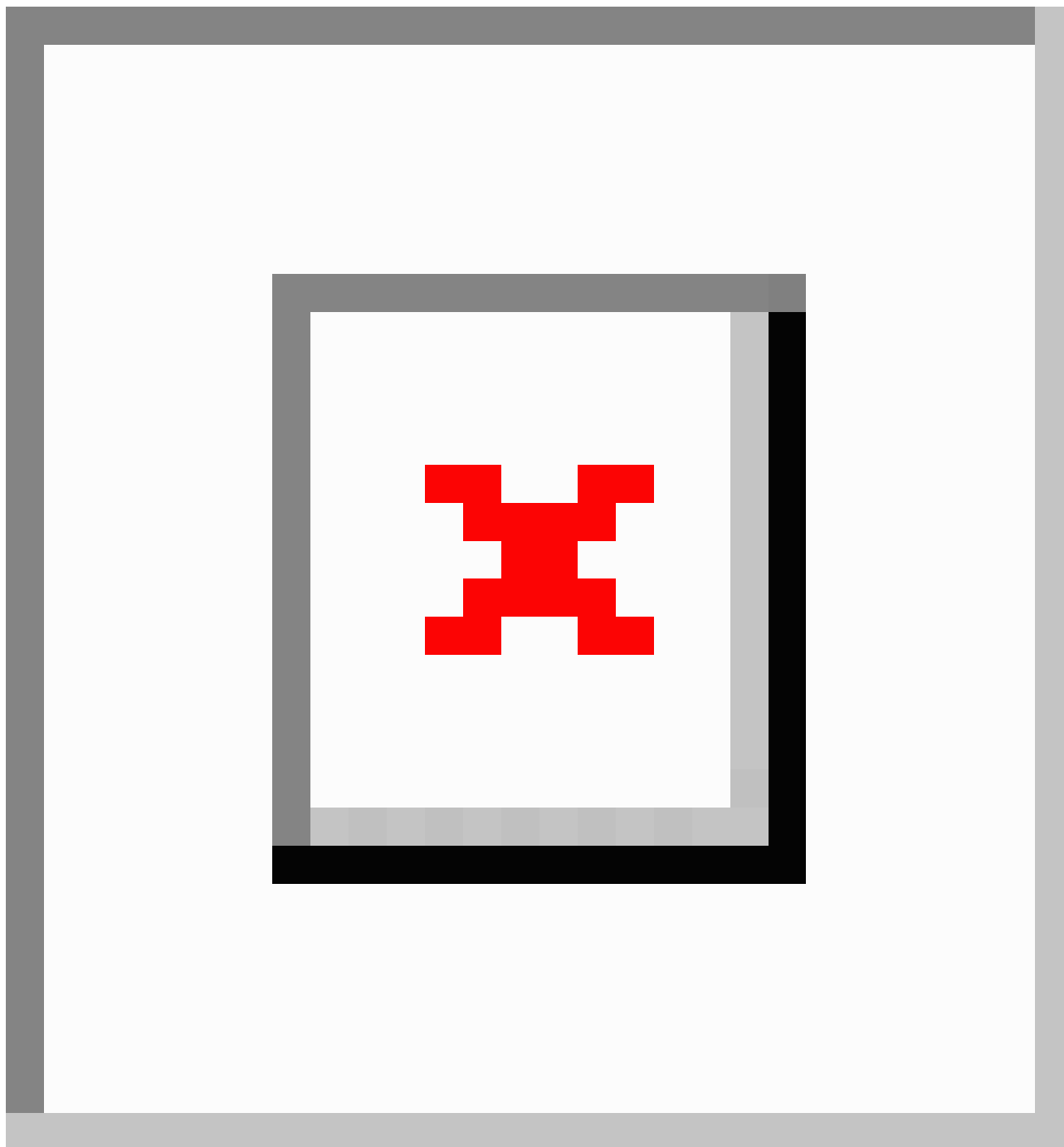


Figure 3. Speech perception and transmission mode. Box plots demonstrating lower quartile, median, and upper quartile, and whiskers representing 1.5 times the IQR (X=outliers): Free-field speech perception performance (correctly repeated words in percentage) from 19 CI users at 5 different Skype transmissions with 0-20% data PL and one landline connection (PSTN). Stars indicate a statistically significant difference between a Skype and a PSTN transmission. PESQ scores measurements (panel B) (assessed with an audio analyzer) indicate the objective measured voice quality (y-axis) for the different test conditions (Skype connection with data PL from 0% to 20% and PSTN). Panel C shows the subjective perceived voice quality under the same conditions.

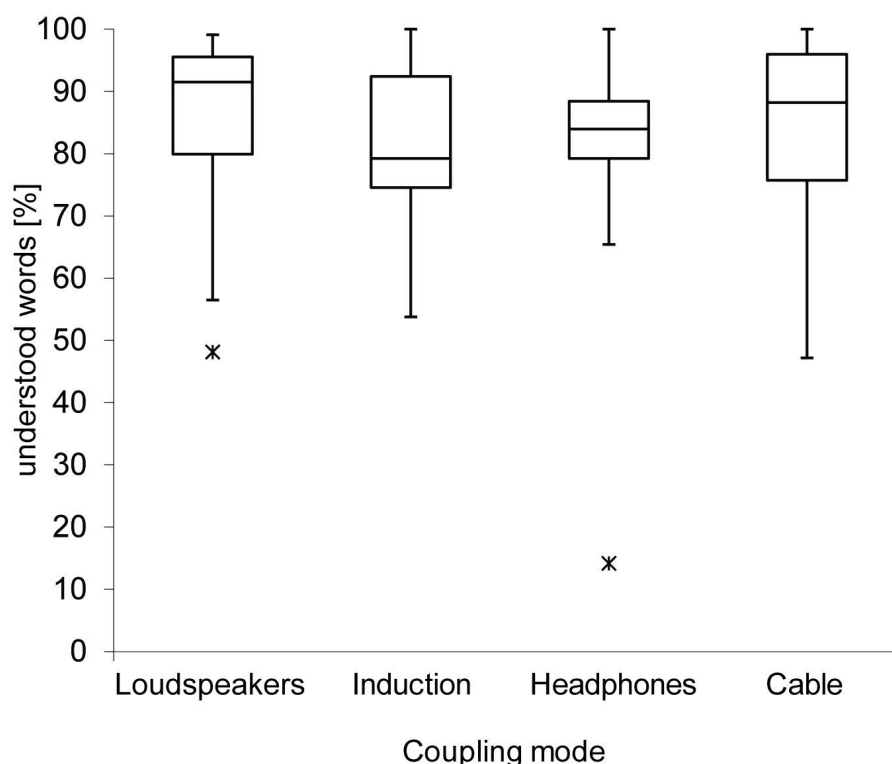


VoIP Accessories for PC and CI

Free-field speech perception with Skype (PC version) using an optimal Internet connection (0% PL) was 91.5 % (median, n=18, range 48.1-99.1 %, IQR 15.6). Speech perception with a connected induction loop was 79.3% (median, n=15, range

53.8-100%, IQR 17.92), with headphones 83.9 % (median, n=18, range 14.2-100 %, IQR 9.215), and with coupled CI cables 88.2 % (median, n=18, range 47.1-100 %, IQR 20.27). There was no significant difference ($P>.05$) between the tested accessories (Figure 4).

Figure 4. Speech perception scores using telephone accessories. Four different accessories have been compared in terms of speech perception under optimal Internet connections (0% data PL): an induction loop, headphones, an audio cable, and active loudspeakers connected to a Skype PC have been tested. Box plots are indicating the median percentage of word recognition. There was no significant difference across the tested accessories.



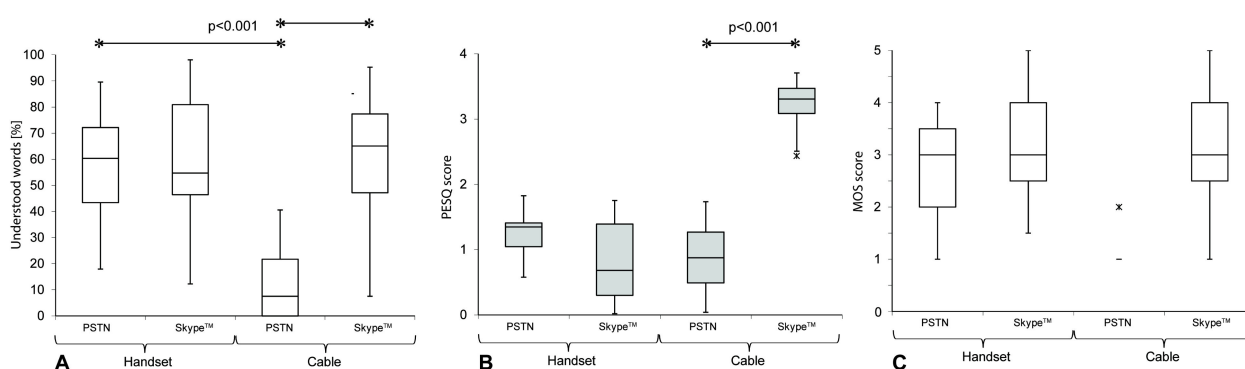
Skype DECT Phone Device and Coupling Mode

Figure 5 shows the speech perception performance with a DECT telephone coupled to the CI either with a cable or the handset. Median speech perception using a CI audio cable was 7.5% for PSTN ($n=17$, range 0-40.6%, IQR 21.7) compared with speech understanding of 65.1% with the Skype app (median, $n=17$, range 47.17-95.3%, IQR 30.2) installed on the DECT telephone. Speech understanding with Skype was significantly superior if the DECT telephone was directly coupled to the CI by cable ($P<.001$). There was, however, no significant advantage seen

for Skype if the handset was held directly near the microphone above the pinna (Figure 5).

The voice quality measurements (PESQ) on the headphone jack socket of a DECT phone (cable connection) showed significantly higher scores for the installed Skype app (median 3.08, $n=19$, range 2.43- 3.71, IQR 0.285) compared with the PSTN connection (median 0.73, $n=19$, range 0.33-1.73, IQR 0.94, $P<.001$) on the same device (Figure 5). The built-in telephone loudspeakers of the handset, however, did not show any significant voice quality differences between PSTN and the Skype app (Figure 5).

Figure 5. Speech perception and coupling mode. It shows speech perception scores using a DECT phone with dual transmission modes (either PSTN or Skype). The handset was either coupled to the CI microphone above the ear (handset coupling mode) or connected directly over the 3.5-mm headphone jack to the CI (cable coupling). Stars indicate significant differences. The objective voice quality (y-axis) was measured on a head and torso simulator for the same conditions (panel B). The subjective perceived voice quality (MOS) was rated by the participants for each condition (panel C).



Mean Opinion Score (MOS)

The subjective perceived voice quality, which describes the level of effort required for understanding, was best with an optimal Skype connection (MOS 4.1 [SD 0.9] at PL 0%; [Table 3](#); [Figure 3](#)). The audio cable was rated as the best accessory

requiring the least effort for understanding speech (MOS 4.2, SD 0.7), while the neck loop was found to offer the most unpleasant sound (MOS 3.5, SD 1.2). [Figures 3](#) and [5](#) show a comparison between psychoacoustic (panels A), technical (panels B), and subjective (panels C) performance.

Table 3. Mean opinion scores (MOS) for each modality

Modality	Method	Mean opinion scores (MOS) ^a
VoIP vs PSTN		
	Skype connection PL 0%	4.1
	Skype connection PL 5%	4
	Skype connection PL 10%	3.6
	Skype connection PL 15%	3.6
	Skype connection PL 20%	2.6
	Telephone connection (PSTN)	2.1
Accessories		
	PC-Speakers	4.1
	Headphones	3.9
	Neck Induction Loop	3.5
	Audio cable	4.2
Cordless DECT telephones		
	PSTN Handset coupling	2.9
	PSTN Audio cable	1.2
	Skype Handset coupling	3.2
	Skype Audio cable	3.2

^aSubjective perceived voice quality was assessed using a 5-point unipolar rating scale (MOS, [Table 2](#)), ranging from 5 points indicating an excellent voice quality to 1 point indicating a bad voice quality.

Discussion

Principal Findings

Speech perception by CI patients using Skype with active loudspeakers was superior to the conventional telephone under perfect or medium network voice transmissions. There was no advantage of Internet telephony for transmissions with severe or very severe data PL (>10% PL). Accessories such as a neck loop with wireless telecoil, an audio cable directly connected to the CI or headphones, did not further improve speech perception; however, the CI users subjectively perceived the voice quality (MOS) as superior, when using direct cable input. A dedicated Skype DECT telephone did not offer the full advantage of the superior voice quality provided by the Skype app because of loudspeaker quality limitations.

Impact of Internet Connection Quality

A proof of concept was given by a previous experimental study showing a superior voice quality of Internet telephony resulting in better speech understanding, provided that the Internet connection quality was ideal [\[21\]](#). This phenomenon was also shown in normal hearing subjects and was not related to any speech coding strategy of a CI [\[21\]](#).

Sound quality might be affected by low bit-rate coding, data PL, background noise, silence suppression, or by network filtering leading to sound delay, jitter, noise, and speech level changes. These parameters might not equally influence speech perception performance by CI users. The PESQ score, however, uses objective parameters to model psychoacoustic and cognitive perception of speech. Such a model was calibrated to predict MOS scores and to improve correlation between PESQ and MOS scores [\[26\]](#).

Live Skype calls (desktop version) transmitted via a deteriorated Internet connection (5-10% PL) still generated a better voice quality on the receiver side compared with a PSTN connection, but not for transmission modes with severe PL (15% and 20%). These findings are in line with speech perception test results performed under simulated laboratory conditions [\[20\]](#).

The main reason for the superiority of VoIP is likely due to technical reasons. The frequency coverage with VoIP is higher than that of conventional telephony (0.1-8 kHz vs 0.3-3.4 kHz; [Figure 2](#)) resulting in better audio quality (see PESQ measurements). Since Skype transmits higher frequencies, it is likely to convey more of the speech-relevant signal content such as consonants. In addition, telecommunication companies digitalize the analogue voice signal using low bit-rate coding (8 kHz sampling rate, G.711 codec, ITU recommendations,

G-series) and maximum bit rates of 33.6 kbits/s compared with Skype, which uses a sampling rate of 16 kHz and variable bit rates up to 40 kbits/s [28].

No improvement in speech perception, however, was measured for the CI users when PL was >10% (Figure 3). This was unexpected after analyzing objective voice quality measurements (PESQ), which proved a better audio quality of Skype, regardless of the Internet connection quality (Figure 3). All tests were performed at a constant background noise, making these tests more sensitive to speech signal deterioration, which might have an impact on speech perception performance.

Telephone Listening Accessories

Accessories may improve telephone listening experience in background noise [29]. We measured no significant speech discrimination differences whether accessories were used or not. Possible ceiling effects might have occurred, since median speech perception performance was around 90% regardless of the accessory used. Test conditions with a lower signal-to-noise ratio (SNR < 10 dB) might have yielded different performance results since assisting listening accessories protect from unwanted noise by either shielding physically (around-the-ear/circumaural headphones) or by routing the signal directly to the device (cable, induction, FM system). Loudspeakers, however, do not offer an improved SNR, which is important for speech discrimination in CI users.

We tested only monaurally to simulate an everyday telephone use, and this might have affected the general performance with accessories. Studies show significant advantages in speech perception with binaural hearing (particularly in environmental noise) [30] or even in bimodal hearing [31]. Assistive listening devices enable users to transmit sound on both ears simultaneously.

Our data suggest that the choice of accessories should be made according to personal preference. Low quality built-in loudspeakers of the Skype DECT device resulted in lower voice quality and consequently in a lower speech perception performance. The Skype DECT device offered a better speech perception if the headphone jack was used. Most subjects also reported the best subjective listening experience with the audio cable connecting the headphone jack with the CI, although a cable connection limits the range of use. This might change with new wireless streaming possibilities [29,32,33], which have not been tested yet with VoIP applications.

Strength and Limitations

This study analyzed telephone speech perception using real-time settings and measurements, which give a better estimate than the previous laboratory tests.

The patients kept the everyday settings of their speech processor, however, another way to improve the telephone listening experience is the fine structure preservation [6] or the application of a special telephone fitting mode [34] by reducing the current level for electrodes stimulating outside the transmitted frequency range.

Although all individuals were blinded regarding the tested condition, performance bias might still be possible, since individuals were aware of the used accessories or might have perceived the presented voice quality. The small sample size might lead to an under- or overestimation of study results. In addition, this is a self-controlled study without any other control group.

We tested one single VoIP software and selected accessories. These results are therefore not generalizable for all available VoIP programs or other Internet transmissions modes.

Further research is mandatory to test voice quality and speech perception performance through wireless and mobile Internet connections. New technologies such as wireless audio streaming [35] may further improve listening experience and performance. New generations of Bluetooth technology with low battery consumption and direct connection to the implant might replace any assisting telephone accessory in the future while preserving voice quality.

Clinical Implications

Internet telephony improves speech perception performance even under real and adverse network conditions. CI users who are not able to have a meaningful telephone conversation could improve their telephone listening experience by using Skype or any other broadband Internet telephony service. This might have a direct impact on social integration, general health, life expectancy [16], and cognitive function in the elderly [18,36,37]. Restrictions in interactions and activities because of hearing loss might result in reduction of the overall health status and thus increase morbidity and mortality [16].

Skype and other VoIP software are freely available and can be used with any computer, microphone and speaker system. Additional assisting listening devices and telephone accessories might help for binaural hearing or hearing in noise.

Conclusions

Broadband VoIP software such as Skype can significantly improve telecommunication experience for CI users even with low quality Internet connections. Listening accessories such as headphones, audio cables, or an induction loop were equivalent in terms of speech perception performance. Microphone and speaker quality of Skype telephone DECT devices do not fully exploit benefits of Skype apps which provide an enhanced broadband audio and voice quality.

Acknowledgments

We thank Prof Dr Thomas Linder and Nicola De Min from the ENT Clinic at the Kantonsspital Luzern, Switzerland, for their support and recruitment of CI patients. Cochlear Europe Ltd, Addlestone, UK provided financial support for this study.

Conflicts of Interest

None declared.

References

1. Géléoc GS, Holt JR. Sound strategies for hearing restoration. *Science* 2014 May 09;344(6184):1241062 [FREE Full text] [doi: [10.1126/science.1241062](https://doi.org/10.1126/science.1241062)] [Medline: [24812404](https://pubmed.ncbi.nlm.nih.gov/24812404/)]
2. Basura GJ, Eapen R, Buchman CA. Bilateral cochlear implantation: current concepts, indications, and results. *Laryngoscope* 2009 Dec;119(12):2395-2401. [doi: [10.1002/lary.20751](https://doi.org/10.1002/lary.20751)] [Medline: [19894280](https://pubmed.ncbi.nlm.nih.gov/19894280/)]
3. Brand Y, Senn P, Kompis M, Dillier N, Allum JH. Cochlear implantation in children and adults in Switzerland. *Swiss Med Wkly* 2014 Feb 04;144:w13909 [FREE Full text] [doi: [10.4414/smw.2014.13909](https://doi.org/10.4414/smw.2014.13909)] [Medline: [24496729](https://pubmed.ncbi.nlm.nih.gov/24496729/)]
4. Hochmair I, Hochmair E, Nopp P, Waller M, Jolly C. Deep electrode insertion and sound coding in cochlear implants. *Hear Res* 2015 Apr;322:14-23 [FREE Full text] [doi: [10.1016/j.heares.2014.10.006](https://doi.org/10.1016/j.heares.2014.10.006)] [Medline: [25456089](https://pubmed.ncbi.nlm.nih.gov/25456089/)]
5. Gaylor J, Raman G, Chung M, Lee J, Rao M, Lau J, et al. Cochlear implantation in adults: a systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg* 2013 Mar;139(3):265-272. [doi: [10.1001/jamaoto.2013.1744](https://doi.org/10.1001/jamaoto.2013.1744)] [Medline: [23429927](https://pubmed.ncbi.nlm.nih.gov/23429927/)]
6. Galindo J, Lassaletta L, Mora RP, Castro A, Bastarrica M, Gavilán J. Fine structure processing improves telephone speech perception in cochlear implant users. *Eur Arch Otorhinolaryngol* 2013 Mar;270(4):1223-1229. [doi: [10.1007/s00405-012-2101-9](https://doi.org/10.1007/s00405-012-2101-9)] [Medline: [22766835](https://pubmed.ncbi.nlm.nih.gov/22766835/)]
7. Holmes AE, Frank T. Telephone listening ability for hearing-impaired individuals. *Ear Hear* 1984;5(2):96-100. [Medline: [6724177](https://pubmed.ncbi.nlm.nih.gov/6724177/)]
8. Kepler LJ, Terry M, Sweetman RH. Telephone usage in the hearing-impaired population. *Ear Hear* 1992 Oct;13(5):311-319. [Medline: [1487091](https://pubmed.ncbi.nlm.nih.gov/1487091/)]
9. Adams J, Hasenstab MS, Pippin GW, Sismanis A. Telephone use and understanding in patients with cochlear implants. *Ear Nose Throat J* 2004 Feb;83(2):96, 99-100, 102-3. [Medline: [15008443](https://pubmed.ncbi.nlm.nih.gov/15008443/)]
10. Wu C, Liu T, Wang N, Chao W. Speech perception and communication ability over the telephone by Mandarin-speaking children with cochlear implants. *Int J Pediatr Otorhinolaryngol* 2013 Aug;77(8):1295-1302. [doi: [10.1016/j.ijporl.2013.05.015](https://doi.org/10.1016/j.ijporl.2013.05.015)] [Medline: [23768483](https://pubmed.ncbi.nlm.nih.gov/23768483/)]
11. Liu C, Fu QJ, Narayanan SS. Effect of bandwidth extension to telephone speech recognition in cochlear implant users. *J Acoust Soc Am* 2009 Feb;125(2):EL77-EL83 [FREE Full text] [doi: [10.1121/1.3062145](https://doi.org/10.1121/1.3062145)] [Medline: [19206836](https://pubmed.ncbi.nlm.nih.gov/19206836/)]
12. Milchard AJ, Cullington HE. An investigation into the effect of limiting the frequency bandwidth of speech on speech recognition in adult cochlear implant users. *Int J Audiol* 2004 Jun;43(6):356-362. [Medline: [15457818](https://pubmed.ncbi.nlm.nih.gov/15457818/)]
13. Ito J, Nakatake M, Fujita S. Hearing ability by telephone of patients with cochlear implants. *Otolaryngol Head Neck Surg* 1999 Dec;121(6):802-804. [doi: [10.1053/hn.1999.v121.a93864](https://doi.org/10.1053/hn.1999.v121.a93864)] [Medline: [10580241](https://pubmed.ncbi.nlm.nih.gov/10580241/)]
14. Cray JW, Allen RL, Stuart A, Hudson S, Layman E, Givens GD. An investigation of telephone use among cochlear implant recipients. *Am J Audiol* 2004 Dec;13(2):200-212. [Medline: [15903146](https://pubmed.ncbi.nlm.nih.gov/15903146/)]
15. Dalton DS, Cruickshanks KJ, Klein BE, Klein R, Wiley TL, Nondahl DM. The impact of hearing loss on quality of life in older adults. *Gerontologist* 2003 Oct;43(5):661-668. [Medline: [14570962](https://pubmed.ncbi.nlm.nih.gov/14570962/)]
16. Barnett S, Franks P. Deafness and mortality: analyses of linked data from the National Health Interview Survey and National Death Index. *Public Health Rep* 1999;114(4):330-336 [FREE Full text] [Medline: [10501133](https://pubmed.ncbi.nlm.nih.gov/10501133/)]
17. Cacciatore F, Napoli C, Abete P, Marciano E, Triassi M, Rengo F. Quality of life determinants and hearing function in an elderly population: osservatorio geriatrico campano study group. *Gerontology* 1999;45(6):323-328. [Medline: [10559650](https://pubmed.ncbi.nlm.nih.gov/10559650/)]
18. Gates GA, Cobb JL, Linn RT, Rees T, Wolf PA, D'Agostino RB. Central auditory dysfunction, cognitive dysfunction, and dementia in older people. *Arch Otolaryngol Head Neck Surg* 1996 Feb;122(2):161-167. [Medline: [8630210](https://pubmed.ncbi.nlm.nih.gov/8630210/)]
19. Mosnier I, Bebear J, Marx M, Fraysse B, Truy E, Lina-Granade G, et al. Improvement of cognitive function after cochlear implantation in elderly patients. *JAMA Otolaryngol Head Neck Surg* 2015 May 01;141(5):442-450. [doi: [10.1001/jamaoto.2015.129](https://doi.org/10.1001/jamaoto.2015.129)] [Medline: [25763680](https://pubmed.ncbi.nlm.nih.gov/25763680/)]
20. Mantokoudis G, Dubach P, Pfiffner F, Kompis M, Caversaccio M, Senn P. Speech perception benefits of internet versus conventional telephony for hearing-impaired individuals. *J Med Internet Res* 2012 Jul 16;14(4):e102 [FREE Full text] [doi: [10.2196/jmir.1818](https://doi.org/10.2196/jmir.1818)] [Medline: [22805169](https://pubmed.ncbi.nlm.nih.gov/22805169/)]
21. Mantokoudis G, Kompis M, Dubach P, Caversaccio M, Senn P. How internet telephony could improve communication for hearing-impaired individuals. *Otol Neurotol* 2010 Sep;31(7):1014-1021. [doi: [10.1097/MAO.0b013e3181ec1d46](https://doi.org/10.1097/MAO.0b013e3181ec1d46)] [Medline: [20634771](https://pubmed.ncbi.nlm.nih.gov/20634771/)]
22. Mantokoudis G, Dähler C, Dubach P, Kompis M, Caversaccio MD, Senn P. Internet video telephony allows speech reading by deaf individuals and improves speech perception by cochlear implant users. *PLoS One* 2013;8(1):e54770 [FREE Full text] [doi: [10.1371/journal.pone.0054770](https://doi.org/10.1371/journal.pone.0054770)] [Medline: [23359119](https://pubmed.ncbi.nlm.nih.gov/23359119/)]
23. Trond U, Stafnsnes F. VoIP speech quality - better than PSTN? *Elektronikk* 2006;1(119):29.
24. Sun LF, Wade G, Lines BM, Ifeachor EC. Impact of packet loss location on perceived speech quality. 2001 Presented at: 2nd IP-Telephony Workshop; 2001/4/2; New York.

25. Ding L, Goubran RA. Assessment of effects of packet loss on speech quality in VoIP. 2003 Presented at: The 2nd IEEE International Workshop; 2003/09/21; Ottawa, Canada p. 49-54.
26. Rix AW, Beerends JG, Hollier MP, Hekstra AP. Perceptual evaluation of speech quality (PESQ) - a new method for speech quality assessment of telephone networks and codecs. 2001 Presented at: IEEE International Conference on Acoustics, Speech and Signal Processing; 2001/05/07; Salt Lake City, UT p. 749-752.
27. Hochmair-Desoyer I, Schulz E, Moser L, Schmidt M. The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. *Am J Otol* 1997 Nov;18(6 Suppl):S83. [Medline: [9391610](#)]
28. Vos K, Jensen S, Soerensen K. IETF. SILK speech codec URL: <https://tools.ietf.org/html/draft-vos-silk-01> [accessed 2017-02-25] [WebCite Cache ID 6oY3Glp0o]
29. Kim JS, Kim CH. A review of assistive listening device and digital wireless technology for hearing instruments. *Korean J Audiol* 2014 Dec;18(3):105-111 [FREE Full text] [doi: [10.7874/kja.2014.18.3.105](#)] [Medline: [25566400](#)]
30. Litovsky RY, Goupell MJ, Godar S, Grieco-Calub T, Jones GL, Garadat SN, et al. Studies on bilateral cochlear implants at the University of Wisconsin's binaural hearing and speech laboratory. *J Am Acad Audiol* 2012 Jun;23(6):476-494 [FREE Full text] [doi: [10.3766/jaaa.23.6.9](#)] [Medline: [22668767](#)]
31. Iwaki T, Matsushiro N, Mah S, Sato T, Yasuoka E, Yamamoto K, et al. Comparison of speech perception between monaural and binaural hearing in cochlear implant patients. *Acta Otolaryngol* 2004 May;124(4):358-362. [Medline: [15224853](#)]
32. Wolfe J, Morais M, Schafer E. Speech recognition of bimodal cochlear implant recipients using a wireless audio streaming accessory for the telephone. *Otol Neurotol* 2016 Feb;37(2):e20-e25. [doi: [10.1097/MAO.0000000000000903](#)] [Medline: [26756151](#)]
33. Qian H, Loizou PC, Dorman MF. A phone-assistive device based on Bluetooth technology for cochlear implant users. *IEEE Trans Neural Syst Rehabil Eng* 2003 Sep;11(3):282-287. [doi: [10.1109/TNSRE.2003.816871](#)] [Medline: [14518792](#)]
34. Di NW, Anzivino R, Gambini G, De CE, Paludetti G. Improvement of telephone communication in elderly cochlear implant patients. *Audiol Neurotol* 2014;19 Suppl 1:27-32 [FREE Full text] [doi: [10.1159/000371605](#)] [Medline: [25733363](#)]
35. Wolfe J, Morais DM, Schafer E, Cire G, Menapace C, O'Neill L. Evaluation of a wireless audio streaming accessory to improve mobile telephone performance of cochlear implant users. *Int J Audiol* 2016;55(2):75-82. [doi: [10.3109/14992027.2015.1095359](#)] [Medline: [26681229](#)]
36. Lin FR, Yaffe K, Xia J, Xue Q, Harris TB, Purchase-Helzner E, Health ABC Study Group. Hearing loss and cognitive decline in older adults. *JAMA Intern Med* 2013 Feb 25;173(4):293-299 [FREE Full text] [doi: [10.1001/jamainternmed.2013.1868](#)] [Medline: [23337978](#)]
37. Wayne R, Johnsrude IS. A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res Rev* 2015 Sep;23(Pt B):154-166. [doi: [10.1016/j.arr.2015.06.002](#)] [Medline: [26123097](#)]

Abbreviations

CI: cochlear implant
DECT: digital enhanced cordless telecommunications
IQR: interquartile range
ITU: International Telecommunication Union
MOS: mean opinion score
PESQ: perceptual evaluation of speech quality
PL: packet loss
PSTN: public switched telephone networks
SNR: signal-to-noise ratio
SPL: sound pressure level
VoIP: voice over Internet protocol

Edited by G Eysenbach; submitted 06.11.16; peer-reviewed by N Dillier, C Stieger, A Paglialonga, M Masalski; comments to author 15.12.16; revised version received 28.01.17; accepted 02.03.17; published 24.04.17.

Please cite as:

Mantokoudis G, Koller R, Guignard J, Caversaccio M, Kompis M, Senn P
Influence of Telecommunication Modality, Internet Transmission Quality, and Accessories on Speech Perception in Cochlear Implant Users
J Med Internet Res 2017;19(4):e135
 URL: <http://www.jmir.org/2017/4/e135/>
 doi: [10.2196/jmir.6954](#)
 PMID: [28438727](#)

©Georgios Mantokoudis, Roger Koller, Jérémie Guignard, Marco Caversaccio, Martin Kompis, Pascal Senn. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 24.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Importance-Performance Analysis of Personal Health Records in Taiwan: A Web-Based Survey

Hsiao-Hsien Rau¹, PhD; Yi-Syuan Wu^{2,3*}, MS; Chi-Ming Chu^{4*}, PhD; Fu-Chung Wang^{5*}, PhD; Min-Huei Hsu^{1*}, PhD, MD; Chi-Wen Chang^{6,7*}, PhD; Kang-Hua Chen⁸, PhD; Yen-Liang Lee^{1,9*}, PhD; Senyeong Kao^{4*}, PhD; Yu-Lung Chiu^{4*}, PhD; Hsyien-Chia Wen^{10*}, PhD; Anis Fuad^{11*}, MSc; Chien-Yeh Hsu^{1,12,13*}, PhD; Hung-Wen Chiu^{1*}, PhD

¹Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei, Taiwan

²Administration Department, Healthconn Corp, Taipei, Taiwan

³Institute of Medical Informatics, Department of Computer Science of Information Engineering, National Cheng Kung University, Tainan, Taiwan

⁴School Of Public Health, National Defence Medical Center, Taipei, Taiwan

⁵National Health Insurance Administration, Ministry of Health and Welfare, Taipei, Taiwan

⁶School of Nursing, College of Medicine, Chang Gung University, Linkou, Taiwan

⁷Division of Endocrinology, Department of Pediatrics, Linkou Chang Gung Memorial Hospital, Linkou, Taiwan

⁸School of Nursing, College of Medicine, Chang Gung University, Taipei, Taiwan

⁹Internet of Things Laboratory, Chunghwa Telecom Laboratories, TaoYuen, Taiwan

¹⁰School of Health Care Administration, Taipei Medical University, Taipei, Taiwan

¹¹Department of Biostatistics, Epidemiology and Population Health, Faculty of Medicine, Universitas Gadjah Mada, Yogyakarta, Indonesia

¹²Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

¹³Master Program in Global Health and Development, Taipei Medical University, Taipei, Taiwan

*these authors contributed equally

Corresponding Author:

Hung-Wen Chiu, PhD

Graduate Institute of Biomedical Informatics

Taipei Medical University

15F, No.172-3, Sec 2, Keelung Rd, Da'an Dist

Taipei, 106

Taiwan

Phone: 886 266382736 ext 1511

Fax: 886 266382736

Email: hwchiu@tmu.edu.tw

Related Article:

This is a corrected version. See correction statement: <http://www.jmir.org/2017/6/e209/>

Abstract

Background: Empowering personal health records (PHRs) provides basic human right, awareness, and intention for health promotion. As health care delivery changes toward patient-centered services, PHRs become an indispensable platform for consumers and providers. Recently, the government introduced “My health bank,” a Web-based electronic medical records (EMRs) repository for consumers. However, it is not yet a PHR. To date, we do not have a platform that can let patients manage their own PHR.

Objective: This study creates a vision of a value-added platform for personal health data analysis and manages their health record based on the contents of the “My health bank.” This study aimed to examine consumer expectation regarding PHR, using the importance-performance analysis. The purpose of this study was to explore consumer perception regarding this type of a platform: it would try to identify the key success factors and important aspects by using the importance-performance analysis, and give some suggestions for future development based on it.

Methods: This is a cross-sectional study conducted in Taiwan. Web-based invitation to participate in this study was distributed through Facebook. Respondents were asked to watch an introductory movie regarding PHR before filling in the questionnaire. The questionnaire was focused on 2 aspects, including (1) system functions, and (2) system design and security and privacy. The questionnaire would employ 12 and 7 questions respectively. The questionnaire was designed following 5-points Likert scale ranging from 1 (“disagree strongly”) to 5 (“Agree strongly”). Afterwards, the questionnaire data was sorted using IBM SPSS Statistics 21 for descriptive statistics and the importance-performance analysis.

Results: This research received 350 valid questionnaires. Most respondents were female (219 of 350 participants, 62.6%), 21-30 years old (238 of 350 participants, 68.0%), with a university degree (228 of 350 participants, 65.1%). They were still students (195 out of 350 participants, 56.6%), with a monthly income of less than NT \$30,000 (230 of 350 participants, 65.7%), and living in the North Taiwan (236 of 350 participants, 67.4%), with a good self-identified health status (171 of 350 participants, 48.9%). After performing the importance-performance analysis, we found the following: (1) instead of complex functions, people just want to have a platform that can let them integrate and manage their medical visit, health examination, and life behavior records; (2) they do not care whether their PHR is shared with others; and (3) most of the participants think the system security design is not important, but they also do not feel satisfied with the current security design.

Conclusions: Overall, the issues receiving the most user attention were the system functions, circulation, integrity, ease of use, and continuity of the PHRs, data security, and privacy protection.

(*J Med Internet Res* 2017;19(4):e131) doi:[10.2196/jmir.7065](https://doi.org/10.2196/jmir.7065)

KEYWORDS

electronic health records; health records, personal; surveys and questionnaires

Introduction

A personal health record (PHR) is an electronic record of health-related information on an individual that conforms to nationally recognized interoperability standards, and that can be drawn from multiple sources while being managed, shared, and controlled by the individual [1]. PHR includes medical records, lab results, physical assessments, medical history, medication history, physiologic measurements, dietary records, exercise records, and so on. These data can be updated by the user, measured by the measurement equipment automatically uploaded, or by hospitals or clinics that allow data import [2]. The purpose of PHR is to integrate patient health information from a variety of sources, including all patient records, and allow authorized persons to access these records any time and at any place.

PHR is that part of the electronic medical records (EMRs) or Electronic Health Records (EHR) that an individual “owns” and controls. In Taiwan, hospitals adopted EMR in 2004. It was related with the introduction of a basic format of EMR. In 2007, the National Health Informatics Project (NHIP) was promoted to implement the infrastructure of health information by the government to prepare the EMR exchange, and encourage hospitals to use EMR. In 2008, clinical document architecture (CDA) was adopted by the government to create 108 basic formats of EMR. The government of Taiwan also provided incentives to hospitals that adopted EMR. Until 2015, 406 hospitals (90% of Taiwan’s hospitals) already earned rewards and could exchange EMR with each other.

The research on PHR has increased. In the past, EMRs were stored in a large database; different medical service providers in the hospital could access the medical records in the database, but transfer between different hospitals was a problem. Thus, the implementation of PHRs allowed patients to achieve the integration of their PHRs and medical records [3]. Traditionally,

most people use paper to write their records, whereas PHRs allow people to easily record and maintain their own health information [4]. With the increasing number of people using EMRs, PHRs were also managed and authorized to be shared through a network in some foreign countries. This approach was found to improve the medical satisfaction of performance, and reduce the cost of medical care with better quality [5].

Some studies listed the benefits of using PHR, including reducing health care costs, improving personal health outcomes, and improving the experience of care for patients and their families [6-10].

The current mechanism of medical record exchange in an EMR exchange center plays a role of a personal EMR platform. Through the Exchange Center, a medical record previously created in any hospital can be accessed from another hospital to reduce duplicate examinations, accelerate diagnosis and treatment, and reduce the medical resource burden. The medication record in the EMR is very conducive to decision-making by doctors to avoid drug allergies and other medical disputes. The importance of a personal EMR platform to a medical institution and its members is like the importance of the clients’ information to an enterprise, which is the basis for providing good service.

Although EMR is already being used well, the medical records are still owned by hospitals; people cannot own and manage their health care record. Therefore, the Ministry of Health and Welfare introduced the concept of “My health bank” in 2015 to let patients own their PHR. In the last three years, this initiative permitted people to download the integrated medical record from the Internet, which includes outpatient and hospital records, diagnosis, drug use records, cost, laboratory test and health examination report, allergies, and so on. It lets people view and manage their health care record at any time and from any place. However, “My health bank” still has some problems. The information it includes is a general report with no detailed

content; it doesn't have medical images, the content is still not standardized, and people can not add their own data regarding diet, exercise, nutrition records, or the health examination report from another examination center.

We still do not have a platform that can let patient manage their own PHR. This study creates a vision of a value-added platform for personal health data analysis and management of health records on the basis of the contents of the "My health bank"; patients can retrieve their health records and medical records through the personal EMR platform to manage their own health conditions. The availability of medical records can contribute to the transparency of the medical records and facilitate immediate access to the medical contents, thereby allowing discussions of disease conditions with the patients' relatives and friends and related medical personnel. Thus, the ownership of the medical record is reverted to the patient.

The National Committee on Vital and Health Statistics found that a critical success factor for PHRs is the provision of software tools that help patients manage their own health conditions [11].

This research gives a vision of the electronic PHR management platform and uses importance-performance analysis (IPA) to identify the important factors from patients' perspectives. The purpose of this study was to explore consumer perception regarding this kind of platform, and try to both identify the key success factors and important aspects using

importance-performance analysis, and give some suggestions for future development based on the findings.

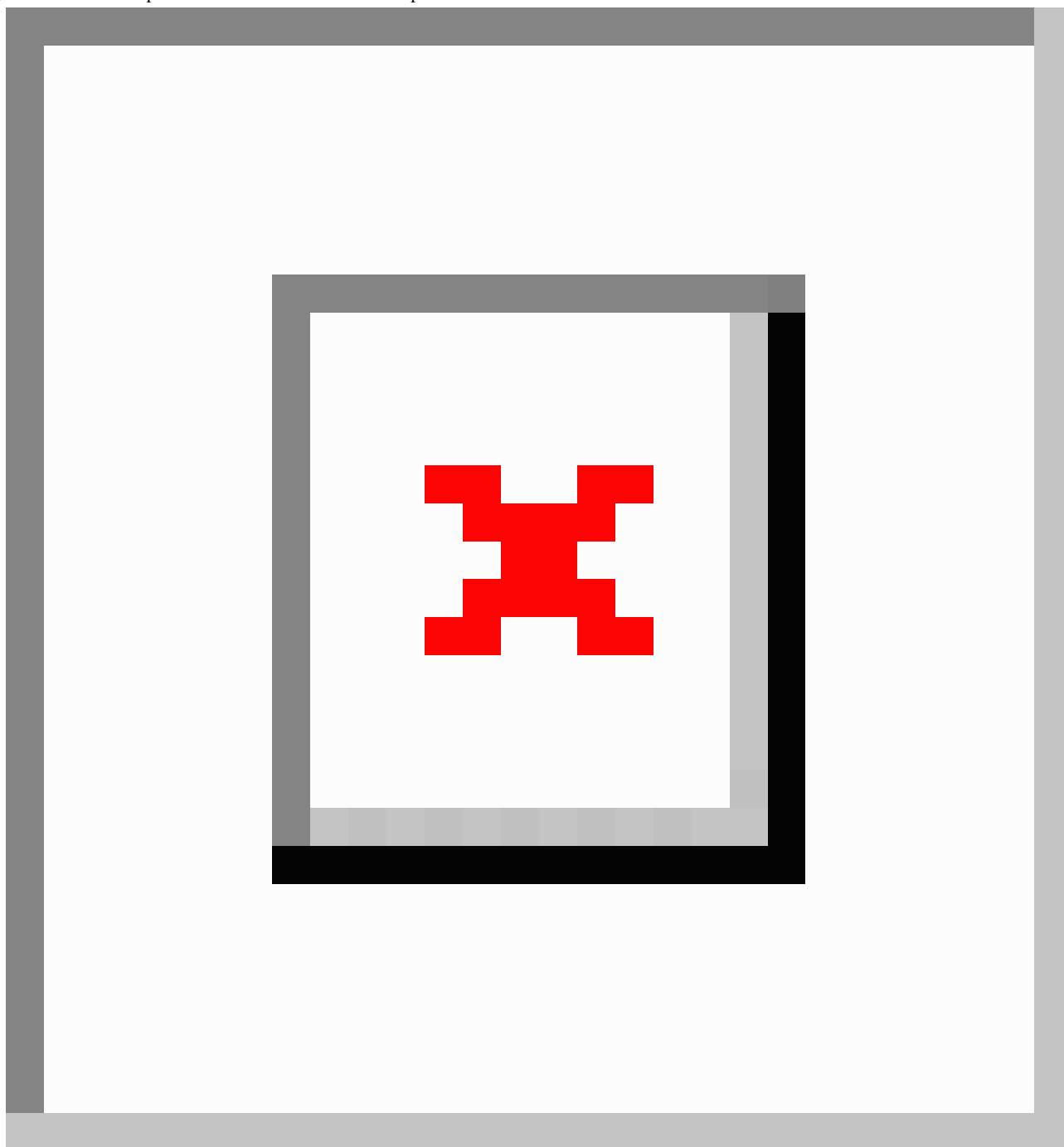
Methods

Study Design

In this study, we used a cross-sectional study design. We let the participants watch a Web-based video on the vision of personal EMR platform. After the video introduction, we asked them to fill a structured questionnaire on the Internet (created by Google Form) for quantization of their acceptance and importance of the personal EMR platform's functions and security.

In this study, we developed a Web-based video simulation describing the scenario of the personal EMR platform's operation, as shown in [Figure 1](#). This 2:17-min video presented the functions of the personal EMR platform using a case of a patient seeking medical treatment in hospital A. Through the personal EMR platform, after returning home, the patient could use or manage in a secure and private home environment their continuous and integrated PHR, which include personal general information, their health examination record, hospital visit record, PHR, and so on. They can also let doctors in hospital B, family members, and insurance companies use these files through the sharing mechanism when needed.

After viewing the video, the participants were asked to fill out a Web-based questionnaire in Google Form regarding the importance and satisfaction of the performance of the personal EMR platform.

Figure 1. Scenario of personal electronic medical record platform.

Questionnaire Design

In order to explore their impact on public acceptance of the personal EMR platform and its importance, this study utilized a questionnaire on importance-performance analysis [12]. The original version of this questionnaire has been used for marketing research to understand consumer perception regarding satisfaction and importance of a product.

We modified this questionnaire into 2 main constructs including (1) system functions and (2) system design and security and

privacy, with 12 and 7 questions, respectively. The questionnaire was designed following 5-point Likert scale with the scale ranging from 1 (“disagree strongly”) to 5 (“Agree strongly”).

Before the study, this study invited 4 experts to do expert validity and was pretested by 30 patients. No major problems emerged during this pretest. As showed in Table 1, the questionnaire is highly reliable. The Cronbach alpha of these 2 parts were both more than .9. It means that this questionnaire has a good consistency, and that the subjects could understand the question and fill it in clearly.

Table 1. Questionnaire design and reliability analysis.

Aspect	Definition	Questions	Cronbach alpha
System functions	People's opinion of satisfaction and importance of personal EMR ^a platform's system function	12	.936
System design and security	People's opinion of satisfaction and importance of personal EMR platform's system design and security	7	.917

^aEMR: electronic medical record.

Sampling and Exclusion Criteria

This study used random sampling of the Internet users. We put the video on the YouTube platform from 31st March 2014 to 9th April and publicized it through email and Facebook. After participants finished watching this video, they were asked to fill the questionnaire (Web-based questionnaire in Google sheet format).

The sample size was calculated by using the Magnani [13] formula. It was calculated by reliability coefficient, population proportion (parameter), and width (CI) / margin of error. After calculation, we got the minimum sample size of 323, which meant we would have at least 323 valid respondents after the announcement for 10 days. In total, 614 people have watched this video and 400 people finished the questionnaire survey (response rate: 65.1%); after excluding incomplete and repeated questionnaires, we got 350 valid respondents, which is more than the minimum required case number.

Data Analysis

This study used Excel (Microsoft) and SPSS (IBM Corp) as statistic tools for analysis. We used descriptive statistics to

calculate the mean, standard deviation, median, frequency distribution, and percentage statistics to ascertain the data distribution in system functions, and system design and security. The objective was to examine the relationship between characteristics of participants, as well as all of the aspects above, and set 0.05 as the significance level; if $P < .05$, then it is statistically significant.

We also performed IPA to measure the participants' attitudes toward the personal EMR platform's functions. We calculated the mean of every factor in "system functions" and "system design and security" aspects, and we put them in a quadrant diagram.

Results

Respondent Characteristics

This research gathered 350 valid questionnaires. As shown in Table 2, most respondents were female ($n=219$), 21-30 years old ($n=238$), with a university degree ($n=228$). Or they were still students ($n=195$), with a monthly income of less than NT \$30,000 ($n=230$), and lived in the north area ($n=236$), with a good self-identified health status ($n=171$).

Table 2. Characteristics of respondents.

Characteristics	n (%)
Gender	
Male	131 (37.4)
Female	219 (62.6)
Age (years)	
<20	46 (13.1)
21-30	238 (68.0)
31-40	56 (16.0)
>40	10 (2.9)
Highest education level	
High school degree	8 (2.3)
University degree	228 (65.1)
Institute or above	114 (32.6)
Employment	
Students	195 (55.7)
Services	61 (17.4)
Manufacturing	29 (8.3)
Financial industry	7 (2.0)
Military and police education	47 (13.4)
Unemployed	11 (3.1)
Monthly income (NT \$)	
<30,000	230 (65.7)
30,001~50,000	94 (26.9)
50,001~70,000	21 (6.0)
>70,001	5 (1.5)
Living area	
Northern Taiwan	236 (67.4)
Central Taiwan	49 (14.0)
Southern Taiwan	55 (15.7)
Eastern Taiwan	7 (2.0)
offshore islands	3 (0.9)
Health status	
Excellent	28 (8.0)
Good	171 (48.9)
Normal	132 (37.7)
Poor	19 (5.4)

Importance-Performance Analysis

The Importance-performance Analysis (IPA) framework was introduced by Martilla and James [12]. It uses the mean of importance and satisfaction of performance of all items, and employs the intersection point as the origin to draw a quadrant diagram. As defined in quadrant 1 of Figure 2, it indicates high importance and high satisfaction of performance, which

indicates that existing systems have strengths and should be maintained. Items in quadrant 2 have high importance but low satisfaction of performance. This quadrant is labeled as “Concentrate Here.” This category is labeled as “Keep up the good work.” In quadrant 3, low importance and low satisfaction of performance items renders it as “Low Priority.” Finally, quadrant 4 represents low importance and high satisfaction of performance, which suggests insignificant strengths and a

possibility that the resources invested may be better used elsewhere.

In this study, the questionnaire was divided into “system functions” and “system design and security,” the two aspects that evaluate the relationship between importance and satisfaction of performance. We focused on the key area for improvement in the second quadrant of the IPA to identify the services showing high importance and low satisfaction of performance to determine the improvement priorities. This analysis is expected to provide direction for future enhancement and implementation of the personal EMR platform as a reference for government agencies and system developers.

In the system function part, as shown in Table 3, F7 (can maintain and keep complete personal health examination record) and F10 (can maintain and keep complete personal medical record) both have the highest satisfaction of performance. They are followed by F6 (can maintain and keep complete personal health record), F3 (can access PHR from other hospitals to avoid duplication of examinations, tests, and medication), and F8 (can maintain and keep complete personal medical image). The lowest satisfaction of performance is F1 (have detailed operating instructions). F3 has the highest importance, followed by F10 and F6.

Table 3. Satisfaction of performance and importance with the order of every question in system function part (order 1: most important or satisfied, 12: least important or satisfied).

Question	Satisfaction of performance, mean (SD)	Order	Importance, mean (SD)	Order	Quadrant	Chuchiming Index ^b
F1. Personal EMR ^a platform have detailed operating instructions	3.42 (0.92)	12	4.19 (0.78)	7	II	0.71
F2. Personal EMR platform lets patients integrate existing paper-based medical history	3.66 (0.83)	10	4.15 (0.73)	9	III	0.11
F3. Personal EMR platform lets patients access their own PHR from other hospitals to avoid duplication of examinations, tests, and medication	3.89 (0.89)	4	4.38 (0.69)	1	I	3.00
F4. Personal EMR platform lets patients integrate their own PHR data and provide continuous numerical statistics	3.72 (0.88)	8	4.20 (0.74)	6	II	0.33
F5. Personal EMR platform allow doctors to add more details on the medical records	3.77 (0.92)	7	4.22 (0.77)	5	I	0.40
F6. Personal EMR platform lets patients maintain and keep complete personal “health record” (such as disease history, medication history, and blood pressure)	3.93 (0.88)	3	4.31 (0.72)	3	I	0.00
F7. Personal EMR platform lets patients maintain and keep complete personal “health examination record” (such as blood test and urine test reports)	3.95 (0.86)	1	4.26 (0.72)	4	I	−0.75
F8. Personal EMR platform lets patients maintain and keep complete personal “medical image” (such as X-ray and MRI)	3.87 (0.86)	5	4.19 (0.76)	7	I	−0.29
F9. Personal EMR platform lets patients maintain and keep complete personal “endoscopic image”	3.80 (0.89)	6	4.14 (0.78)	10	IV	−0.40
F10. Personal EMR platform lets patients maintain and keep complete personal medical record (such as diagnosis and prescriptions)	3.95 (0.86)	1	4.34 (0.73)	2	I	−0.50
F11. Personal EMR platform lets patients share their PHR with family (friends) to enable them understand their health condition	3.53 (1.01)	11	3.74 (1.02)	12	III	−0.08
F12. Personal EMR platform lets patients share their medical records with another physician as a reference when diagnosis is carried out	3.72 (0.91)	8	4.08 (0.84)	11	III	−0.27

^aEMR: electronic medical record.

^bindex developed by Dr Chuchiming. In this study, Chuchiming index>0 indicates items need concerted improvement (perceiving targets), Chuchiming index<0 indicates resources can be drawn from items (shifting resources), Chuchiming index=0 indicates items can fit people’s expectation (balancing items).

Figure 2 shows the quadrant distribution of the importance and satisfaction of performance of every system function of the personal EMR platform. The first quadrant (high importance and high satisfaction of performance) has F3, F5 (allows doctor to add more details on the medical records), F6, F7, F8, and F10. The items in the second quadrant (high importance and low satisfaction of performance) are F1 and F4 (can integrate their own PHR data and provide continuous numerical statistics).

The items in the third quadrant (low importance and low satisfaction of performance) are F2 (existing paper-based medical history can be integrated), F11 (can be shared PHR with family), and F12 (can share PHR with physicians). The item in the fourth quadrant (low importance and high satisfaction of performance) is F9 (can maintain and keep complete personal endoscopic images).

In [Figure 2](#), there are 4 functions with below average importance, including “sharing PHR with family and physician” (F11, F12), “integrating existing paper-based medical history” (F2), and “maintaining and keeping complete personal endoscopic image” (F9).

It also has 5 functions with below average satisfaction of performance, including “detailed personal EMR platform operating instructions” (F1), “integrating existing paper-based medical history” (F2), “sharing PHR with family or physician” (F11, F12), and “integrating their own PHR data and provide continuous numerical statistics” (F4).

It means a paperless process or sharing with others is not seen as important functions by the participants. They just want to have a platform that can let them maintain and keep complete personal records and basic medical images.

Table 4. Satisfaction of performance and importance with their order of every question in system design and security part (order 1: most important or satisfied, 7: least important or satisfied).

Question	Satisfaction of performance, mean (SD)	Order	Importance, mean (SD)	Order	Quadrant	Chuchiming Index ^c
S1. The interface should be simple and easy to understand	3.82 (0.91)	3	4.28 (0.76)	4	IV	-0.25
S2. Personal EMR ^a platform lets patients access their own PHR ^b quickly	3.91 (0.86)	1	4.25 (0.71)	7	IV	-0.86
S3. Personal EMR platform lets patients login by multiple methods such as citizen digital certificate, id, and password.	3.87 (0.92)	2	4.39 (0.86)	1	I	1.00
S4. Personal EMR platform lets patients access their own PHR under a secure environment	3.78 (0.96)	4	4.34 (0.76)	2	II	1.00
S5. Personal EMR platform lets patients login by id and password	3.73 (0.85)	7	4.27 (0.82)	5	III	0.40
S6. Personal EMR platform lets patients login by “Citizen Digital Certificate”	3.76 (0.95)	6	4.26 (0.80)	6	III	0.00
S7. Personal EMR platform lets patients set access rights for every physician, family, or friend.	3.78 (0.92)	5	4.29 (0.76)	3	III	0.67

^aEMR: electronic medical record.

^bPHR: personal health record.

^cindex developed by Dr Chuchiming. In this study, Chuchiming index>0 indicates items need concentrated improving (perceiving targets), Chuchiming index<0 indicates resources can be drawn from items (shifting resources), Chuchiming index=0 indicates items can fit people’s expectation (balancing items).

[Figure 3](#) shows the quadrant distribution of the importance and satisfaction with performance of the system design and security aspects of the Personal EMR Platform. The item in the first quadrant (high importance and high satisfaction with performance) is S3. The item in the second quadrant (high importance and low satisfaction with performance) is S4. The items in the third quadrant (low importance and low satisfaction with performance) are S5, S6 (lets patients log in using a citizen digital certificate) and S7. The fourth quadrant (low importance and high satisfaction with performance) includes S1 and S2.

In [Figure 3](#), we also can find most of the system design and security below average in importance, besides letting patients log in by multiple methods (S3) and letting patients access their own PHR under a secure environment (S4); however, even the next personal EMR platform has been designed with multiple security protection methods. Participants still feel concerned

After conducting the Analysis of Variance (ANOVA) test, we didn’t find any significant difference between the various age groups, annual income groups, and health status groups, meaning that the results described above are consistent in every group.

In the system design and security part, as shown in [Table 4](#), S2 (can let patients access their own PHR quickly) have highest satisfaction of performance and followed by S3 (lets patients login by multiple methods), S1 (interface should be simple and easy to understand) and S4 (can let patients access their own PHR under a secure environment). The lowest satisfaction of performance is S5 (let patients can login by id and password). S3 have highest importance, followed by S4 and S7 (can let patients set access right).

about the system security issue; the satisfaction of performance of most questions is low, besides S1, S2, and S3. Thus, system security is still a big issue when building this kind of platform in the future.

After conducting the ANOVA test, we found some differences in system design and security part in terms of the importance attributed to system design and security part, as shown in [Table 5](#). The mean importance of the age group of less than 20 years is lower than the values of other groups ($P=.05$). It means this group is not very concerned about the security issue. Also, in satisfaction with performance of system security (according to a group’s monthly income and health status), the group of monthly income between NT \$50K and NT \$70K have lowest satisfaction with performance ($P=.02$). This was also found in the group which thinks they are poor health participants. ($P=.002$).

Table 5. Differences found in analysis of importance and satisfaction of performance between every demographic class in system design and security part by analysis of variance (ANOVA) test.

Attitude	Category	N	Mean (SD)	F	P value
Importance	Age (years)				
	<20	46	4.034 (0.657)	2.586	.05
	21-30	238	4.315 (0.637)		
	31-40	56	4.258 (0.616)		
	>40	10	4.371 (0.629)		
	Highest education level				
	High school	8	4.161(0.442)	0.504	.60
	University degree	228	4.251(0.640)		
	Institute or above	114	4.316(0.656)		
	Employment				
	Students	195	4.315(0.624)	0.734	.60
	Services	61	4.253(0.661)		
	Manufacturing	29	4.118(0.615)		
	Financial industry	7	4.122(0.550)		
	Military and police education	47	0.204(0.720)		
	Unemployed	11	4.351(0.611)		
	Monthly income (NT \$)				
	<30,000	230	4.276(0.634)	0.270	.76
	30,001~50,000	94	4.280(0.615)		
	50,001~70,000	26	4.181(0.727)		
	Living area				
	Northern Taiwan	236	4.308(0.636)	1.510	.20
	Central Taiwan	49	4.262(0.629)		
	Southern Taiwan	55	4.171(0.648)		
	Eastern Taiwan	7	3.796(0.552)		
	offshore islands	3	4.333(1.033)		
	Health status				
	Poor	19	4.180(0.789)	0.349	.79
	Normal	132	4.260(0.664)		
	Good	171	4.272(0.615)		
	Excellent	28	4.367(0.594)		
Satisfaction of performance	Age (years)				
	<20	46	3.637(0.739)	1.533	.20
	21-30	238	3.867(0.773)		
	31-40	56	3.737(0.794)		
	>40	10	3.614(1.149)		
	Highest education level				
	High school	8	3.821(0.457)	1.097	.34
	University degree	228	3.853(0.766)		

Attitude	Category	N	Mean (SD)	F	P value
	Institute or above	114	3.719(0.839)		
	Employment				
	Students	195	3.834(0.762)		
	Services	61	3.763(0.872)		
	Manufacturing	29	3.828(0.732)		
	Financial industry	7	3.918(0.605)		
	Military and police education	47	3.699(0.869)		
	Unemployed	11	3.948(0.652)		
	Monthly income (NT \$)				
	<30,000	230	3.822(0.763)	4.074	.02
	30,001~50,000	94	3.888(0.786)		
	50,001~70,000	26	3.401(0.889)		
	Living area				
	Northern Taiwan	236	3.810(0.733)	0.962	.43
	Central Taiwan	49	3.843(0.764)		
	Southern Taiwan	55	3.813(0.765)		
	Eastern Taiwan	7	3.306(0.871)		
	offshore islands	3	4.238(0.719)		
	Health status				
	Poor	19	3.143 (0.838)	5.077	.002
	Normal	132	3.874 (0.723)		
	Good	171	3.825 (0.774)		
	Excellent	28	3.852 (0.934)		

Table 6 shows the result of the ANOVA test of importance and satisfaction regarding performance between every demographic class in system functions. In this part, only the group with the monthly income between NT \$50k and NT \$70k, the mean

satisfaction of performance, is lower than that of others ($P=.01$). This means income has a strong correlation to the satisfaction with performance of system functions.

Table 6. Differences found in analysis of importance and satisfaction with performance between every demographic class in system functions part by analysis of variance (ANOVA) test.

Attitude	Category	N	Mean (SD)	F	P value
Importance	Age (years)				
	<20	46	3.984(0.677)	2.406	.07
	21-30	238	4.209(0.588)		
	31-40	56	4.198(0.536)		
	>40	10	4.417(0.626)		
	Highest education level				
	High school	8	4.208(0.396)	1.204	.30
	University degree	228	4.148(0.607)		
	Institute or above	114	4.254(0.588)		
	Employment				
	Students	195	4.233(0.594)	1.565	.17
	Services	61	4.111(0.636)		
	Manufacturing	29	3.994(0.440)		
	Financial industry	7	3.845(0.667)		
	Military and police education	47	4.234(0.632)		
	Unemployed	11	4.205(0.503)		
	Monthly income (NT \$)				
	<30,000	230	4.198(0.610)	0.387	.68
	30,001~50,000	94	4.230(0.560)		
	50,001~70,000	26	4.154(0.635)		
	Living area				
	Northern Taiwan	236	4.215(0.596)	1.037	.27
	Central Taiwan	49	4.197(0.528)		
	Southern Taiwan	55	4.089(0.594)		
	Eastern Taiwan	7	3.786(0.829)		
	offshore islands	3	4.111(1.197)		
	Health status				
	Poor	19	4.154(0.709)	0.267	.85
	Normal	132	4.152(0.605)		
	Good	171	4.204(0.594)		
	Excellent	28	4.232(0.521)		
Satisfaction of performance	Age (years)				
	<20	46	3.601(0.699)	1.282	.28
	21-30	238	3.815(0.717)		
	31-40	56	3.702(0.757)		
	>40	10	3.725(1.069)		
	Highest education level				
	High school	8	3.844(0.649)	0.522	.59
	University degree	228	3.792(0.713)		

Attitude	Category	N	Mean (SD)	F	P value
	Institute or above	114	3.710(0.780)		
	Employment				
	Students	195	3.788(0.710)	0.734	.60
	Services	61	3.745(0.820)		
	Manufacturing	29	3.690(0.600)		
	Financial industry	7	3.738(0.598)		
	Military and police education	47	3.722(0.845)		
	Unemployed	11	3.924(0.619)		
	Monthly income (NT \$)				
	<30,000	230	3.747(0.702)	4.366	.01
	30,001~50,000	94	3.904(0.754)		
	50,001~70,000	26	3.442(0.831)		
	Living area				
	Northern Taiwan	236	3.757(0.765)	0.526	.72
	Central Taiwan	49	3.847(0.610)		
	Southern Taiwan	55	3.759(0.679)		
	Eastern Taiwan	7	3.476(0.830)		
	offshore islands	3	4.208(1.018)		
	Health status				
	Poor	19	4.180(0.789)	0.349	.79
	Normal	132	4.260(0.664)		
	Good	171	4.272(0.615)		
	Excellent	28	4.367(0.594)		

Figure 2. Quadrant diagram of the importance and performance of product establishment and functionality.

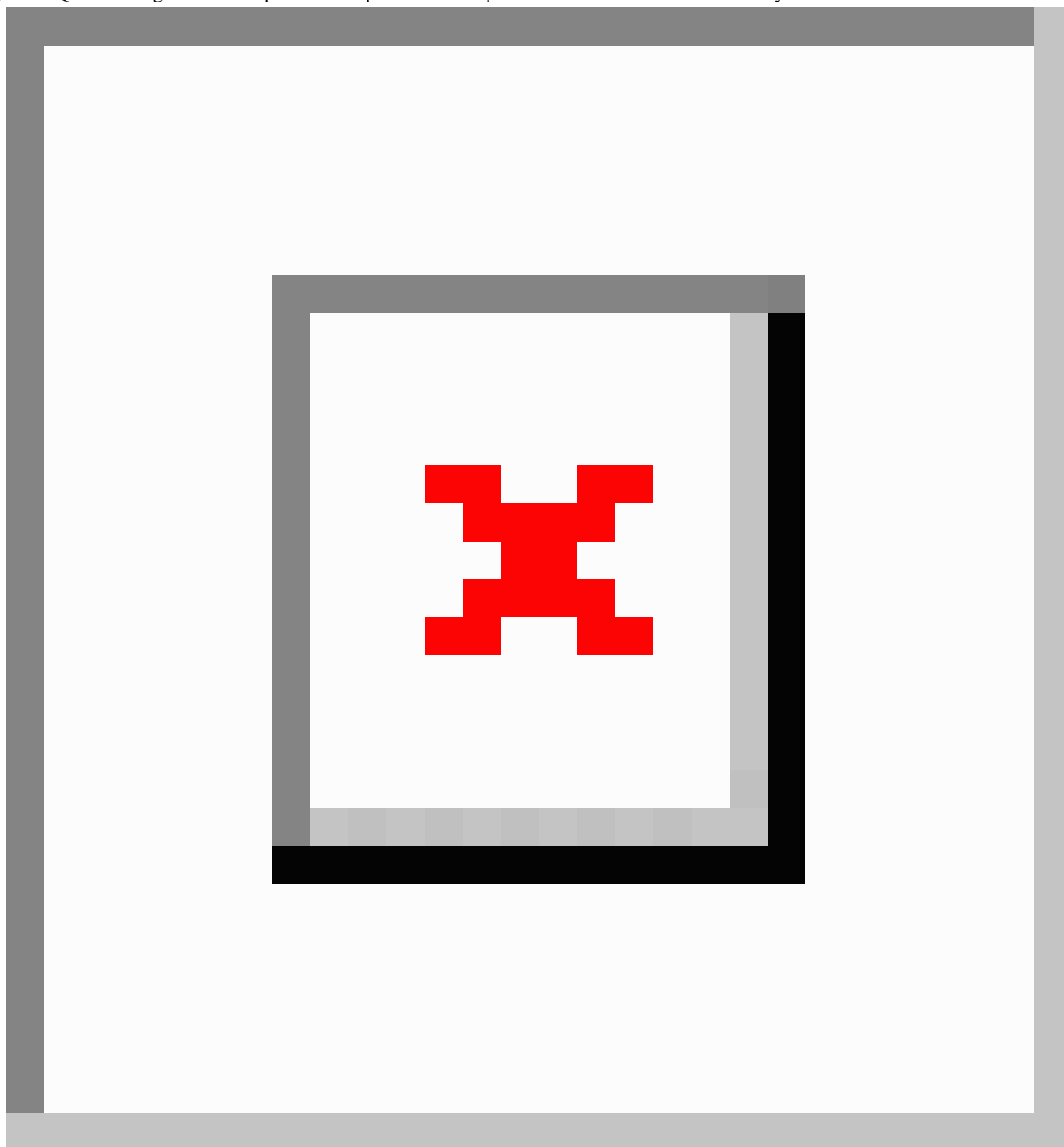
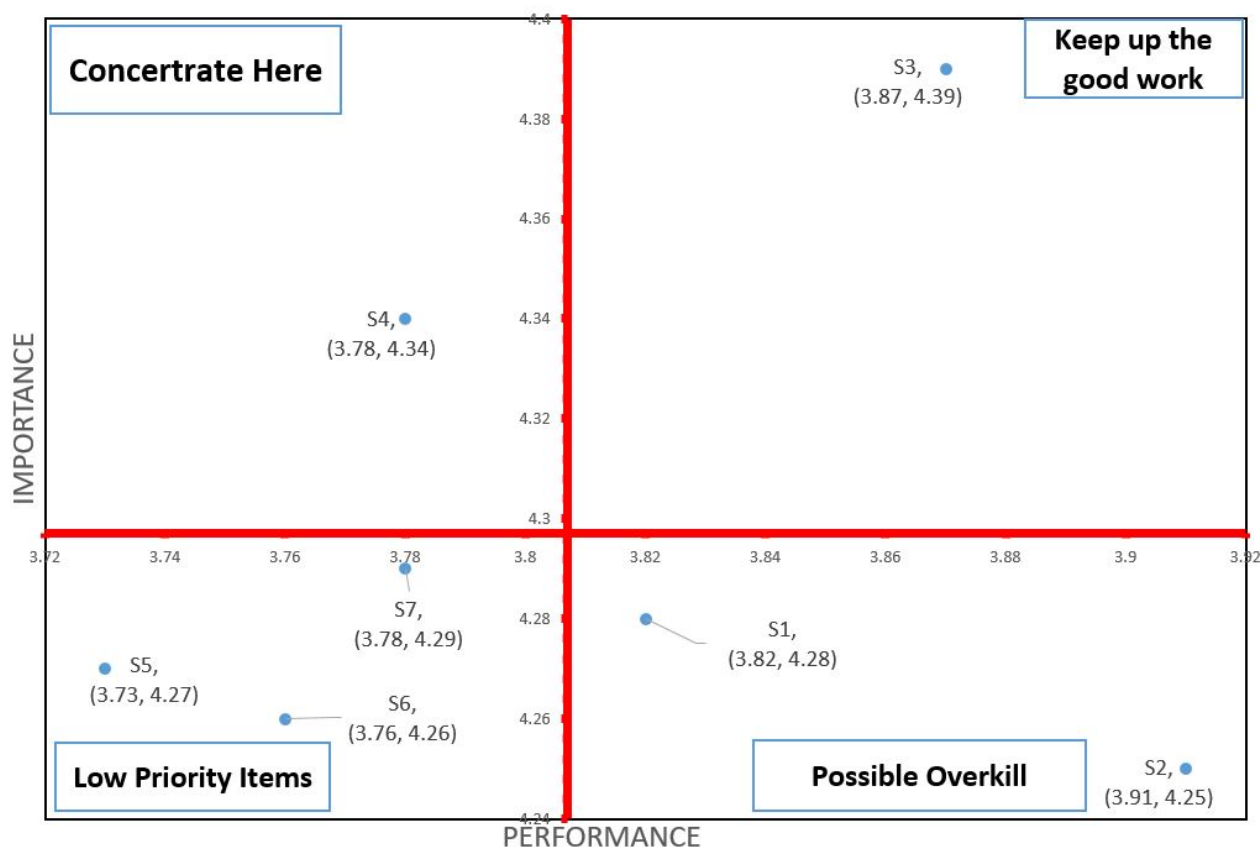


Figure 3. Quadrant diagram of the importance and performance of the security of the system.

Discussion

Principal Findings

Liang's [14] awareness, want, and adoption gap ratio (AWAG) segment matrix that analyzes the digital divide in eHealth services divides into four groups: strong opened group, generic opened group, generic closed group, and strong closed group by the participants' awareness and want rate. Results showed that for both eHealth services, the digital divides of awareness, want, and adoption existed across demographic variables, as well as between computer owners and nonowners, and between Internet users and nonusers. It means users' attitudes and experience are the most important factors determining whether and how eHealth services will be used.

This study focused on system functions and system design and security to identify the key functions of patients' views by using IPA to determine which functions people think are important for them. The study can serve as a guide when this kind of platform is built in the future. We found people do not need overly complex functions; they just want a platform that lets them integrate and manage their medical visits, health examination, and life behavior record. However, viewing the Web-based medical image is not an important function for them, and they also do not care if their PHR is shared with others. Surprisingly, most of the participants think the system security design is not important. The importance of 5 of 7 questions in total is below average (4.29, 1: most not important, 5: most important); they only think "letting patients login by multiple methods, such as citizen digital certificate, id and password"

and "letting patients access their own PHR under a secure environment" is important. Satisfaction in performance is also not high. The significance of only 3 of 7 questions is higher than average (3.8). We also found that the satisfaction regarding performance is of lower importance not only in system functions aspects but also in system design and security aspects. This means that although they think some of the system functions and system security design are not very important, they are not satisfied even when the system has these functions. However, although future personal EMR platform could have many system security designs, people may still not think these are enough to make sure the system is secure.

With increased acceptance of PHR, users increasingly believe that the use of a personalized EMR could help them understand their own medical records. Research on PHR became more popular in recent years; however, most of the research was focused on investigating the use of doctors and nurses' satisfaction. Rarely were the patients' views given consideration. Some studies have also pointed out that if the patients are satisfied with the use of information in this context, one could improve the feasibility of using PHR [15]. Some surveys have indicated that consumers want Web-based access to their PHR [16,17]; however, they still have some doubts, including the possibility that full and open access to personal medical information could bring up privacy concerns [18]; the problems of usability and security issues, complexity in the use of PHRs [2]; and perceived usefulness, motivation, patient and health professional interaction, lack of time and workload, resources availability, management, outcome expectancy, and interoperability [19]. Agarwal et al [20] found provider

satisfaction, interactions between environmental factors (communication tactics and value of the tool functionality), and interactions between patient activation and tool empowerment potential were significantly associated with behavioral intentions to use the PHR. Patients who believed the tool to be empowering demonstrated greater intention to use it, which was further enhanced for highly motivated patients. Baird [21] and Liu [22] also thinks that the concerns and challenges of using PHR are more focused on discussions regarding confidentiality, integrity, authorization, access control, portability, efficiency, scalability of solutions, and issues related to user experiences.

Compared with these studies, our results are similar; however, our study focused on system functions and system design and security. These questions are related to perceived usefulness, patient and health professional interaction, and management and interoperability aspects for perceived usefulness. Our research found participants pay more attention to how PHR can help them manage their health records, such as medical records, health examination records, and medical images, while also helping them avoid duplication of examinations, tests, and medication. For patient and health professional interactions, participants think PHR can let physicians add more details to the medical records; regarding management and interoperability, it was found in our study that detailed operating instructions, log in methods, and operation with secure methods are important for the participants, but the satisfaction with performance are below average.

Overall, the issues receiving the most user attention were the system functions, circulation, integrity, ease of use, continuity of the PHR, and data security and privacy protection. "My health bank" query service implemented by the government in recent years allows patients to check medical records (including the

date of medical treatment, drug use, inspection report, and the doctor's advice) for 1 year through a personal certificate. Through this access, the integrity and continuity of PHRs can be achieved, but the propaganda for publicity still needs to be strengthened. In future, the government should provide functions and services that can meet the needs of the users, which will also enable the users to understand their own medical records, enhance understanding of the disease for the doctors, enhance the quality of EMR writing, reduce duplicated examinations, and develop holistic care.

Limitations

The personal EMR platform concept proposed in this study is relatively new in Taiwan. Most people are not yet aware of this process. Therefore, their understanding of the personal EMR platform may be poor. In this study, the system was introduced by way of a video to ensure that the participants understood as much as possible before they began to fill out the Web-based questionnaire, which we expected to reduce possible errors. Although the accuracy and validity of the data were not validated due to the lack of sample representation and extrapolation of the results, this prospective study can act as a reference for future studies on the development of the domestic personal EMR platform.

This study was conducted only by a Web-based questionnaire survey. Most of the participants were people with high Internet usage, young, and the areas of residence were concentrated in the north, which could cause the findings to be generalized to the whole population. Therefore, if the results were to extend to other regions or remote areas, there may be a gap. We suggest that future researches could use diversification methods for the survey.

Conflicts of Interest

None declared.

References

1. NAHIT. Amazonaws. 2008 Apr 28. The National Alliance for Health Information Technology Report to the Office of the National Coordinator for Health Information Technology on Defining Key Health Information Terms URL: <http://s3.amazonaws.com/rdcms-himss/files/production/public/HIMSSorg/Content/files/Code%205%20Defining%20Key%20Health%20Information%20Technology%20Terms.pdf> [accessed 2017-04-13] [WebCite Cache ID 6pgqGckTQ]
2. Roehrs A, de Costa CA, Righi RD, de Oliveira KS. Personal health records: a systematic literature review. J Med Internet Res 2017 Jan 06;19(1):e13 [FREE Full text] [doi: [10.2196/jmir.5876](https://doi.org/10.2196/jmir.5876)] [Medline: [28062391](https://pubmed.ncbi.nlm.nih.gov/28062391/)]
3. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. Br Med J 2007 Aug 18;335(7615):330-333 [FREE Full text] [doi: [10.1136/bmj.39279.482963.AD](https://doi.org/10.1136/bmj.39279.482963.AD)] [Medline: [17703042](https://pubmed.ncbi.nlm.nih.gov/17703042/)]
4. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126 [FREE Full text] [doi: [10.1197/jamia.M2025](https://doi.org/10.1197/jamia.M2025)] [Medline: [16357345](https://pubmed.ncbi.nlm.nih.gov/16357345/)]
5. Reti SR, Feldman HJ, Ross SE, Safran C. Improving personal health records for patient-centered care. J Am Med Inform Assoc 2010;17(2):192-195 [FREE Full text] [doi: [10.1136/jamia.2009.000927](https://doi.org/10.1136/jamia.2009.000927)] [Medline: [20190063](https://pubmed.ncbi.nlm.nih.gov/20190063/)]
6. Baudendistel I, Winkler E, Kamradt M, Brophy S, Längst G, Eckrich F, et al. The patients' active role in managing a personal electronic health record: a qualitative analysis. Support Care Cancer 2015 Feb 5;23(9):2613-2621. [doi: [10.1007/s00520-015-2620-1](https://doi.org/10.1007/s00520-015-2620-1)] [Medline: [25652149](https://pubmed.ncbi.nlm.nih.gov/25652149/)]

7. Jilka SR, Callahan R, Sevdalis N, Mayer EK, Darzi A. "Nothing About Me Without Me": an interpretative review of patient accessible electronic health records. *J Med Internet Res* 2015 Jun 29;17(6):e161 [[FREE Full text](#)] [doi: [10.2196/jmir.4446](https://doi.org/10.2196/jmir.4446)] [Medline: [26123476](#)]
8. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. *J Med Internet Res* 2015 Feb 10;17(2):e44 [[FREE Full text](#)] [doi: [10.2196/jmir.3171](https://doi.org/10.2196/jmir.3171)] [Medline: [25669240](#)]
9. Ronda MC, Dijkhorst-Oei LT, Rutten GE. Reasons and barriers for using a patient portal: survey among patients with diabetes mellitus. *J Med Internet Res* 2014 Nov 25;16(11):e263 [[FREE Full text](#)] [doi: [10.2196/jmir.3457](https://doi.org/10.2196/jmir.3457)] [Medline: [25424228](#)]
10. Otte-Trojel T, de Bont A, van de Klundert J, Rundall TG. Characteristics of patient portals developed in the context of health information exchanges: early policy effects of incentives in the meaningful use program in the United States. *J Med Internet Res* 2014 Nov 21;16(11):e258 [[FREE Full text](#)] [doi: [10.2196/jmir.3698](https://doi.org/10.2196/jmir.3698)] [Medline: [25447837](#)]
11. U.S. Department of Health and Human Services, National Cancer Institute National Institutes of Health, National Center for Health Statistics Centers for Disease Control and Prevention. NCVHS. 2006 Feb. Personal Health Records and Personal Health Record Systems URL: <http://ncvhs.us/wp-content/uploads/2014/05/0602nhiirpt.pdf> [accessed 2017-04-13] [[WebCite Cache ID 6pgqfslqd](#)]
12. Martilla J, James JC. Importance-performance analysis. *Journal of Marketing* 1977 Jan;41(1):77. [doi: [10.2307/1250495](https://doi.org/10.2307/1250495)]
13. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 2005 May;19(Suppl 2):S67-S72. [Medline: [15930843](#)]
14. Liang TH. Matrix analysis of the digital divide in eHealth services using awareness, want, and adoption gap. *J Med Internet Res* 2012 Feb 13;14(1):e11 [[FREE Full text](#)] [doi: [10.2196/jmir.1670](https://doi.org/10.2196/jmir.1670)] [Medline: [22329958](#)]
15. Littlejohns P, Wyatt JC, Garvican L. Evaluating computerised health information systems: hard lessons still to be learnt. *Br Med J* 2003 Apr 19;326(7394):860-863 [[FREE Full text](#)] [doi: [10.1136/bmj.326.7394.860](https://doi.org/10.1136/bmj.326.7394.860)] [Medline: [12702622](#)]
16. Goldzweig CO, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Mlake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. *Ann Intern Med* 2013 Nov 19;159(10):677-687. [doi: [10.7326/0003-4819-159-10-201311190-00006](https://doi.org/10.7326/0003-4819-159-10-201311190-00006)] [Medline: [24247673](#)]
17. Phelps RT, 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 Oct 24;16(10):e241 [[FREE Full text](#)] [doi: [10.2196/jmir.3371](https://doi.org/10.2196/jmir.3371)] [Medline: [25344621](#)]
18. D'Amore JD, Mandel JC, Kreda DA, Swain A, Koromia GA, Sundareswaran S, et al. Are meaningful use stage 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA collaborative. *J Am Med Inform Assoc* 2014;21(6):1060-1068 [[FREE Full text](#)] [doi: [10.1136/amiainl-2014-002883](https://doi.org/10.1136/amiainl-2014-002883)] [Medline: [24970839](#)]
19. McGinn CA, Gagnon M, Shaw N, Sicotte C, Mathieu L, Leduc Y, et al. Users' perspectives of key factors to implementing electronic health records in Canada: a Delphi study. *BMC Med Inform Decis Mak* 2012 Sep 11;12:105 [[FREE Full text](#)] [doi: [10.1186/1472-6947-12-105](https://doi.org/10.1186/1472-6947-12-105)] [Medline: [22967231](#)]
20. Agarwal R, Anderson C, Zarate J, Ward C. If we offer it, will they accept? Factors affecting patient use intentions of personal health records and secure messaging. *J Med Internet Res* 2013 Feb 26;15(2):e43 [[FREE Full text](#)] [doi: [10.2196/jmir.2243](https://doi.org/10.2196/jmir.2243)] [Medline: [23470453](#)]
21. Baird AN, Raghu T. Personal Health Records (PHR) and the Future of the Physician-patient Relationship. 2011 Feb 08 Presented at: Proceedings of the iConference; 2011; New York.
22. Liu LS, Shih PC, Hayes GR. Barriers to the adoption and use of personal health record systems. 2011 Feb 08 Presented at: 2011 iConference; 2011; Seattle.

Abbreviations

EMR: electronic medical records

PHR: Personal Health Records

Edited by A Keepanasseril; submitted 29.11.16; peer-reviewed by P Teixeira, A Azizi; comments to author 11.01.17; revised version received 18.03.17; accepted 27.03.17; published 27.04.17.

Please cite as:

Rau HH, Wu YS, Chu CM, Wang FC, Hsu MH, Chang CW, Chen KH, Lee YL, Kao S, Chiu YL, Wen HC, Fuad A, Hsu CY, Chiu HW
Importance-Performance Analysis of Personal Health Records in Taiwan: A Web-Based Survey
J Med Internet Res 2017;19(4):e131

URL: <http://www.jmir.org/2017/4/e131/>

doi: [10.2196/jmir.7065](https://doi.org/10.2196/jmir.7065)

PMID:

©Hsiao-Hsien Rau, Yi-Syuan Wu, Chi-Ming Chu, Fu-Chung Wang, Min-Huei Hsu, Chi-Wen Chang, Kang-Hua Chen, Yen-Liang Lee, Senyeong Kao, Yu-Lung Chiu, Hsyien-Chia Wen, Anis Fuad, Chien-Yeh Hsu, Hung-Wen Chiu. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.04.2017. 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.

Original Paper

The Multimodal Assessment of Adult Attachment Security: Developing the Biometric Attachment Test

Federico Parra^{1,2}, MA; Raphaële Miljkovitch², PhD; Gwenaëlle Persiaux³, MA; Michelle Morales^{1,4}, MS; Stefan Scherer¹, PhD

¹Institute for Creative Technologies, University of Southern California, Los Angeles, CA, United States

²Paragraphe Laboratory, Paris VIII University, Saint-Denis, France

³University Hospital Center Sainte-Étienne, Sainte-Étienne, France

⁴The Graduate Center, City University of New York, New York, NY, United States

Corresponding Author:

Federico Parra, MA

Institute for Creative Technologies

University of Southern California

12015 E Waterfront Dr

Los Angeles, CA, 90094

United States

Phone: 33 782132695

Fax: 1 (310) 574 5725

Email: federico.parra@hotmail.com

Abstract

Background: Attachment theory has been proven essential for mental health, including psychopathology, development, and interpersonal relationships. Validated psychometric instruments to measure attachment abound but suffer from shortcomings common to traditional psychometrics. Recent developments in multimodal fusion and machine learning pave the way for new automated and objective psychometric instruments for adult attachment that combine psychophysiological, linguistic, and behavioral analyses in the assessment of the construct.

Objective: The aim of this study was to present a new exposure-based, automatic, and objective adult-attachment assessment, the Biometric Attachment Test (BAT), which exposes participants to a short standardized set of visual and music stimuli, whereas their immediate reactions and verbal responses, captured by several computer sense modalities, are automatically analyzed for scoring and classification. We also aimed to empirically validate two of its assumptions: its capacity to measure attachment security and the viability of using themes as placeholders for rotating stimuli.

Methods: A total of 59 French participants from the general population were assessed using the Adult Attachment Questionnaire (AAQ), the Adult Attachment Projective Picture System (AAP), and the Attachment Multiple Model Interview (AMMI) as ground truth for attachment security. They were then exposed to three different BAT stimuli sets, whereas their faces, voices, heart rate (HR), and electrodermal activity (EDA) were recorded. Psychophysiological features, such as skin-conductance response (SCR) and Bayesky stress index; behavioral features, such as gaze and facial expressions; as well as linguistic and paralinguistic features, were automatically extracted. An exploratory analysis was conducted using correlation matrices to uncover the features that are most associated with attachment security. A confirmatory analysis was conducted by creating a single composite effects index and by testing it for correlations with attachment security. The stability of the theory-consistent features across three different stimuli sets was explored using repeated measures analysis of variances (ANOVAs).

Results: In total, 46 theory-consistent correlations were found during the exploration (out of 65 total significant correlations). For example, attachment security as measured by the AAP was correlated with positive facial expressions ($r=.36$, $P=.01$). AMMI's security with the father was inversely correlated with the low frequency (LF) of HRV ($r=-.87$, $P=.03$). Attachment security to partners as measured by the AAQ was inversely correlated with anger facial expression ($r=-.43$, $P=.001$). The confirmatory analysis showed that the composite effects index was significantly correlated to security in the AAP ($r=.26$, $P=.05$) and the AAQ ($r=.30$, $P=.04$) but not in the AMMI. Repeated measures ANOVAs conducted individually on each of the theory-consistent features revealed that only 7 of the 46 (15%) features had significantly different values among responses to three different stimuli sets.

Conclusions: We were able to validate two of the instrument's core assumptions: its capacity to measure attachment security and the viability of using themes as placeholders for rotating stimuli. Future validation of other of its dimensions, as well as the ongoing development of its scoring and classification algorithms is discussed.

(*J Med Internet Res* 2017;19(4):e100) doi:[10.2196/jmir.6898](https://doi.org/10.2196/jmir.6898)

KEYWORDS

psychometrics; linguistics; heart rate; facial expression; psychophysiology; psychopathology; COVAREP; attachment

Introduction

The Relevance of Adult Attachment in Mental Health

Attachment theory originated with the work of a British psychiatrist, John Bowlby [1]. Inspired by ethological observations and evolution theory, he theorized that the chance for survival of human genes had increased by the natural selection of behaviors that augmented proximity and bonding between infants and their caregivers, leading to a greater probability of protection for the children [1,2]. Attachment theory posits an innate psychobiological behavioral system, the *attachment system*, which activates specially in times of perceived threat, inciting the child to seek the proximity and care of their caregivers, the *attachment figures*. The system deactivates once a felt sense of security and safety is reestablished [1,3]. Despite the universality of attachment proximity-seeking behaviors in children [4,5], the security and care sought are only found when the attachment figures are capable of responding promptly and adequately [6]. The quality and outcome of these repeated early attachment interactions leave an enduring mark in the developing person [7-10]. The nature of this mark is threefold: it is cognitive, since dynamic representational models of the attachment figures and the relationship with them develop [11,12], contributing in adulthood to appraisals of the self as worthy of care and of others as capable of providing care [13]; behavioral, because our innate attachment behaviors accommodate to the environment [6], for example, in case of continuous unavailability of caregivers children might stop proximity-seeking behaviors entirely and act as if they were totally independent, a pattern that is then carried into adulthood [14]; and psychobiological, because negative early attachment experiences can lead to differences in the response of the bilateral amygdala and left ventral striatum during stressful situations, and to an overall higher sympathetic activation baseline [15-17].

The different adult attachment patterns have been extensively described in the literature using both dimensional and categorical models [2,18]. In the dimensional approach, the single most important dimension is *attachment security* [12].

Attachment theory sparked some of the largest and more rigorous longitudinal studies in psychology to date [8,19], proving itself essential in three overlapping research domains of mental health: the study of psychopathology, the study of psychological development, and the study of the psychology of adult interpersonal relationships.

In terms of clinical research, longitudinal studies have shown that negative early attachment interactions in childhood predict

childhood attachment security [9], which in turn partially predict adult psychopathology [7,20], whereas cross-sectional studies have consistently linked adult *attachment insecurity* to several psychopathologies [21,22], such as depression [23], post-traumatic stress disorder (PTSD) [24], or borderline personality disorder [25]. Positive attachment experiences in adulthood, whether naturally occurring or the outcome of therapeutic interventions, can help increase attachment security, which in turn improves mental health [26-28].

In terms of developmental psychology, studies show that developmental competencies that are essential to sustain mental health and to cope with mental health disorders, such as emotional regulation, social skills, or cognitive ability, are associated and interdependent with attachment across the lifespan [7,29,30].

Finally, adult attachment is key in the psychology of interpersonal relationships, including long-term romantic relationships [31,32], which tend to function as a buffer in coping with psychopathology and stress [33,34]. Attachment insecurity has been associated with having more interpersonal problems in general [35], and these problems explain insecure persons' self-reported loneliness, social isolation, low relationship satisfaction, more frequent relationship breakups, greater physiological stress reaction to interpersonal conflict, and more frequent conflicts and violence [21,36-38]. Secure attachment, in relationship with social support, has been acknowledged as a protective factor for psychological stress [34], with perceived social support mediating the relationship between attachment security and depressive symptoms [33].

Current Limitations in the Assessment of Adult Attachment

Since 1985 (Findings by George C, Kaplan N, and Main M, unpublished data, 1985), various validated instruments for the assessment of adult attachment developed concomitantly within the fields of social psychology and developmental psychology (for a review, see [39]). Social psychology has spurred the development of several questionnaires, such as the Adult Attachment Questionnaire (AAQ [40]) or the Adult Attachment Scale (AAS [41]). Developmental psychology, on the other hand, has relied on a variety of broadly defined semistructured interview methods, beginning with the Adult Attachment Interview (AAI; Findings by George C, Kaplan N, and Main M, unpublished data, 1985) which is considered the "gold standard" [39]. Both approaches suffer from several limitations that affect both research and clinical assessments, and that are reflective of the current state of psychometrics.

Questionnaire-based assessments are self-report measures. As such, they are prone to self-report biases that have been well

described in the literature [42]. In terms of construct validity, there has been no longitudinal association demonstrated between attachment in childhood as measured for example with the Strange Situation Procedure (SSP [43]) and adult attachment as measured with questionnaires [19]. Furthermore, almost no concurrent validity has been found between questionnaires and interview-based assessments of adult attachment, adding to the construct validity controversy [39]. On the other hand, questionnaires of adult attachment are easy, economic, and fast to both administer and score. Administration can be done remotely, and automatic scoring is possible. These positive practical psychometric characteristics may explain the surge of studies that have chosen questionnaires of adult attachment as their measure [39].

Interview-based assessments of adult attachment rely on some form of semistructured interview, which is later transcribed and scored by a trained judge, that has undergone substantial training in a specific standardized scoring tradition. In a way, this form of assessment is closer to child assessments which also rely on third-party experts for scoring and classification. However, in childhood-attachment assessments such as the SSP, the scoring experts observe behavior in general, whereas in interview methods only transcribed language is analyzed during scoring, thus limiting the scope of dimensions evaluated in this process. The Attachment Multiple Model Interview (AMMI [12]) circumvents this limitation in part, by including in the interview specific probes to gather self-reported information about behaviors.

In terms of construct validity, the AAI has consistently shown a link between parents and their children's attachment patterns, which is considered strong evidence of its validity [2]. Moreover, a substantial longitudinal link has been found using the AMMI [12], further supporting the consensus that interview methods based on expert judgment can produce results with higher construct validity than self-report measures.

But despite this consensus, interview methods are not without their own limitations. In contrast to their questionnaire-based counterparts, interview methods are difficult, costly, and lengthy to both administer and score. They add additional layers to the process, that is, the manual transcription and coding of the interview. There is a training load required for both administering and scoring. This process is costly.

Interview methods are impacted by an additional problem: the subjectivity inherent to an expert judge [44]. This limitation might decrease the replicability of attachment studies, adding to psychology's current "replicability crisis" [45].

Finally and contrary to questionnaires, interview methods cannot be administered remotely, limiting their application, for instance, in Internet-based research.

Advances in Multimodal Analysis and Automatic Detection of Psychological Markers

Finding psychophysiological and behavioral markers of psychological conditions is gaining traction within mainstream psychiatry [46,47], as part of a quest to provide more objective and precise clinical assessments to patients. The American National Institute of Mental Health released a statement in 2013

[48] in which it made explicit its desire of moving toward more objective and precise diagnostic methodologies. Several attempts to tackle this problem have arisen from the Computer Sciences. In a recent review, Cummins et al [46] reviewed the state-of-the-art in the automatic detection of depression and suicidality through the analysis of speech and its paralinguistic acoustic features. Scherer et al [49] described, in 2013, a set of automatically extracted audiovisual nonverbal behavioral features helpful in the identification of depression, anxiety, and PTSD [49]. The link between the objective measure of singular biometric or behavioral markers, and the sought ability to offer more precise diagnoses, relies on the use of machine learning algorithms that can fuse multiple modalities of data at once [50]. This allows for the uncovering of complex multimodal data patterns that can serve in the automatic assessment of specific mental conditions. In recent studies, such multimodal systems have approached human performance in the detection of indicators of PTSD [51]. Since several studies on the specific psychophysiological [2] and linguistic [44] traits of adult attachment already exist and show promise, we decided it was time to use this new technology in the assessment of adult attachment.

The Biometric Attachment Test

The Biometric Attachment Test (BAT) was created with the objective of automatically and objectively measuring attachment in adults. At its core, the BAT is an exposure-based test, which means that the participant being tested is exposed to a short (9 min) standardized set of visual and music stimuli, whereas their immediate reactions and verbal responses, captured by several computer sense modalities, are automatically analyzed for scoring and classification.

There are two aspects of the development of the BAT that require separate attention: the instrument itself, meaning its assumptions, stimuli selection, and administration protocols, which will be articulated in this work; and the test's automatic classification and scoring algorithms, a work-in-progress that we will briefly touch upon in "Discussion" section.

Construction of the Biometric Attachment Test (BAT)

The BAT was strongly influenced by three previous instruments: Bowlby's first Separation Anxiety Test (SAT [52]), the previously mentioned SSP [43], and the AAP [13].

The SAT (1976 version [52]) is a projective attachment test for children aged 4-7 years consisting of a set of 6 pictures depicting situations in which a child, separated from their family, must cope on their own without help from their parents. The tested child is asked to interpret the protagonist's feelings and predict their behavior, and their transcribed responses are later scored and classified.

The SSP is a structured observation protocol for assessing attachment in children aged between 12 and 24 months. During 20 min, the child undergoes a series of separations and reunions from their caregiver, while they are also exposed to the arrival and presence of a stranger. The child's behavior is videotaped and then analyzed for attachment scoring and classification.

The AAP is an adult attachment test based on a set of black and white drawings, some of which are ambiguous, depicting more diverse situations that activate the attachment system: separation, loss, solitude, and physical threat [13]. Participants are asked to tell a short story about the pictures, which are transcribed and analyzed, and an attachment classification and continuous scores are obtained [13].

Like the SAT, our BAT uses photos, of real people, in explicit situations. Like the AAP, our stimuli depict a variety of attachment-sensitive situations. Like the SSP, the BAT is meant to produce an *alternating activation and deactivation* of the attachment system, with stimuli representing themes such as loss, death, or separation alternating with stimuli representing themes such as intimate connection, soothing, or protection.

Unlike other exposure-based and projective tests, the BAT uses music stimuli in addition to visual stimuli, both on its own and concomitantly with visual stimuli. Music was included because of its ability to trigger strong emotional feelings and experiences [53].

Like the SSP, scoring and classification in the BAT take into consideration observed behaviors. In fact, unlike other tests in which verbatim transcripts of verbal responses are analyzed, the BAT captures the participants' reactions and responses in a variety of modalities: physiological (heart rate [HR] and electrodermal response [EDA]) from which psychophysiological features can be derived (eg, Bayevsky stress index [54]), behavioral (facial expressions, gaze, face distance from stimuli, paralinguistic speech characteristics), and verbal.

The Concept of Themes in the Biometric Attachment Test (BAT)

Exposure-based and projective psychometric tests typically rely on a fixed set of stimuli selected by the authors [13,55]. We pose the following critiques to this approach: first, stimuli can eventually leak into the public domain, such as in the case of the Rorschach [56], and this might undermine a test's effectiveness due to priming effects. Second, longitudinal studies such as clinical trials require participants to be tested several times using the same instruments, and if stimuli are always the same this might also lead to priming effects. Finally, we believe ideally stimuli should be selected based on input from the general population toward which it is destined.

Our BAT innovates introducing the concept of *themes*: placeholders for actual stimuli. A theme is a narrative that describes a specific situation to be evoked by a stimulus, with a specific objective. For example, in terms of adult attachment, a theme could be "the loss of a close one," its objective being to activate the attachment system (ie, to cause attachment-related distress).

Themes thus can solve the aforementioned problems with fixed-stimuli test designs: since themes are placeholders for stimuli as opposed to fixed stimuli, there is no risk if a stimuli set becomes widely known. All the contrary: stimuli in the BAT can—and should—be replaced from time-to-time and from

context-to-context. In the case of clinical trials, stimuli sets in the BAT could rotate between assessments. Finally, the process for stimuli selection in the BAT is standardized and crowdsourced, as we will see briefly.

About the themes' objectives, each is meant to evoke a reaction in the participant depending on the participant's attachment patterns. The themes were inspired by the SAT, the AAP, the SSP, and attachment theory core principles. In total, 14 themes resulted from this work (see Figure 1).

Theme 1 ("baseline") was designed to measure the participants' reactions to being in the test situation, where they are still not being confronted with any attachment-specific stimulus. This provides proper baselines for all biometric and behavioral measures.

Themes 2, 8, and 10 were designed to elicit specific reactions depending on the underlying attachment pattern of the test participant, to help in classification.





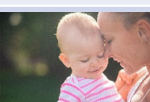
All other BAT themes have per objective to either activate (ie, stress) or deactivate (ie, calm) the attachment system. We would like to clarify that throughout this paper we use the terms "attachment activation" and "attachment deactivation" in their literal sense, that is, the way in which the attachment system is activated when under specific relational stress and how it becomes deactivated when that relational stress is sufficiently addressed. This is not to be confused with "avoidant deactivation," a "Minimizing strategy (...) conceived by Main (1990) as a shift of attention away from conditions normally eliciting attachment behavior, leading to the apparent absence of attachment behaviors in such circumstances" [12].

Stimuli Selection in the Biometric Attachment Test (BAT): A Standardized Process

A set of objective and subjective criteria were developed for each of the BAT's themes. The objective criteria were directly derived from the themes' narratives: for example, for a stimulus to be appropriate to represent the "attuned mother-child" theme, there should be a mother and a child in the picture. Subjective criteria are notions that require more complex judgments: for example, for a stimulus to be appropriate to depict the "attuned mother-child" theme, the child and the mother must seem attuned to each other and, thanks to said attunement, they should both seem relatively relaxed. To decide whether a mother or a child seem relaxed or not just by looking at them in a picture is a subjective process that should not be arbitrarily decided by researchers.

We used the straightforward objective criteria to preselect stimuli: three large picture databases conceived for the study of emotion were used: the Nencki Affective Picture System (NAPS [57]), the International Affective Picture System (IAPS [58]), and the Geneva Affective Picture Database (GAPED [59]). In some cases, none of these databases had enough pictures for some of the themes, so we turned to a stock picture service, iStockPhoto. We ended up with 126 preselected pictures.

Figure 1. Biometric Attachment Test (BAT) themes, goals, and stimuli set sample.

Theme	Theme name	Objective	Stimuli sample
1	Baseline	-	
2	Child solitude	Depends on attachment pattern	
3	Couple separation	Attachment activating or stressing	
4	Attuned father-child	Attachment deactivating or soothing	
5	Violent father	Attachment activating or stressing	
6	Unresponsive mother	Attachment activating or stressing	
7	Attuned couple	Attachment deactivating or soothing	
8	Exploration	Depends on attachment pattern	
9	Raw sadness	Attachment activating or stressing	Sad music
10	Vulnerable baby	Depends on attachment pattern	
11	Raw fear	Attachment activating or stressing	Tense music
12	Death or loss	Attachment activating or stressing	
13	Attuned mother-child	Attachment deactivating or soothing	
14	Family connection	Attachment deactivating or soothing	

The Web-Based Survey

We then created an anonymous web-based survey using SurveyGizmo services. The survey randomly introduced each of the preselected pictures, accompanied by sliders that

participants could adjust to the right or to the left, signaling a judgment about a specific criterion. In the case of the “attuned mother-child” theme, for example, one of the available sliders allowed the participant to judge the perceived level of stress of the child in the picture. We always opposed two traits (eg,

stressed vs relaxed), randomizing their order and starting with the slider in the center among them.

Our survey was made available in Spanish, English, and French and was distributed through social media and email campaigns in the United States, France, and Argentina. A total of 520 participants (female=72.3%, 376/520, male=27.7%, 144/520), of a variety of ages (mean 37.53, SD 10.87) responded. The survey was kept online for a period of 10 days between March 3, 2016 and March 13, 2016. Results were then cleaned-up using standard survey results cleaning best practices [60].

We created composite scores formulas for each of the 14 BAT themes, allowing to combine the subjective criteria measured in the survey. For the “attuned mother-child” theme, for example, the composite score formula was composed by the perceived level of genuineness of the picture, plus the perceived attunement between mother and child in the picture, minus the perceived levels of stress in the child and in the mother in the picture, individually.

A minimum required composite score was set for each theme to prevent stimulus that are not evocative enough from being used in the future.

The list of themes, objective and subjective criteria, as well as the survey design, are available for other researchers to generate new stimuli sets for the BAT in the future (contact corresponding author).

Music selection was easier and did not require a survey process. Music themes were conceived to convey basic raw emotions (eg, theme 9, “raw sadness”). A total of 25 short music clips were selected from a music set conceived to elicit emotion and that already provides scores on discrete perceived emotions [61]. We simply chose the music clips with higher scores in the required emotion per theme.

Biometric Attachment Test (BAT) Administration Procedure

The BAT automatic administration procedure was constructed using OpenSesame software, version 3.1.2 [62]. The full test duration is of 9 min.

Before beginning, the participant is instructed to observe the visual or listen to the music stimuli as long as it is visible or audible, and then to describe aloud what they felt about it.

During the test, each theme stimulus is automatically presented for 15 s, followed by a black screen displaying the phrase “What did you feel?” which shows for 25 s, whereas a countdown slider displays the available time to respond. It then shows the phrase “Thank you. Here is the following stimulus...,” for 5 s, followed by the following stimulus, and so forth.

The test is administered with the person being alone with the computer in a room; aloneness can facilitate the activation of the attachment system [1] and renders the test situation closer to Ainsworth’s SSP [6]. It also removes possible interference from researchers.

Hypotheses of This Study

This study was designed to empirically evaluate two core assumptions of the BAT:

H1: Those adults with higher attachment security will more successfully use the BAT’s attachment-deactivating themes to reassure and soothe themselves, and this will be in turn reflected in specific psychophysiological, behavioral and linguistic markers. Theme 4 (“attuned father-child”), 7 (“attuned couple”), and 13 (“attuned mother-child”) are evocative of the availability of attachment figures and will be used to test this hypothesis.

H2: Different stimuli sets, selected through our standardized process, are interchangeable and cause very similar responses or reactions in participants. Specifically, the features most associated with attachment security will remain consistent across three different stimuli sets.

Methods

Sample

The sample consisted of 59 French francophone participants (45 females, 14 males) that were interviewed between March and May, 2016. The sample was formed from multiple sources in different regions of France: 9 psychiatric patients recruited at University Hospital Center Sainte-Étienne and 7 recruited at the Ville Evrard Center of Psychotherapy in Saint Denis; 29 volunteers enrolled in Mornant, Paris, and Rouen; and 14 college students enrolled at Paris 8 University in Saint Denis. It was intended for the sample to be as diverse as possible in terms of age (mean 35.7, SD 12.2), occupational status (10% unemployed, 6/59, 32% employed, 19/59, 33% students, 20/59, 23% other, 14/59), as well as relationship status (37% in a relationship, 22/59, 23% married, 14/59, 11% separated or divorced, 7/59, 25% single, 14/59, 3% unknown, 2/59) and psychopathology (27%, 16/59 were patients). Since questions about ethnicity or race are not allowed in French research, we don’t have information to report about the ethnic diversity of our sample. All participants signed informed consent forms in accordance to best practices in French Universities.

Measures

Adult Attachment Questionnaire

Fifty of our participants completed the AAQ before the interview, as a web-based questionnaire. The AAQ is a 17-item measure that asks individuals to indicate how they relate to romantic partners in general. It yields a continuous measure of attachment security with regards to romantic partners [40].

Adult Attachment Projective Picture System

All our participants completed the AAP test, which was introduced earlier. Transcripts of the AAP were scored by a trained member of our team blind to all information about the participants. Interjudge reliability was obtained for 5 cases that were double-coded by one of the AAP’s creators, with 80% of interrater agreement for both classifications and scores. The AAP outputs a continuous attachment security score, called “agency of self,” which per George [13] has both an inward and outward aspects. For this study, we’ll focus on the latter, which

evaluates the degree to which an individual seeks for, and trust, attachment figures to provide for them a haven of safety in times of stress [13].

Attachment Multiple Model Interview

The AMMI is a validated semistructured interview that investigates participants' reactions in attachment-related situations. By analyzing and scoring transcriptions of the interviews, AMMI provides scores for three different attachment relationships: attachment to mother, father, and partner [12]. Since each relationship requires a specific amount of interview time, not all participants were able to complete all the interviews: attachment to the mother was evaluated for 27 participants, attachment to the father for 23, and attachment to the partner for 17. Transcripts of the AMMI were scored by a trained member of our team. Six cases were double-coded by the AMMI's creator, and interrater reliability was satisfying (83% of agreement).

The aforementioned measures have been validated in several languages including English. Their French version was used during this study.

Biometric Attachment Test

In order to evaluate our second hypothesis, we produced three BAT stimuli sets for this study: two fixed ones (ie, that show the same stimuli each time they are used) and a randomized one (ie, that shows different stimuli each time it is used). We have used the results from the French subsample of the survey respondents ($n=194$) to select the best pictures for a French population. The higher ranked pictures for each theme were put together in a stimuli set; the second higher ranked pictures were put together in a second stimuli set; and the pictures ranked third, fourth, and fifth were used to create a third set that randomly chooses one of those pictures each time it is played.

All 59 participants were exposed to the first stimuli set, 41 of them were also exposed to the second set, and 50 to the third rotating-stimuli set. Sets were presented one after the other.

Physiological Measures

HR was measured using the photoplethysmography sensor of an Empatica E4 wristband device. The sensor's reliability has been established [63]. Like all heart sensors, the E4 is subject to artifacts produced by movement. Quality readings were obtained for 29 participants during the first BAT set, 19 during the second BAT set, and 9 during the third. Electrodermal activity (EDA), with a specific interest in skin conductance response (SCR), was measured using the EDA sensor of the Empatica E4 wristband device. Quality readings were obtained for all participants during all BATs. The EDA sensor's reliability has been tested by the manufacturer [64]. Deliberately choosing to use a wireless wristband to measure physiological signals allowed our participants a more natural experience during the test.

Video and Audio Recording

Video of the participants' faces was obtained through the computer's webcam (Microsoft Surface Pro 4) and their speech was recorded using a USB Microphone (Samson GoMic). Since the BAT stimuli were presented using the same computer,

gazing toward the stimuli was almost equivalent to gazing in the direction of the camera, facilitating gaze tracking.

Feature Extraction

We conducted extensive feature extraction from each of the sense modalities captured during the BAT. All feature extraction procedures described below, including noise filtering processes, were performed programmatically without the need for human supervision.

The interbeat interval (IBI) was automatically calculated from the HR data by proprietary algorithms of the Empatica E4 research wristband [65]. The IBI files were cleaned of artifacts using Artiifact software, version 209 [66]. The same software was used for the extraction of heart rate variability (HRV) features (for a review of most standard HRV features, see [67]). We created a function in Microsoft Excel's Visual Basic for Applications version 7.1 to automatically calculate Bayevsky stress index [54] from the IBI files.

From the EDA data, SCR, phasic maximal activity, and tonic skin conductance features were extracted using LedaLab software version 349 [68].

From the video data, facial expressions (such as anger and contempt, as well as the composites negative, neutral, and positive) were extracted using FACET's Emotient [69]. A face size measure was extracted by the same software, which permits to establish the movement toward or away from the camera and thus the stimuli. Gaze direction was extracted using OpenFace [70].

From the audio recordings of the participants' responses, paralinguistic acoustic features were extracted using the Cooperative Voice Analysis Repository for Speech Technologies (COVAREP) version 1.2 [71]. They help identify a breathy, relaxed voice from a tense voice.

We used Python and the French language model of Google's Cloud Speech API to generate automated transcripts of all responses. We then processed the transcripts using Python and Linguistic Inquiry and Word Count (LIWC) French dictionary [72]. This dictionary is organized in 64 psychologically meaningful word categories. The frequency of each word category in the response to each theme was calculated, to be used as linguistic features.

Extracted features per theme were then treated in two different ways:

Subtracted baseline: results on the first theme (baseline theme) were subtracted from all other themes' results. In theory, the resultant score should be more individualized to each person's individual characteristics (eg, their specific mean HR baseline).

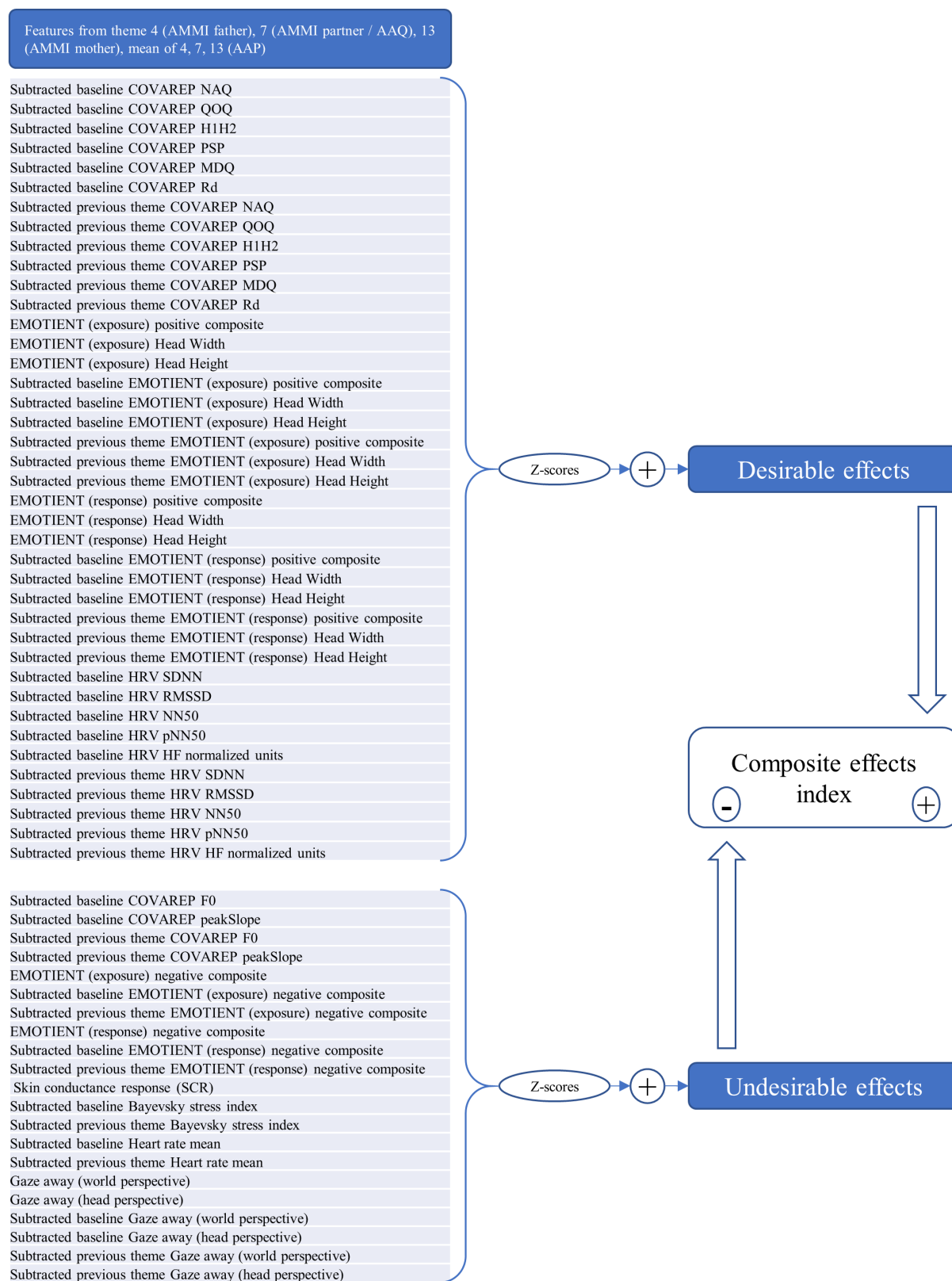
Subtracted previous theme: results on each theme were subtracted from the following one. In theory, the resultant score would isolate results from the exposure to the theme under analysis from the cumulative score due to exposure to all precedent themes (eg, the specific augmentation or decrease in mean HR when exposed to theme 7).

For many features, we further separated the reaction during exposure to the stimuli from the reaction during the verbal

response to the stimuli, for example, facial expressions during exposure versus response.

Due to the high number of features extracted, the number of BAT themes, the two treatments we just described, and the separation between exposure and response, feature extraction led to a total of 4264 features per participant per stimuli set.

In this study, we will focus on specific themes instead of the entire stimuli set, and each theme has 202 features (see [Figure 2](#)). In total, 2436 features pertain to the entire stimuli set as opposed to any single theme (eg, total stimuli-set-wise mean HR).

Figure 2. Development of the composite effects index.

Analyses

A first exploratory analysis, conducted in MathWorks Matlab version R2016A, consisted of performing correlation matrices to uncover associations between attachment security as measured

by the AAQ, the AAP, and the AMMI, and the features extracted from BAT responses. For the AAQ and the AAP, we used Pearson correlations, whereas Spearman rank was used for the AMMI due to the small number of assessed participants. Since AAQ measures attachment to romantic partners, it was evaluated

vis-à-vis theme 7 (“attuned couple”). AAP attachment security is concerned with attachment figures in general, thus we evaluated it with regards to a composite formed by the mean of responses to theme 4 (“attuned father-child”), theme 7 (“attuned couple”), and theme 13 (“attuned mother-child”). As per the AMMI, since it yields security scores for mother, father, and partner separately, we explored each in relationship to the corresponding BAT theme (themes 13, 4, and 7, respectively).

A second, confirmatory analysis was conducted using IBM SPSS Statistics version 23 to verify if our exploratory findings were not a mere spurious byproduct of multiple hypotheses testing [73]. We proceeded with a stringent approach consisting of producing a single “composite effects index” out of all available features (weighted in the same direction), then testing such index for a Pearson correlation vis-à-vis the variable of interest [74,75]. This approach circumvents the problem of type I errors often encountered in exploratory analyses. It also accounts for the problem of type II errors, which are likely when statistical correction procedures to control for family wise error rate (eg, Bonferroni correction) or false discovery rate (eg, Benjamini-Hochberg correction) are performed in studies with small samples, an elevated number of features, or small effect sizes such as ours [76-78].

For this analysis, our features’ scores were first transformed into z scores. Next, they were added to either an undesirable effects group or a desirable effects group. The decision was based on the literature available on each of the set of features, with the following results: the high frequency (HF) component of HRV (associated with parasympathetic “relaxing” activation), HRV’s SDNN, RMSSD, NN50, and pNN50 features (all of which convey slightly different aspects of the same, desirable construct: HRV), COVAREP features associated with a “breathy” relaxed voice, Emotient’s “positive emotions” composite, as well as the head size (proximity of participant to stimuli source), were all summed up within a desirable effects group. On the other hand, Bayevsky stress index, HR, gazing away from the stimuli, COVAREP features associated with a “tense” voice, Emotient’s “negative emotions” composite, as well as SCR levels were all summed up within an undesirable effects group. For each of the aforementioned features, scores extracted from the exposure phase and those from the response phase of the BAT were summed up (when available). Score treatments (subtracted baseline, subtracted previous theme) described earlier were also summed up, when available. Finally, a single composite effects index was created by subtracting the total score of the undesirable effects group from that of the desirable effects group. This index therefore is weighed in such a way that a higher score means more desirable effects and vice versa. Figure 2 illustrates this analysis.

Unfortunately, we could not include LIWC (linguistic features) in the analysis because they cannot easily be distributed among simple desirable or undesirable effects groups (eg, features such as “frequency of future tense verbs”). Finally, a few mathematically redundant (ie, equal information) features were omitted from this analysis, namely, the LF component of HRV in normalized units, as well as the HF/LF ratio of HRV (because their information is mathematically redundant with respect to the HF component in normalized units, see [79]); the percentage

(%) and absolute power versions of the HF component feature (because the normalized units version of the feature controls for the very low frequency (VLF) component of HRV and thus is a more realistic measure of the same construct); the mean and median R-R features of HRV (because they are redundant with respect to HR). Specific EMOTIENT emotion features (eg, sadness) were not included separately since they are all included in two composites already produced by the software, one for negative expressions and the other for positive expressions. The Phasic maximal activity feature of EDA was not included for being redundant with respect to EDA’s SCR. The tonic skin conductance feature was not included because it requires longer measuring durations to be meaningful (they were calculated for future analyses focusing on the totality of the BAT instead of just isolated themes). The total amount of features per theme that ended being added up in the composite effects index is of 61 (see Figure 2).

A third analysis, conducted using IBM SPSS Statistics version 23, consisted in performing repeated measures analysis of variances (ANOVAs) on the BAT responses extracted features that were revealed as both statistically significant in their correlation to attachment security as well as theory consistent with attachment theory. The objective was to evaluate if those features yielded different results across different BAT stimuli sets or if they were consistently similar.

Results

Correlation Exploratory Analyses

Adult Attachment Questionnaire (AAQ; Pearson Correlations)

In the responses to BAT’s theme 7 (“attuned couple”), AAQ romantic attachment security was negatively correlated with negative facial expressions in general during exposure ($r=-.32$, $P=.02$) and anger in particular during response ($r=-.43$, $P=.001$) and exposure ($r=-.38$, $P=.006$). AAQ attachment security was also negatively correlated with the inhibition ($r=-.38$, $P=.008$), tentative ($r=-.34$, $P=.02$), and feeling ($r=-.41$, $P=.004$) categories of LIWC.

Adult Attachment Projective Picture System (AAP; Pearson Correlations)

In the responses to BAT’s theme 4 (“attuned father-child”), theme 7 (“attuned couple”), and theme 13 (“attuned mother-child”), using the mean of the responses to the three themes as a composite score, AAP attachment security was correlated with the NN50 after subtracting baseline ($r=0.48$, $P=.007$) and pNN50 after subtracting baseline ($r=.38$, $P=.04$), features of HRV, while it was negatively correlated with Bayevsky’s stress index after subtracting baseline ($r=-.45$, $P=.01$). AAP security was also correlated with the H1-H2 ratio of COVAREP after subtracting baseline ($r=.30$, $P=.02$). It was correlated as well with the hearing ($r=.31$, $P=.02$), we ($r=.45$, $P<.001$), leisure ($r=.28$, $P=.04$), and they ($r=.40$, $P=.002$) categories of LIWC. It was also correlated with positive facial expressions in general during exposure ($r=.32$, $P=.01$) and response ($r=.36$, $P=.005$) after subtracting previous theme, and joy in particular during response ($r=.38$, $P=.003$) after

subtracting previous theme, as well as disgust in both exposure ($r=.31$, $P=.02$) and response ($r=.33$, $P=.01$) after subtracting previous theme. It was also negatively correlated with surprise on both exposure ($r=-.29$, $P=.03$) and response ($r=-.32$, $P=.01$), confusion during both exposure ($r=-.27$, $P=.04$) and response ($r=-.27$, $P=.04$), confusion during response ($r=-.28$, $P=.03$) after subtracting previous theme, anger during both exposure ($r=-.37$, $P=.004$) and response ($r=-.31$, $P=.02$) after subtracting previous theme, sadness during response ($r=-.26$, $P=.04$) after subtracting previous theme, neutral expressions during response ($r=-.29$, $P=.02$) after subtracting previous theme, fear during exposure after subtracting baseline ($r=-.26$, $P=.05$), and contempt during response ($r=-.26$, $P=.05$). It was negatively correlated with head size during exposure ($r=-.32$, $P=.01$) after subtracting previous theme and with head size during both exposure ($r=-.37$, $P=.003$) and response ($r=-.35$, $P=.01$) after subtracting baseline.

Attachment Multiple Model Interview (AMMI; Spearman Rank Correlations)

In the responses to BAT's theme 4 ("attuned father-child"), AMMI father attachment security was correlated with the past tense ($r=.47$, $P=.02$), and negatively correlated with the discrepancy ($r=-.46$, $P=.03$) categories of LIWC. It was also negatively correlated with the LF feature of HRV after subtracting baseline ($r=-.87$, $P=.03$), and with SDNN feature of HRV after subtracting baseline ($r=-.87$, $P=.03$). It was also negatively correlated with head size in both exposure ($r=-.43$, $P=.04$) and response ($r=-.45$, $P=.03$). Finally, it was negatively correlated with facial expression of frustration during response ($r=-.67$, $P<.001$) and frustration after subtracting baseline during exposure ($r=-.44$, $P=.04$), as well as expressions of frustration ($r=-.53$, $P=.01$) and surprise ($r=-.43$, $P=.04$) after subtracting previous theme.

In the responses to BAT's theme 13 ("attuned mother-child"), AMMI mother attachment security was correlated with SCR ($r=.46$, $P=.02$). It was correlated with facial expressions of sadness during exposure after subtracting baseline ($r=.42$, $P=.03$). It was negatively correlated with facial expression of sadness ($r=-.51$, $P=.01$) and with gazing away from the stimuli during exposure ($r=-.41$, $P=.03$) after subtracting previous theme.

In the responses to BAT's theme 7 ("attuned couple"), AMMI partner attachment security was correlated with the high frequency feature of HRV in both normalized units ($r=1$, $P=.02$) and percentage ($r=1$, $P=.02$) after subtracting previous theme, and it was negatively correlated with the low frequency feature of HRV in both normalized units ($r=-1$, $P=.02$) and percentage ($r=-1$, $P=.02$) after subtracting previous theme, negatively correlated with the ratio of low versus high frequency of HRV ($r=-1$, $P=.02$) after subtracting previous theme, as well as negatively correlated with the mean heart rate ($r=-1$, $P=.02$). It was correlated with SCR ($r=.61$, $P=.01$). It was also correlated with face closeness to screen after subtracting baseline ($r=.56$, $P=.02$). It was negatively correlated with COVAREP Rd feature after subtracting baseline ($r=-.60$, $P=.01$). It was correlated with facial expressions of surprise after subtracting baseline

($r=.59$, $P=.01$). Finally, it was correlated with LIWC *exclusion* category ($r=.49$, $P=.05$).

Confirmatory Analysis

Our composite effects index was significantly correlated to attachment security in the Adult Attachment Projective Picture System ($r=.26$, $P=.05$) by using the mean score from features of themes 4, 7, and 13 (like in the previous analysis), and significantly correlated to attachment security in the AAQ ($r=.30$, $P=.04$) by using scores from features of theme 7. Security with father, mother, and partner in the AMMI were unrelated to the composite effects index by using scores from features of themes 4, 13, and 7, respectively.

Analysis of Variances (ANOVAs)

Repeated measures ANOVAs were conducted on each of the 46 features that were both statistically significant in their correlation to attachment security as well as theory consistent across the responses to the 3 BAT stimuli sets during the aforementioned specific themes. They revealed that only 7 (15%) of the 46 features had significantly different values depending on the stimuli set. Those features were the tentative category of LIWC during theme 7, $F_{1,56,49,92}=4.81$, $P=.02$ (after a Greenhouse-Geisser correction); gazing away from the stimuli (after subtracting previous theme) during exposure to theme 13, $F_{2,76}=5.75$, $P=.005$; and during themes 4, 7, and 13 (mean of the three), the hearing category of LIWC, $F_{2,64}=4.37$, $P=.02$; the leisure category of LIWC, $F_{2,64}=4.63$, $P=.01$; confusion facial expressions during exposure, $F_{1,7,66}=5$, $P=.01$ (after a Greenhouse-Geisser correction) and during response, $F_{1,7,63}=6.3$, $P=.005$ (after a Greenhouse-Geisser correction), as well as anger facial expressions during response after subtracting previous theme, $F_{1,7,62}=3.7$, $P=.04$ (after a Greenhouse-Geisser correction).

Discussion

Principal Findings

Earlier in this work, we discussed the relevance that attachment theory has earned in mental health research, and we commented on the current limitations of psychometric instruments for assessing adult attachment.

We presented the BAT, a new adult-attachment assessment instrument, explicating its sources as well as its rationales and assumptions.

We then set to empirically evaluate two of the BAT's core assumptions: that its themes can help measure attachment security as assessed by validated measures such as the AAQ, the AAP, and the AMMI; and that rotating the stimuli sets in the BAT would not alter the participants' responses to the test.

Regarding the first hypothesis H1, during our exploratory analysis we were able to find physiological, behavioral, and linguistic markers of attachment security, both in general, and specific to romantic partners, mother, and father. These markers were elicited by the BAT's specifically designed attachment-deactivating themes, which counts as preliminary evidence for the instrument's internal and construct validity.

There was an important level of coherence, as well as theory consistency within our findings: in the presence of attachment-deactivating, reassuring stimuli, the more securely attached individuals experienced parasympathetic activation and sympathetic deactivation, a relaxation response revealed by increase in the HF and decrease in the LF of HRV [67], which also produced a decrease in overall stress as shown by the Bayevsky stress index [54]. During the verbal responses, the more securely attached participants' voices became *breathier*, as revealed by COVAREP, indicating relaxation [71], and used more words that can convey attunement, like words related to hearing (eg, listen, heard) and the we pronoun, and conversely less words that can convey relational distress, like words related to inhibition (eg, block, constrain) and to tentativeness (eg, maybe, guess). Positive facial expressions, including joy, were related to attachment security, and their negative counterparts were mostly negatively correlated, as the theory would suggest. The more secure participants tended to not gaze away from the reassuring stimuli but, instead, got physically closer to them. Findings were not all theory consistent, however, as we'll see below.

Since our exploratory analysis was based on multiple hypotheses testing, a statistical concern arose: could these findings be just the product of chance? But when all the available features, including the many that were not *significantly* correlated with attachment in the exploratory analyses, were summed up in a single composite effects index, said index was significantly correlated with two of our three attachment security "ground truth" measures, attesting to the robustness of the findings. This analysis might also suggest that some of the features not showing a statistically significant association with attachment security might not achieve so because of a small sample size or small effect sizes. Composites help increase the effect size of features weighting in the same direction, statistically revealing their direction [74].

We believe this sort of multimodal automatic appraisal of "the whole picture" that is an attachment-deactivating reaction, from a behavioral, psychophysiological, and linguistic standpoint, is a taste of what is becoming possible for psychometrics. Moreover, the fact that our data was obtained outside of a lab setting, using a consumer tablet and its webcam, a consumer-grade USB microphone, and a wireless wristband, attests to the pace at which sensing technology is advancing, offering a glimpse at how effortlessly these measures could be obtained in a close future.

Regarding the second hypothesis H2, 84% (39/46) of the features revealed as associated with attachment security and that were theory consistent were stable across three different BAT stimuli sets. This is especially remarkable given that one of those stimuli sets (always presented second in order) was randomized every time the test was administrated, which means that no person, among our sample, saw or listened to the exact same second stimuli set. This finding suggests that during the BAT, participants react and respond mainly to the themes (ie, the attachment narratives) which are being evoked by the stimuli, and not so much to the stimulus details themselves (eg, the color of a person's hair or the specific background). This also suggests that, as long as stimuli are selected using the

standardized procedure described in this work and our minimum fitness scores are respected, new stimuli sets could be developed for the BAT without affecting its capacity to evoke and measure attachment. H2 results suggest, of course, that we should stop including the 7 features that did change across stimuli sets in further developments of the BAT's scoring algorithms, as they seem to be less reliable when stimuli sets are varied.

We have chosen not to perform multiple comparison corrections (eg, Bonferroni) in our repeated measures ANOVAs to, counterintuitively, increase the rigor of the analysis. This is because within our ANOVA analysis, we compared 46 features across 3 different stimuli sets, for a total of 138 *F* tests. A Bonferroni correction would imply that the alpha level is divided by the number of comparisons (.05/138) for a corrected alpha level of .0003. It would be very difficult for any difference to be found under this alpha level with our sample size. This would be a convenient result, but probably a false one.

Limitations

Out of the 65 significant correlations revealed by our exploratory analysis, 19 (29%) seemed to go against what would be expected from an attachment theory perspective. Some of the most striking examples were the increase in sadness expressions correlated with AMMI mother model security, or the increase of disgust expressions correlated with AAP security. Thus when developing our scoring and classification algorithms in the future, it will be important to discard such features, unless we can find theoretical underpinnings for them.

The confirmatory analysis we performed, based on creating a composite effects index, was designed to prove that *overall* the many features studied when put together weighted in the right direction in correlation with attachment security. We argue that this is confirmatory evidence for the construct validity of the BAT, namely that the test activates people in a way that can be captured by multiple modalities and that is correlated with attachment security; but we do not present this as confirmatory evidence for the relationship of any specific feature and attachment security. For example, our study suggests, but does not confirm, that a *breathier* voice can be a watermark of a more securely attached person during an attachment-soothing situation. In this sense, specially associations of LIWC features and attachment security should be seen as merely exploratory since LIWC was not included in our confirmatory analysis (for a rigorous confirmatory study about LIWC features and attachment, see [80]). As for security in the AMMI, which was not significantly correlated with our composite effects index (albeit it was with several individual features), it is important to restate that only a fraction of our sample took that test (*n*=27 for the mother model, *n*=23 for the father model, *n*=17 for the partner model) and since the effect sizes of most features are small, this might explain the lack of association.

In our ANOVA analysis, we decided to include only features that were found to be both statistically significant in their associations to attachment security as well as theory consistent in that association. Why? We compared incredibly different variables in these ANOVAs; from word count on a variety of categories to facial expressions to HRV, and so on, as a reaction to very different image and music combinations. All these

features were calculated without human supervision out of the raw input data, introducing some degree of random error that should favor a finding of difference between answers to different stimuli sets. As a result, odds were stacked in favor of finding differences, and avoiding a multiple comparison correction made it more so as explained above. Including more features (eg, theory-inconsistent features) without controlling for multiple comparisons would generate just too many type I errors for the analysis to be useful.

Future Directions

The BAT was designed to test far more than the attachment security dimension. The different themes in the BAT were designed to also test for the other three main attachment dimensions: attachment anxiety, attachment avoidance [40], and attachment disorganization [12]. They were also designed to measure attachment defenses, such as deactivation, cognitive disconnection, and segregated systems [13]. Different themes heighten differences in the reactions of the four classic adult attachment groups (dismissing, preoccupied, secure, unresolved) to help in classification of attachment. Finally, some themes in the BAT were designed to measure emotional regulation, as measured by instruments such as the Difficulties in Emotion Regulation Scale (DERS [81]), and relational trauma, as measured by the trauma system of the AAP [13]. Empirical validation of the BAT's fitness to measure these constructs is, of course, warranted.

An important area of our work with the BAT is the development of algorithms to automatically score and classify attachment based on extracted features from responses to the test, like the ones highlighted by this study.

This endeavor is complex. It entails finding the right fusion formula for the different BAT features so that the emerging multimodal pattern can accurately predict attachment continuous scores and classifications. It also entails extensive cross-validation to verify the generalizability of the prediction capability to new cohorts. We are underway in this work, and in fact we have developed preliminary regression and classification algorithms capable of predicting ground truth attachment continuous scores and classifications better than chance, cross-validating our results to prevent over-fitting and to warrant generalizability.

Conclusions

Overall, this study brings us one step closer to our goal of developing an automatic and objective adult attachment test. In the future, a 9-min BAT test could be deployed through the Internet to participants or patients residing in remote areas. The test could be scored instantaneously and automatically, with the results becoming available to the researcher or clinician just minutes later. We hope that this could unleash a new wave of attachment research as well as favor clinical attachment testing, in turn benefiting patients by offering them more cost-effective and efficient mental health assessments and treatments.

Acknowledgments

We wish to thank Ville Evrard Clinical Research Unit's Director, Dr Dominique Januel, for her support in the form of financial help to purchase material for this study, as well as granting us access to patients in the Ville Evrard Psychotherapy Center. The work depicted here is sponsored by the US Army Research Laboratory (ARL) under contract number W911NF-14-D-0005. Statements and opinions expressed and content included do not necessarily reflect the position or the policy of the US Government and no official endorsement should be inferred.

Conflicts of Interest

None declared.

References

1. Bowlby J. Attachment and loss. New York: Basic Books; 1969.
2. Cassidy J, Shaver P. editors. Handbook of attachment: Theory, research, and clinical applications. 3rd edition. New York: Guilford Press; 2016.
3. Sroufe LA, Waters E. Attachment as an Organizational Construct. *Child Development* 1977 Dec;48(4):1184. [doi: [10.2307/1128475](https://doi.org/10.2307/1128475)]
4. Mesman J, van IJzendoorn MH, Sagi-Schwartz A. Cross-cultural patterns of attachment: universal and contextual dimensions. In: Cassidy J, Shaver PR, editors. Handbook of Attachment: Theory, Research, and Clinical Applications. New York: Guilford Press; 2016.
5. Grossmann K, Grossmann K, Keppler A. Universal and Culture-Specific Aspects of Human Behavior: The Case of Attachment. In: Friedlmeier W, Chakkarath P, Schwarz B, editors. Culture and Human Development: The Importance of Cross-Cultural Research to the Social Sciences. Hove: Psychology Press; 2005.
6. Ainsworth MD. Attachments across the life span. *Bull N Y Acad Med* 1985 Nov;61(9):792-812 [FREE Full text] [Medline: [3864511](https://pubmed.ncbi.nlm.nih.gov/3864511/)]
7. Sroufe LA, Coffino B, Carlson EA. Conceptualizing the role of early experience: lessons from the minnesota longitudinal study. *Dev Rev* 2010 Mar 1;30(1):36-51 [FREE Full text] [doi: [10.1016/j.dr.2009.12.002](https://doi.org/10.1016/j.dr.2009.12.002)] [Medline: [20419077](https://pubmed.ncbi.nlm.nih.gov/20419077/)]
8. Fraley C, Roisman G. Do early caregiving experiences leave an enduring or transient mark on developmental adaptation? *Curr Opin Psychol* 2015 Feb;1:101-106. [doi: [10.1016/j.copsyc.2014.11.007](https://doi.org/10.1016/j.copsyc.2014.11.007)]

9. De Wolff MS, van IJzendoorn MH. Sensitivity and attachment: a meta-analysis on parental antecedents of infant attachment. *Child Development* 1997 Aug;68(4):571-591. [doi: [10.1111/j.1467-8624.1997.tb04218.x](https://doi.org/10.1111/j.1467-8624.1997.tb04218.x)]
10. Lorber MF, Egeland B. Infancy parenting and externalizing psychopathology from childhood through adulthood: developmental trends. *Dev Psychol* 2009 Jul;45(4):909-912 [FREE Full text] [doi: [10.1037/a0015675](https://doi.org/10.1037/a0015675)] [Medline: [19586169](https://pubmed.ncbi.nlm.nih.gov/19586169/)]
11. Miljkovitch R, Moran G, Roy C, Jaunin L, Forcada-Guex M, Pierrehumbert B, et al. Maternal interactive behaviour as a predictor of preschoolers' attachment representations among full term and premature samples. *Early Hum Dev* 2013 May;89(5):349-354. [doi: [10.1016/j.earlhumdev.2012.11.006](https://doi.org/10.1016/j.earlhumdev.2012.11.006)] [Medline: [23265254](https://pubmed.ncbi.nlm.nih.gov/23265254/)]
12. Miljkovitch R, Moss E, Bernier A, Pascuzzo K, Sander E. Refining the assessment of internal working models: the Attachment Multiple Model Interview. *Attach Hum Dev* 2015;17(5):492-521. [doi: [10.1080/14616734.2015.1075561](https://doi.org/10.1080/14616734.2015.1075561)] [Medline: [26325611](https://pubmed.ncbi.nlm.nih.gov/26325611/)]
13. George C, West M. The adult attachment projective picture system: attachment theory and assessment in adults. New York: Guilford Press; 2012.
14. Mikulincer M, Dolev T, Shaver PR. Attachment-related strategies during thought suppression: ironic rebounds and vulnerable self-representations. *J Pers Soc Psychol* 2004 Dec;87(6):940-956. [doi: [10.1037/0022-3514.87.6.940](https://doi.org/10.1037/0022-3514.87.6.940)] [Medline: [15598116](https://pubmed.ncbi.nlm.nih.gov/15598116/)]
15. Coan J. Towards a neuroscience of attachment. In: Cassidy J, Shaver PR, editors. *Handbook of Attachment: Theory, Research, and Clinical Applications*. New York: Guilford Press; 2016.
16. Hane A, Fox N. Studying the biology of human attachment. In: Cassidy J, Shaver PR, editors. *Handbook of Attachment: Theory, Research, and Clinical Applications*. New York: Guilford Press; 2016.
17. Lemche E, Giampietro VP, Surguladze SA, Amaro EJ, Andrew CM, Williams SC, et al. Human attachment security is mediated by the amygdala: evidence from combined fMRI and psychophysiological measures. *Hum Brain Mapp* 2006 Aug;27(8):623-635. [doi: [10.1002/hbm.20206](https://doi.org/10.1002/hbm.20206)] [Medline: [16284946](https://pubmed.ncbi.nlm.nih.gov/16284946/)]
18. Fraley RC, Spieker SJ. Are infant attachment patterns continuously or categorically distributed? a taxometric analysis of strange situation behavior. *Dev Psychol* 2003;39(3):387-404. [doi: [10.1037/0012-1649.39.3.387](https://doi.org/10.1037/0012-1649.39.3.387)]
19. Grossmann KE, Grossmann K, Waters E. *Attachment from Infancy to Adulthood: The Major Longitudinal Studies*. New York: The Guilford Press; 2005.
20. Carlson EA. A prospective longitudinal study of attachment disorganization/disorientation. *Child Development* 1998;69(4):1107-1128. [doi: [10.1111/j.1467-8624.1998.tb06163.x](https://doi.org/10.1111/j.1467-8624.1998.tb06163.x)]
21. Mikulincer M, Shaver PR. An attachment perspective on psychopathology. *World Psychiatry* 2012 Feb;11(1):11-15. [doi: [10.1016/j.wpsyc.2012.01.003](https://doi.org/10.1016/j.wpsyc.2012.01.003)]
22. Bakermans-Kranenburg MJ, van IJzendoorn MH. The first 10,000 adult attachment interviews: distributions of adult attachment representations in clinical and non-clinical groups. *Attach Hum Dev* 2009 May;11(3):223-263. [doi: [10.1080/14616730902814762](https://doi.org/10.1080/14616730902814762)] [Medline: [19455453](https://pubmed.ncbi.nlm.nih.gov/19455453/)]
23. Cantazaro A, Wei M. Adult attachment, dependence, self-criticism, and depressive symptoms: a test of a mediational model. *J Pers* 2010 Aug 01;78(4):1135-1162. [doi: [10.1111/j.1467-6494.2010.00645.x](https://doi.org/10.1111/j.1467-6494.2010.00645.x)] [Medline: [20545820](https://pubmed.ncbi.nlm.nih.gov/20545820/)]
24. Ein-Dor T, Doron G, Solomon Z, Mikulincer M, Shaver PR. Together in pain: attachment-related dyadic processes and posttraumatic stress disorder. *J Couns Psychol* 2010;57(3):317-327. [doi: [10.1037/a0019500](https://doi.org/10.1037/a0019500)]
25. Levy KN, Meehan KB, Weber M, Reynoso J, Clarkin JF. Attachment and borderline personality disorder: implications for psychotherapy. *Psychopathology* 2005;38(2):64-74. [doi: [10.1159/000084813](https://doi.org/10.1159/000084813)] [Medline: [15802944](https://pubmed.ncbi.nlm.nih.gov/15802944/)]
26. Brown DP, Elliott DS. *Attachment disturbances in adults: Treatment for comprehensive repair*. New York: W. W. Norton & Company; 2016.
27. Mikulincer M, Shaver PR. The psychological effects of the contextual activation of security-enhancing mental representations in adulthood. *Curr Opin Psychol* 2015 Feb;1:18-21. [doi: [10.1016/j.copsyc.2015.01.008](https://doi.org/10.1016/j.copsyc.2015.01.008)]
28. Parnell L, Siegel DJ. *Attachment-focused EMDR: Healing relational trauma*. New York: W. W. Norton & Company; 2015.
29. Mikulincer M, Shaver P. Adult attachment and emotion regulation. In: Cassidy J, Shaver PR, editors. *Handbook of Attachment: Theory, Research, and Clinical Applications*. New York: Guilford Press; 2016.
30. Bergman K, Sarkar P, Glover V, O'Connor TG. Maternal prenatal cortisol and infant cognitive development: moderation by infant-mother attachment. *Biol Psychiatry* 2010 Jun 1;67(11):1026-1032 [FREE Full text] [doi: [10.1016/j.biopsych.2010.01.002](https://doi.org/10.1016/j.biopsych.2010.01.002)] [Medline: [20188350](https://pubmed.ncbi.nlm.nih.gov/20188350/)]
31. Miljkovitch R. *Les fondations du lien amoureux*. Paris: Presses universitaires de France; 2009.
32. Feeney J. Adult romantic attachment: developments in the study of couple relationships. In: Cassidy J, Shaver PR, editors. *Handbook of Attachment: Theory, Research, and Clinical Applications*. New York: Guilford Press; 2016.
33. Rodin G, Walsh A, Zimmermann C, Gagliese L, Jones J, Shepherd FA, et al. The contribution of attachment security and social support to depressive symptoms in patients with metastatic cancer. *Psychooncology* 2007 Dec;16(12):1080-1091. [doi: [10.1002/pon.1186](https://doi.org/10.1002/pon.1186)] [Medline: [17464942](https://pubmed.ncbi.nlm.nih.gov/17464942/)]
34. Ditzen B, Schmidt S, Strauss B, Nater UM, Ehler U, Heinrichs M. Adult attachment and social support interact to reduce psychological but not cortisol responses to stress. *J Psychosom Res* 2008 May;64(5):479-486. [doi: [10.1016/j.jpsychores.2007.11.011](https://doi.org/10.1016/j.jpsychores.2007.11.011)] [Medline: [18440400](https://pubmed.ncbi.nlm.nih.gov/18440400/)]
35. Bartholomew K, Horowitz LM. Attachment styles among young adults: a test of a four-category model. *J Pers Soc Psychol* 1991;61(2):226-244. [doi: [10.1037/0022-3514.61.2.226](https://doi.org/10.1037/0022-3514.61.2.226)]

36. Larose S, Bernier A. Social support processes: mediators of attachment state of mind and adjustment in late adolescence. *Attach Hum Dev* 2001 Apr;3(1):96-120. [doi: [10.1080/14616730010024762](https://doi.org/10.1080/14616730010024762)] [Medline: [11708386](https://pubmed.ncbi.nlm.nih.gov/11708386/)]
37. Powers SI, Pietromonaco PR, Gunlicks M, Sayer A. Dating couples' attachment styles and patterns of cortisol reactivity and recovery in response to a relationship conflict. *J Pers Soc Psychol* 2006 Apr;90(4):613-628. [doi: [10.1037/0022-3514.90.4.613](https://doi.org/10.1037/0022-3514.90.4.613)] [Medline: [16649858](https://pubmed.ncbi.nlm.nih.gov/16649858/)]
38. Wallace JL, Vaux A. Social support network orientation: the role of adult attachment style. *J Soc Clin Psychol* 1993 Sep;12(3):354-365. [doi: [10.1521/jscp.1993.12.3.354](https://doi.org/10.1521/jscp.1993.12.3.354)]
39. Ravitz P, Maunder R, Hunter J, Sthankiya B, Lancee W. Adult attachment measures: a 25-year review. *J Psychosom Res* 2010 Oct;69(4):419-432. [doi: [10.1016/j.jpsychores.2009.08.006](https://doi.org/10.1016/j.jpsychores.2009.08.006)] [Medline: [20846544](https://pubmed.ncbi.nlm.nih.gov/20846544/)]
40. Simpson JA, Rholes WS, Phillips D. Conflict in close relationships: An attachment perspective. *J Pers Soc Psychol* 1996;71(5):899-914. [doi: [10.1037/0022-3514.71.5.899](https://doi.org/10.1037/0022-3514.71.5.899)]
41. Collins NL, Read SJ. Adult attachment, working models, and relationship quality in dating couples. *J Pers Soc Psychol* 1990;58(4):644-663. [doi: [10.1037/0022-3514.58.4.644](https://doi.org/10.1037/0022-3514.58.4.644)]
42. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol* 2003 Oct;88(5):879-903. [doi: [10.1037/0021-9010.88.5.879](https://doi.org/10.1037/0021-9010.88.5.879)] [Medline: [14516251](https://pubmed.ncbi.nlm.nih.gov/14516251/)]
43. Ainsworth MDS, Bell SM. Attachment, exploration, and separation: illustrated by the behavior of one-year-olds in a strange situation. *Child Development* 1970 Mar;41(1):49. [doi: [10.2307/1127388](https://doi.org/10.2307/1127388)]
44. Cassidy J, Sherman LJ, Jones JD. What's in a word? linguistic characteristics of adult attachment interviews. *Attach Hum Dev* 2012 Jan;14(1):11-32. [doi: [10.1080/14616734.2012.636649](https://doi.org/10.1080/14616734.2012.636649)] [Medline: [22191604](https://pubmed.ncbi.nlm.nih.gov/22191604/)]
45. Maxwell SE, Lau MY, Howard GS. Is psychology suffering from a replication crisis? what does "failure to replicate" really mean? *Am Psychol* 2015 Sep;70(6):487-498. [doi: [10.1037/a0039400](https://doi.org/10.1037/a0039400)] [Medline: [26348332](https://pubmed.ncbi.nlm.nih.gov/26348332/)]
46. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. *Speech Commun* 2015 Jul;71:10-49. [doi: [10.1016/j.specom.2015.03.004](https://doi.org/10.1016/j.specom.2015.03.004)]
47. Scherer S, Gale L, Gratch J, Rizzo A, Morency LP. Self-reported symptoms of depression and PTSD are associated with reduced vowel space in screening interviews. *IEEE Trans Affective Comput* 2016 Jan 1;7(1):59-73. [doi: [10.1109/TAFFC.2015.2440264](https://doi.org/10.1109/TAFFC.2015.2440264)]
48. NIMH. 2013. Director's blog: transforming diagnosis URL: <https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2013/transforming-diagnosis.shtml> [accessed 2017-03-14] [WebCite Cache ID 6oxwh0qdc]
49. Scherer S, Stratou G, Lucas G, Mahmoud M, Boberg J, Gratch J, et al. Automatic audiovisual behavior descriptors for psychological disorder analysis. *Image Vis Comput* 2014 Oct;32(10):648-658. [doi: [10.1016/j.imavis.2014.06.001](https://doi.org/10.1016/j.imavis.2014.06.001)]
50. Novak D, Mihelj M, Munih M. A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interact Comput* 2012 May;24(3):154-172. [doi: [10.1016/j.intcom.2012.04.003](https://doi.org/10.1016/j.intcom.2012.04.003)]
51. Lucas G, Gratch J, Scherer S, Boberg J, Stratou G. Towards an affective interface for assessment of psychological distress. 2015 Presented at: International Conference on Affective Computing and Intelligent Interaction (ACII); 2015; Washington, DC p. 539-545.
52. Klagsbrun M, Bowlby J. Responses to separation from parents: a clinical test for young children. *British Journal of Projective Psychology & Personality Study* 1976;21(2):7-27.
53. Baumgartner T, Lutz K, Schmidt CF, Jäncke L. The emotional power of music: how music enhances the feeling of affective pictures. *Brain Res* 2006 Feb 23;1075(1):151-164. [doi: [10.1016/j.brainres.2005.12.065](https://doi.org/10.1016/j.brainres.2005.12.065)] [Medline: [16458860](https://pubmed.ncbi.nlm.nih.gov/16458860/)]
54. Bayevsky RM. Analysis of heart rate variability in space medicine. *Hum Physiol* 2002;28(2):202-213. [doi: [10.1023/A:1014866501535](https://doi.org/10.1023/A:1014866501535)]
55. Tuber S. Understanding personality through projective testing. Lanham: Rowman & Littlefield Publishers; Mar 2014.
56. Schultz DS, Loving JL. Challenges since Wikipedia: the availability of Rorschach information online and internet users' reactions to online media coverage of the Rorschach-Wikipedia debate. *J Pers Assess* 2012;94(1):73-81. [doi: [10.1080/00223891.2011.627963](https://doi.org/10.1080/00223891.2011.627963)] [Medline: [22176268](https://pubmed.ncbi.nlm.nih.gov/22176268/)]
57. Marchewka A, Zurawski L, Jednoróg K, Grabowska A. The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav Res Methods* 2014 Jun;46(2):596-610 [FREE Full text] [doi: [10.3758/s13428-013-0379-1](https://doi.org/10.3758/s13428-013-0379-1)] [Medline: [23996831](https://pubmed.ncbi.nlm.nih.gov/23996831/)]
58. Lang P, Bradley M, Cuthbert B. International affective picture system (IAPS): Technical manual and affective ratings. Gainesville: University of Florida, Center for Research in Psychophysiology; 1999.
59. Dan-Glauser ES, Scherer KR. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behav Res Methods* 2011 Jun;43(2):468-477. [doi: [10.3758/s13428-011-0064-1](https://doi.org/10.3758/s13428-011-0064-1)] [Medline: [21431997](https://pubmed.ncbi.nlm.nih.gov/21431997/)]
60. Zapin M. SurveyGizmo. Why and how to clean your online survey data URL: <https://www.surveygizmo.com/survey-blog/clean-your-online-survey-data/> [accessed 2016-10-27] [WebCite Cache ID 6lYu19NYf]
61. Eerola T, Vuoskoski JK. A comparison of the discrete and dimensional models of emotion in music. *Psychol Music* 2010 Aug 27;39(1):18-49. [doi: [10.1177/0305735610362821](https://doi.org/10.1177/0305735610362821)]

62. Mathôt S, Schreij D, Theeuwes J. OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav Res Methods* 2012 Jun;44(2):314-324 [FREE Full text] [doi: [10.3758/s13428-011-0168-7](https://doi.org/10.3758/s13428-011-0168-7)] [Medline: [22083660](https://pubmed.ncbi.nlm.nih.gov/22083660/)]
63. McCarthy C, Pradhan N, Redpath C, Adler A. Validation of the Empatica E4 wristband. 2016 Presented at: InIEEE EMBS International Student Conference (ISC); 2016; Ottawa.
64. Empatica. 2015. Comparison Procomp Vs Empatica E3 Skin Conductance Signal URL: https://dl.dropboxusercontent.com/u/1421169/ProcompVsE3_final20150309.pdf [accessed 2017-03-14] [WebCite Cache ID 60xxj7DaT]
65. Garbarino M, Lai M, Bender D. Empatica E3 A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. Gent, Belgium: ICST; 2014 Nov Presented at: 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming healthcare through innovations in mobile and wireless technologies (MOBIHEALTH); 2014; Athens, Greece p. 03.
66. Kaufmann T, Sütterlin S, Schulz SM, Vögele C. ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. *Behav Res Methods* 2011 Dec;43(4):1161-1170. [doi: [10.3758/s13428-011-0107-7](https://doi.org/10.3758/s13428-011-0107-7)] [Medline: [21573720](https://pubmed.ncbi.nlm.nih.gov/21573720/)]
67. Task Force of the European Society of Cardiology, The North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 1996;93(5):1043-1065.
68. Benedek M, Kaernbach C. A continuous measure of phasic electrodermal activity. *J Neurosci Methods* 2010 Jun 30;190(1):80-91 [FREE Full text] [doi: [10.1016/j.jneumeth.2010.04.028](https://doi.org/10.1016/j.jneumeth.2010.04.028)] [Medline: [20451556](https://pubmed.ncbi.nlm.nih.gov/20451556/)]
69. Littlewort G, Whitehill J, Wu T, Fasel I, Frank M, Movellan J, et al. The computer expression recognition toolbox (CERT). *Gesture Recognition* 2011:298-305.
70. Amos B, Ludwiczuk B, Satyanarayanan M. ADM. 2016. Openface: A general-purpose face recognition library with mobile applications URL: <http://reports-archive.adm.cs.cmu.edu/anon/anon/2016/CMU-CS-16-118.pdf> [accessed 2017-03-14] [WebCite Cache ID 60xyEWbGa]
71. Degottex G, Kane J, Drugman T, Raitio T, Scherer S. COVAREP: A collaborative voice analysis repository for speech technologies. 2014 Presented at: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2014; Florence.
72. Piolat A, Booth R, Chung C, Davids M, Pennebaker J. La version française du dictionnaire pour le LIWC : modalités de construction et exemples d'utilisation. *Psychologie Française* 2011 Sep;56(3):145-159. [doi: [10.1016/j.psfr.2011.07.002](https://doi.org/10.1016/j.psfr.2011.07.002)]
73. Bender R, Lange S. Adjusting for multiple testing-when and how? *J Clin Epidemiol* 2001 Apr;54(4):343-349. [doi: [10.1016/S0895-4356\(00\)00314-0](https://doi.org/10.1016/S0895-4356(00)00314-0)]
74. Anderson ML. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the abecedarian, perry preschool, and early training projects. *J Am Stat Assoc* 2008 Dec;103(484):1481-1495. [doi: [10.1198/016214508000000841](https://doi.org/10.1198/016214508000000841)]
75. Kling JR, Liebman JB, Katz LF, Sanbonmatsu L. Princeton. 2004. Moving to opportunity and tranquility: neighborhood effects on adult economic self-sufficiency and health from a randomized housing voucher experiment URL: <http://dataspace.princeton.edu/jspui/handle/88435/dsp01qz20ss50t> [accessed 2017-03-14] [WebCite Cache ID 60xyPqYa3]
76. Rothman KJ. Six persistent research misconceptions. *J Gen Intern Med* 2014 Jul;29(7):1060-1064 [FREE Full text] [doi: [10.1007/s11606-013-2755-z](https://doi.org/10.1007/s11606-013-2755-z)] [Medline: [24452418](https://pubmed.ncbi.nlm.nih.gov/24452418/)]
77. Pascovici D, Handler DC, Wu JX, Haynes PA. Multiple testing corrections in quantitative proteomics: a useful but blunt tool. *Proteomics* 2016 Sep;16(18):2448-2453. [doi: [10.1002/pmic.201600044](https://doi.org/10.1002/pmic.201600044)] [Medline: [27461997](https://pubmed.ncbi.nlm.nih.gov/27461997/)]
78. Feise RJ. Do multiple outcome measures require p-value adjustment? *BMC Med Res Methodol* 2002 Jun 17;2(1):1450. [doi: [10.1186/1471-2288-2-8](https://doi.org/10.1186/1471-2288-2-8)]
79. Burr RL. Interpretation of normalized spectral heart rate variability indices in sleep research: a critical review. *Sleep* 2007;30(7):913-919.
80. Waters TE, Steele RD, Roisman GI, Haydon KC, Booth-LaForce C. A linguistic inquiry and word count analysis of the adult attachment interview in two large corpora. *Can J Behav Sci* 2016 Jan 01;48(1):78-88 [FREE Full text] [doi: [10.1037/cbs0000035](https://doi.org/10.1037/cbs0000035)] [Medline: [27065477](https://pubmed.ncbi.nlm.nih.gov/27065477/)]
81. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *J Psychopathol Behav Assess* 2004 Mar;26(1):41-54. [doi: [10.1023/B:JOBA.0000007455.08539.94](https://doi.org/10.1023/B:JOBA.0000007455.08539.94)]

Abbreviations

AAI: Adult Attachment Interview
AAP: Adult Attachment Projective Picture System
AAQ: Adult Attachment Questionnaire
AAS: Adult Attachment Scale
AMMI: Attachment Multiple Model Interview
ANOVA: analysis of variance
ARL: Army Research Laboratory
BAT: Biometric Attachment Test

COVAREP: Cooperative Voice Analysis Repository for Speech Technologies
DEERS: Difficulties in Emotion Regulation Scale
EDA: electrodermal activity
GAPED: Geneva Affective Picture Database
HF: high frequency
HR: heart rate
HRV: heart rate variability
IAPS: International Affective Picture System
IBI: interbeat interval
LF: low frequency
LIWC: Linguistic Inquiry and Word Count
NAPS: Nencki Affective Picture System
PTSD: posttraumatic stress disorder
SAT: Separation Anxiety Test
SCR: skin conductance response
SPT: Standard Penetration Test
SSP: Strange Situation Procedure
VLF: very low frequency

Edited by R Calvo; submitted 27.10.16; peer-reviewed by A Jaroszewski, D Messinger; comments to author 28.11.16; revised version received 27.01.17; accepted 05.03.17; published 06.04.17.

Please cite as:

Parra F, Miljkovitch R, Persiaux G, Morales M, Scherer S

The Multimodal Assessment of Adult Attachment Security: Developing the Biometric Attachment Test

J Med Internet Res 2017;19(4):e100

URL: <http://www.jmir.org/2017/4/e100/>

doi: [10.2196/jmir.6898](https://doi.org/10.2196/jmir.6898)

PMID: [28385683](https://pubmed.ncbi.nlm.nih.gov/28385683/)

©Federico Parra, Raphaële Miljkovitch, Gwenaëlle Persiaux, Michelle Morales, Stefan Scherer. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 06.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

An Analysis of 2.3 Million Participations in the Continuing Medical Education Program of a General Medical Journal: Suitability, User Characteristics, and Evaluation by Readers

Hildegard Christ¹, PhD; Jeremy Franklin¹, PhD; Reinhard Griebenow², MD; Christopher Baethge³, MD

¹Institute of Medical Statistics, Informatics and Epidemiology, University of Cologne, Cologne, Germany

²Municipal Hospital Cologne-Merheim, University of Cologne, Cologne, Germany

³Deutsches Ärzteblatt, Editorial Offices, Cologne, Germany

Corresponding Author:

Hildegard Christ, PhD

Institute of Medical Statistics, Informatics and Epidemiology

University of Cologne

Kerpener Str 62

Cologne, 50937

Germany

Phone: 49 221 478 6502

Fax: 49 221 478 6520

Email: Hildegard.Christ@uni-koeln.de

Abstract

Background: Physicians frequently use continuing medical education (CME) in journals. However, little is known of the evaluation of journal CME by readers and also user and participation characteristics. Deutsches Ärzteblatt, the journal of the German Medical Association, is distributed to every physician in Germany and regularly offers its readers CME articles. Therefore, it provides a unique opportunity to analyze a journal CME program directed at an entire population of physicians.

Objective: The aim is to show key sociodemographic characteristics of participants, frequency and temporal distributions of participations, and to analyze whether the articles are suitable for a general medical audience, how physicians rate the CME articles, how successful they were in answering simple multiple-choice questions, and to detect distinct clusters of participants.

Methods: Using obligatory online evaluation forms and multiple-choice questions, we analyzed all participations of the entire 142 CME articles published between September 2004 and February 2014. We compared demographic characteristics of participants with official figures on those characteristics as provided by the German Medical Association.

Results: A total of 128,398 physicians and therapists (male: 54.64%, 70,155/128,393; median age class 40 to 49 years) participated 2,339,802 times (mean 16,478, SD 6436 participations/article). Depending on the year, between 12.33% (44,064/357,252) and 16.15% (50,259/311,230) of all physicians in the country participated at least once. The CME program was disproportionately popular with physicians in private practice, and many participations took place in the early mornings and evenings (4544.53%, 1,041,931/2,339,802) as well as over the weekend (28.70%, 671,563/2,339,802). Participation by specialty (ranked in descending order) was internal medicine (18.25%, 23,434/128,392), general medicine (16.38%, 21,033/128,392), anesthesiology (10.00%, 12,840/128,392), and surgery (7.06%, 9059/128,392). Participants rated the CME articles as intelligible to a wider medical audience and filling clinically relevant knowledge gaps; 78.57% (1,838,358/2,339,781) of the sample gave the CME articles very good or good marks. Cluster analysis revealed three groups, one comprised of only women, with two-thirds working in private practice.

Conclusions: The CME article series of Deutsches Ärzteblatt is used on a regular basis by a considerable proportion of all physicians in Germany; its multidisciplinary articles are suitable to a broad spectrum of medical specialties. The program seems to be particularly attractive for physicians in private practice and those who want to participate from their homes and on weekends. Although many physicians emphasize that the articles address gaps in knowledge, it remains to be investigated how this impacts professional performance and patient outcomes.

(*J Med Internet Res* 2017;19(4):e49) doi:[10.2196/jmir.6052](https://doi.org/10.2196/jmir.6052)

KEYWORDS

education; medical; continuing interactive tutorial; journal article

Introduction

Proof of continuing medical education (CME) has become mandatory for physicians in many countries across the world. For example, in Germany proof of CME is required by law for those who finished residency training. As a result, medical specialists have to show that they continue learning.

The type of CME through which physicians in Germany satisfy this requirement is largely left open to the physicians themselves. Certified courses offered by various institutions (eg, hospitals, medical associations, private providers) are wide-ranging in form. They include lectures, symposia, sitting-in at a clinic, workshops, and structured interactive courses via print and online media, as defined by the Model Regulations on Continuous Medical Education and Certification of the German Medical Association. The Chambers of Physicians accredit a wide range of CME activities [1].

Although data on the “CME mix” (live events, print media, e-learning) as practiced by the individual physician are lacking, there are some data on the use of CME modalities in general on the national level. In 2015, all German Chambers together accredited approximately 360,000 CME activities. CME in print media constitute approximately 1% of all CME activities. However, this type of CME generated approximately 20% of all CME points earned by German physicians, demonstrating wide acceptance of this print article-based CME. For comparison, live CME events represented approximately 95% of all activities accredited, but generated approximately 70% of all CME points. This situation has shown only minor variations since these data have been recorded in 2006 (written personal communication, Reinhard Griebenow, Member of the Senate for medical education of the German Medical Association, August 2016). Accordingly, many journals in Germany offer CME articles. In 2014, there were 75 medical journals with an impact factor, and CME activities were offered in 60% of these.

International data are scarce, but in the United States where CME activity is an educational offer that is planned, implemented, and evaluated in accordance with the Accreditation Council for Continuing Medical Education (ACCME) Accreditation Criteria, journal CME represented 4% of all CME courses without commercial support. However, journal CME contributed 9% to all physician interactions related to CME in 2014 [2].

Potentials and Challenges of Journal Continuing Medical Education

Offering CME in print media has several advantages. In contrast to almost all other forms of CME, print articles are often peer reviewed. Also, print CME is accessible independent of time and place, and a wide variety of courses with considerable content diversity can be offered. The challenges facing print CME include very diverse (and sometimes unknown) backgrounds of participants, which makes it more difficult to

tailor the course to the participants' abilities and interests. Also, feedback and the possibility to raise questions from participants are lacking or delayed. Similarly, discussion among participants is restricted, although online lists and communities can be a substitute for direct interaction. Finally, there is less control over evaluation of the participants' success.

Accredited CME in print media consists of an article accompanied by 10 multiple-choice questions, of which at least seven have to be answered correctly (see [Multimedia Appendix 1](#)). For quality control, readers are asked to give feedback on several aspects of the articles (eg, presence of a current treatment strategy on the part of the participant, comprehensibility of the paper, completeness). In addition, participants are asked to provide basic demographic data and information regarding their specialty and work setting. It is mandatory for participants to submit these data to receive CME credits.

So far, studies analyzing print CME programs on the basis of such or similar data have mainly been related to specialist journals with CME units addressing a certain group of qualified specialist physicians. They are also based on a limited number of participants. For instance, in 2004 [3] and in 2005 [4] participations in CME offered in specialty journals were evaluated in two doctoral theses. Participants gave largely positive ratings, including the appropriateness of the content for their specialty.

In contrast to specialty journals, a comprehensive analysis of a CME print program addressing the general medical audience is lacking. *Deutsches Ärzteblatt* (Christopher Baethge is chief scientific editor), the journal of the German Medical Association, offers an ideal setting for such a study because of its interdisciplinary nature and its unique geographical coverage: all physicians in Germany receive a copy of *Deutsches Ärzteblatt*. For international readers, English translations of all CME, original, and review articles, as well as all editorials and letters to the editors, are provided in *Deutsches Ärzteblatt International*, the global edition of *Deutsches Ärzteblatt*.

In *Deutsches Ärzteblatt*, 142 CME articles have been published since the start of accredited CME in *Deutsches Ärzteblatt* in September 2004 until February 2014. There was a mean of 16,000 participants per unit. Because personal details have to be provided to receive CME points, it can be assumed that data have been conscientiously and accurately recorded. Furthermore, this provides a good opportunity to analyze the characteristics and evaluations of the physicians. A pilot study showed the feasibility of such an analysis, although the data basis was small (only 37 CME articles were evaluated) [5-7].

Accordingly, this investigation aimed to describe the personal, professional, and participation characteristics of all participants in the CME program of the *Deutsches Ärzteblatt* between 2004 and 2014. Specifically, we aimed to (1) describe the participants in terms of age, gender, and professional position; (2) describe the success rate and the evaluation of the scheme by the participants; (3) estimate the percentage of German medical

professionals using this CME scheme; and (4) identify participant subgroups using cluster analysis and to assess the frequency and temporal distribution of participation.

Methods

This is a retrospective analysis of primary data based on the records of 142 accredited CME articles in *Deutsches Ärzteblatt*. Because membership in a Chamber of Physicians is mandatory for all licensed physicians in Germany and the subscription fee for *Deutsches Ärzteblatt* is covered by the membership fee for the Chamber, every German physician receives *Deutsches Ärzteblatt* for free, resulting in a circulation of approximately 450,000 copies as of the first quarter of 2015.

Topics and authors of the CME articles are selected by the scientific editorial board of *Deutsches Ärzteblatt*. All articles have to pass peer review and the accreditation process of the Chamber of Physicians North-Rhine, which demands that not only the content of the article, but also the article pages, have to be free of any commercial influence and advertisements. Terms used in the German health care system are explained in [Multimedia Appendix 2](#).

A CME unit consists of an article and a corresponding knowledge test with 10 multiple-choice questions, in which only one of five answers is correct (with a pass mark of 70%). The articles can be read in the printed edition of the journal or on the Internet, but test participation always takes place online. During the period of this study, each test was available online for six weeks (since then it has been extended to 12 weeks). In addition to the assessment of knowledge, all participants had to answer the following evaluation form:

1. Time required to finish the course: up to 30 minutes / 31 to 45 minutes / 46 to 60 minutes / more than 60 minutes.
2. Prevalence of the topic in everyday clinical practice: "The course topic arises in my professional activities frequently / regularly / seldom / never / not relevant for me."
3. Presence of a treatment strategy for the clinical problem addressed in the article: "For the diagnosis and treatment of the disease concerned, I had already before my involvement with the course: a fixed treatment regimen / unsolved problems / no fixed strategy / not relevant for me."
4. Completeness of information: "From the point of view of daily practice, important aspects of the course topic were not mentioned / treated too briefly / overemphasized / very well presented / not relevant for me."
5. Comprehensibility of the article: "The article is comprehensible to specialists only / comprehensible to all physicians."
6. Difficulty of the test questions asked: "The questions can be answered by studying the article only / only with additional literature."

7. Overall satisfaction with the course (1=very good; 6=insufficient).

In addition, the date and time of participation was recorded. The data were stored in a database by an independent data host. For comparison with official figures, data on demographic characteristics and on medical specialties were taken from official statistics of the German Medical Association (Bundesärztestatistik).

Data analysis was performed using the statistical software SPSS version 22.0.0. Nominal and ordinal data are presented in tables using frequencies and percentages, contingency tables or bar charts. To identify groups of participants, a two-step cluster analysis (maximum: 4 clusters; distance measure: loglikelihood; cluster criterion: Schwarz-Bayes criterion) was performed. In view of the large sample size and because this is a descriptive investigation, *P* values are not presented.

Results

Participants

Of the 128,398 participants 98.36% (126,302/128,398) were active physicians. Per year, between 12.33% (44,064/357,252) in 2013 and 16.15% (50,259/311,230) in 2006 of all licensed physicians had participated [8].

The participants' ages were distributed as follows: 5.78% (7420/128,393) were younger than 30 years; 21.34% (27,398/128,393) were aged between 30 and 39 years; 32.79% (42,096/128,393) were aged between 40 and 49 years; 28.51% (36,610/128,393) were between 50 and 59 years; and 11.58% (14,869/128,393) were 60 years or older. The age distribution showed a declining proportion of women with increasing age (see [Figure 1](#)).

The largest fraction of participants worked in private practice, with a larger proportion of men than women (8% more, see [Figure 2](#)). In all, 40.56% (52,069/128,390) of participants were employed in hospitals, with fewer women in senior positions (9.04%, 5265/58,237 vs 20.75%, 14,561/70,152). The proportion of residents who were female was higher (31.68%, 18,448/58,237 vs 19.66%, 13,795/70,152).

The most frequent specialty of all participating physicians was internal medicine with 18.25% (23,434/128,392), followed by general medicine (16.38%, 21,033/128,392), anesthesiology (10.00%, 12,840/128,392), and surgery (7.06%, 9059/128,392). In the cluster analysis, three groups of participants could be identified. The variables age, gender, position in the health care system, specialty, number of participations, median handling time (article and multiple-choice questions, hours), and mean number of CME points earned per course were employed. Age and gender emerged as the most strongly heterogeneous variables for characterizing the clusters (see [Table 1](#)).

Table 1. Cluster analysis of the participants.

Characteristics	Cluster 1	Cluster 2	Cluster 3
Participants, n (%)	55,101 (42.92)	37,873 (29.50)	35,415 (27.58)
Gender, n (%)			
Male	55,101 (100)	0 (0)	15,051 (42.50)
Female	0 (0)	37,873 (100)	20,364 (57.50)
Age (years), n (%)			
<40	5249 (9.53)	5952 (15.72)	23,616 (66.68)
40-60	39,235 (71.21)	28,354 (74.87)	11,115 (31.39)
Position in health care system, n (%)			
Outpatient care	34,821 (63.19)	24,606 (64.97)	329 (0.93)
Senior specialist	11,730 (21.29)	4942 (13.05)	57 (0.16)
Other areas of health care	420 (8.02)	7278 (19.22)	1437 (4.06)
Residents	1 (0.00)	0 (0)	33,531 (94.63)
Specialty, n (%)			
Internal medicine	9858 (17.89)	4717 (12.45)	8859 (25.01)
General medicine	9988 (18.13)	8214 (21.69)	2831 (7.99)
Anesthesiology	4411 (8.01)	2937 (7.75)	5491 (15.50)
Surgery	4490 (8.15)	983 (2.60)	3586 (10.13)
Participations, mean (SD)	20.00 (27.20)	21.26 (27.48)	12.21 (19.54)

One cluster consisted of male physicians, mainly aged between 40 and 60 years, with 63.19% (34,821/55,101) working in outpatient care and 21.29% (11,730/55,101) being senior specialists.

A further cluster of exclusively female physicians had a similar composition, but with a lower fraction of senior specialists; instead, nearly one-fifth (19.22%, 7278/37,873) were occupied in other institutions.

The third cluster consisted almost exclusively of residents, 66.68% (23,616/35,415) being younger than 40 years and more than half (57.50%, 20,364/35,415) of them were women.

The mean number of course participations in cluster 1 was mean 21.26 (SD 27.48), in cluster 2 mean 21.26 (27.48), and in cluster 3 mean 12.21 (SD 19.54).

Using the example from 2013, a comparison of the most important demographic data of the participants with the active physicians of Germany is given. Of the 357,252 active physicians in Germany, 44,064 (12.33%) were participants [8]. The proportion of women among CME participants (46.46%, 20,471/44,064) was very similar to that among all physicians in Germany (45.03%, 160,869/357,252). The proportion of those younger than 35 years among CME participants (7.92%, 3489/44,063) was less than half of that among all physicians in Germany (18.01%, 64,355/357,252; see [Figure 3](#)), whereas the participants aged between 40 and 60 years were overrepresented (67.56%, 29,772/44,063 vs 54.86%, 195,983/357,252).

The proportion employed in inpatient care was 14.68% lower among CME participants (35.69%, 15,728/44,064) compared with all physicians in Germany (50.67%, 181,012/357,252; see [Figure 4](#)).

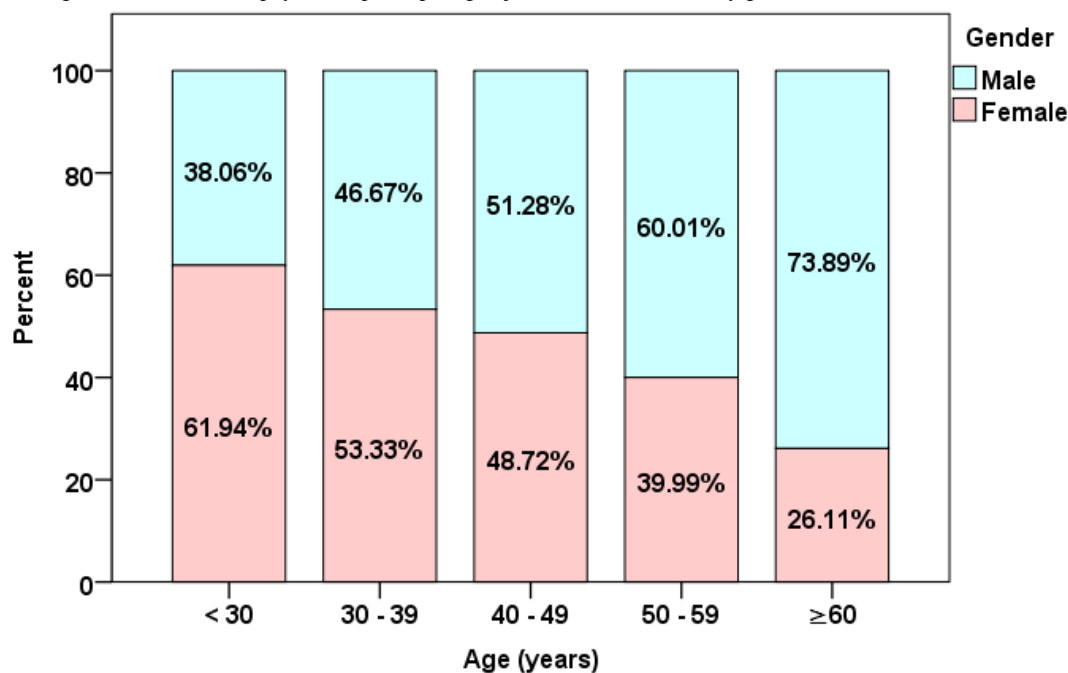
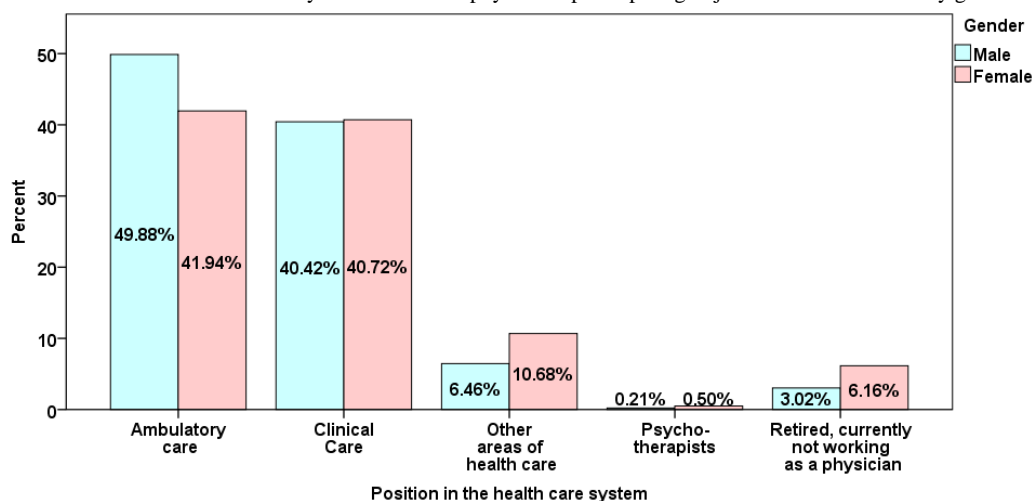
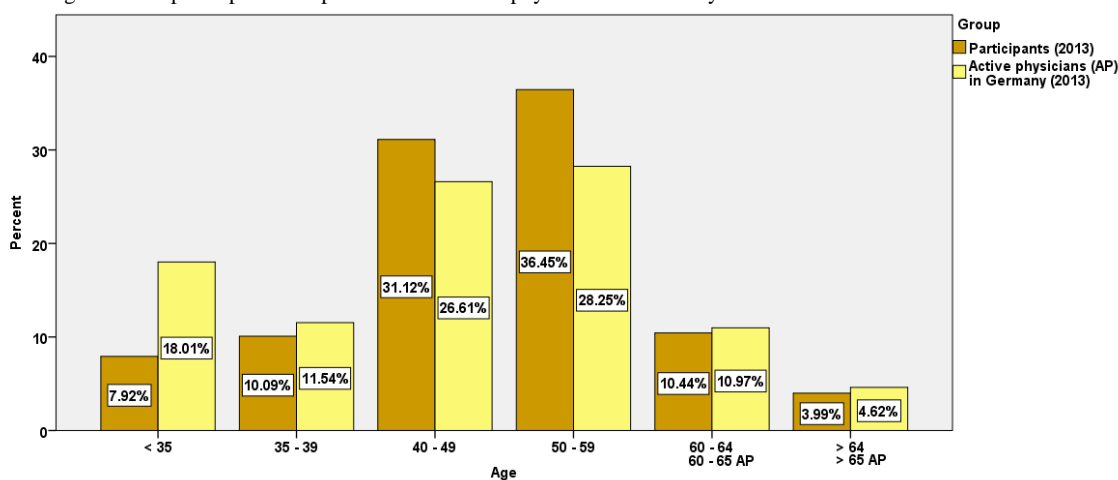
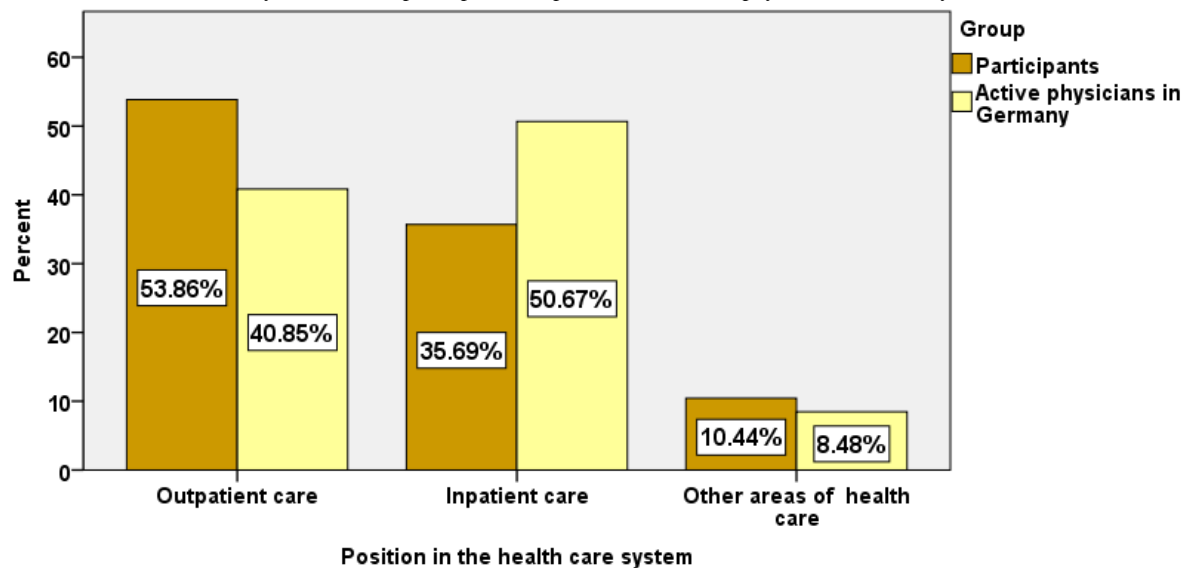
Figure 1. Age classes of German physicians participating in journal CME activities by gender.**Figure 2.** Positions in the health care system of German physicians participating in journal CME activities by gender.**Figure 3.** Age of CME participants compared with all active physicians in Germany.

Figure 4. Position in health care system of CME participants compared with all active physicians in Germany.**Table 2.** Articles with highest and lowest numbers of participants between September 2004 and February 2014.

Article	Participants, (n)	Overall satisfaction, mean (SD)	Questions correctly answered, mean (SD)
Top ranking			
“Lyme borreliosis: current state of knowledge”	25,886	1.75 (0.68)	9.83 (0.46)
“The acute abdomen from the internists’ point of view”	25,547	1.87 (0.68)	9.86 (0.46)
“Key symptom diarrhea”	25,494	2.21 (0.83)	9.32 (0.83)
Lowest ranking			
“Asperger syndrome: a disorder of the autism spectrum”	2407	2.17 (0.86)	8.98 (1.26)
“Unwanted weight loss in the elderly”	2401	2.05 (0.82)	8.96 (0.85)
“Obligatory reporting of infectious diseases”	2341	2.01 (0.85)	8.95 (1.06)

Course Participations

The number of participants in each of the 142 CME courses published between September 2004 and February 2014 ranged from 2341 to 25,886 (mean 16,477, SD 6436). Table 2 displays the articles with the highest and lowest numbers of participants, respectively.

During this period there were 2,339,802 participations by 128,398 physicians and therapists (female: 45.68%, 1,068,559/2,339,032; male: 54.32%, 1,270,473/2,339,032) whose courses had been evaluated and whose data could therefore be analyzed.

The analysis of number of participations considered both inter- and intraindividual differences because 81.09%

(104,122/128,398) of the participants took part more than once. For instance, both the workloads of the various participants in a course unit and the workloads of the same participant in various course units were analyzed.

Concerning participation frequency, approximately one-third (30.00%, 38,522/128,398) of participants each read 1 to 2, 3 to 10 (29.55%, 37,947/128,398), and 11 to 50 articles (29.15%, 37,422/128,398), respectively; 11.30% (14,507/128,398) of participants read more than 50 articles. Of those participating 1 to 10 times, 29.47% (11,352/38,514) were residents versus 19.03% (9883/51,929) of those participating more than 10 times.

Participations were recorded for all days of the week, although most frequently on Sundays and Mondays (see Table 3).

Table 3. Days of participations in journal CME activities.

Day	n (%)
Monday	394,232 (16.85)
Tuesday	345,400 (14.76)
Wednesday	334,516 (14.30)
Thursday	298,312 (12.75)
Friday	295,779 (12.64)
Saturday	244,447 (10.45)
Sunday	427,116 (18.25)

Although the great majority of participations (90.25%, 2,111,675/2,339,802) took place between 09:00 and 22:00 hours, 3.66% (85,702/2,339,802) occurred between 23:00 and 24:00 hours, 1.18% (27,809/2,339,802) between 01:00 and 06:00 hours, and 4.89% (114,616/2,339,802) between 07:00 and 08:00 hours (see [Figure 5](#)).

The temporal distribution of participations mirrors the general working times of the various sectors of health care. The proportion of physicians working in hospitals or being retired was larger at nighttime. Physicians not active in patient care tended to participate from 06:00 to 09:00.

More than three-quarters of participants (78.56%, 1,838,358/2,339,781) ranked the CME article as 1 or 2 (1=very good, 6=insufficient) and the knowledge test was passed by 99.07% (2,318,040/2,339,781) of participants (see [Table 4](#)). Between participations with ratings of 1 and 6, there was a difference of 6 percentage points in pass rates.

The materials were rated as “intelligible for all physicians, not only for specialists” in 95.65% (2,236,963/2,338,664) of participations. In 92.36% (2,156,478/2,334,986) of participations, the knowledge questions could be answered by studying the article materials only. However, for psychological psychotherapists (0.16%, 3677/2,338,664 of all participants), this number was lower (84.96%, 3124/3677).

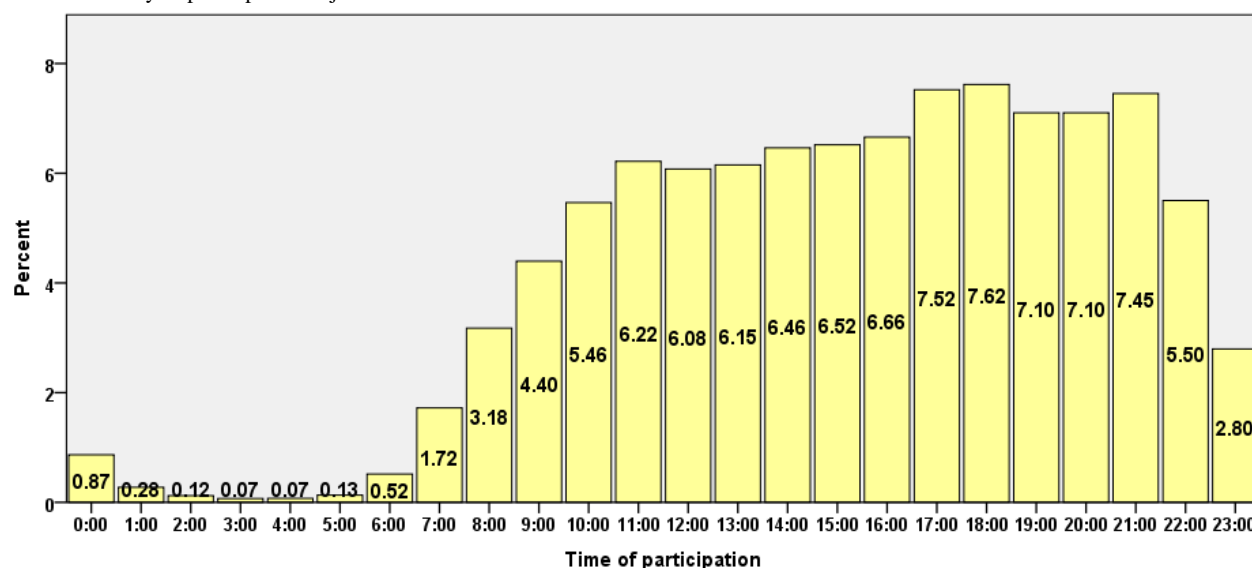
Figure 5. Time of day of participation in journal CME activities.

Table 4. Overall satisfaction and pass rates of participants taking CME courses.

Overall satisfaction ^a	Frequency, n (%)	Passed, n (%)
1	527,951 (22.56)	525,622 (99.56)
2	1,310,407 (56.01)	1,301,059 (99.29)
3	377,073 (16.12)	370,874 (98.36)
4	87,803 (3.75)	85,563 (97.45)
5	27,633 (1.18)	26,575 (96.17)
6	8914 (0.38)	8347 (93.64)

^a 1=very good; 6=insufficient.

The question “The topic occurs in my clinical routine” was answered “frequently” in 16.66% (389,624/2,339,049) of participations, “regularly” in 21.99% (514,296/2,339,049), “rarely” in 39.17% (916,280/2,339,049), “irrelevant” in 13.71% (32,0724/2,339,049), and “never” in 8.74% (198,125/2,339,049) of participations (see [Figure 6](#)).

In 25.61% (599,046/2,339,175) of participations, physicians had already “an established treatment strategy for the disease concerned” before reading the article, 32.88% (769,185/2,339,175) had “unresolved problems,” 20.44% (478,175/2,339,175) had “no established strategy,” and for

21.07% (492,769/2,339,175) the question was irrelevant. Physicians in private practice, senior specialists, and principal physicians showed the highest percentages for having a definite treatment regimen (see [Figure 7](#)).

In 70.27% (1,643,825/2,339,275) of the participations, important issues were treated “very well” from the point of view of daily practice. Exclusion of participations that rated the content of the article as irrelevant for their medical practice increased this value to 84.44% (1,643,825/1,946,694). Physicians in private practice, principal physicians, senior specialists, and retired physicians gave higher ratings (see [Figure 8](#)).

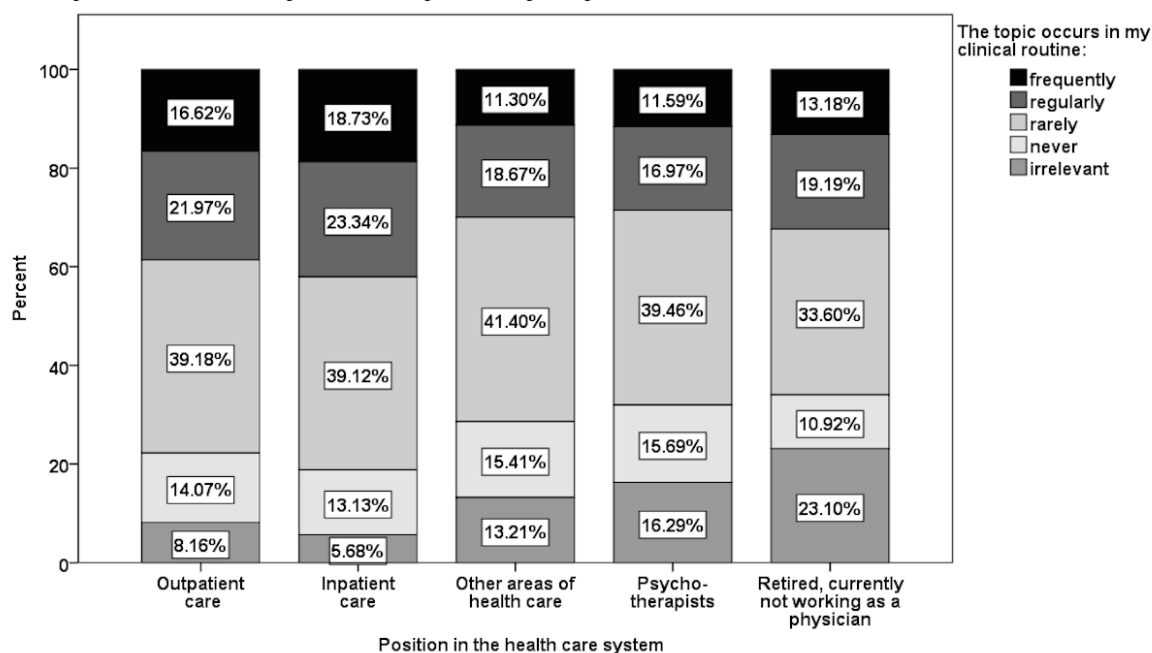
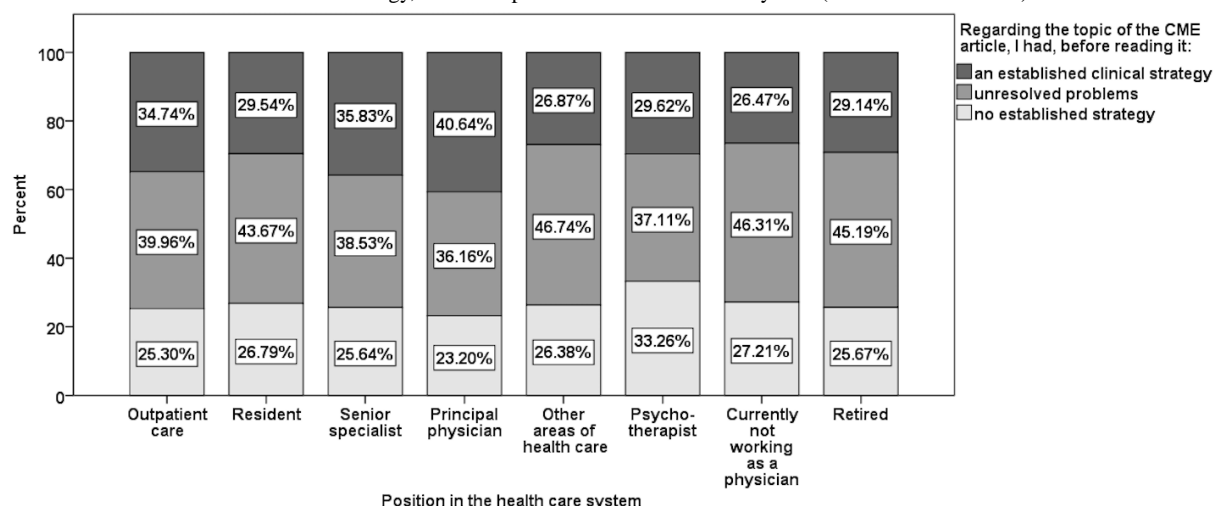
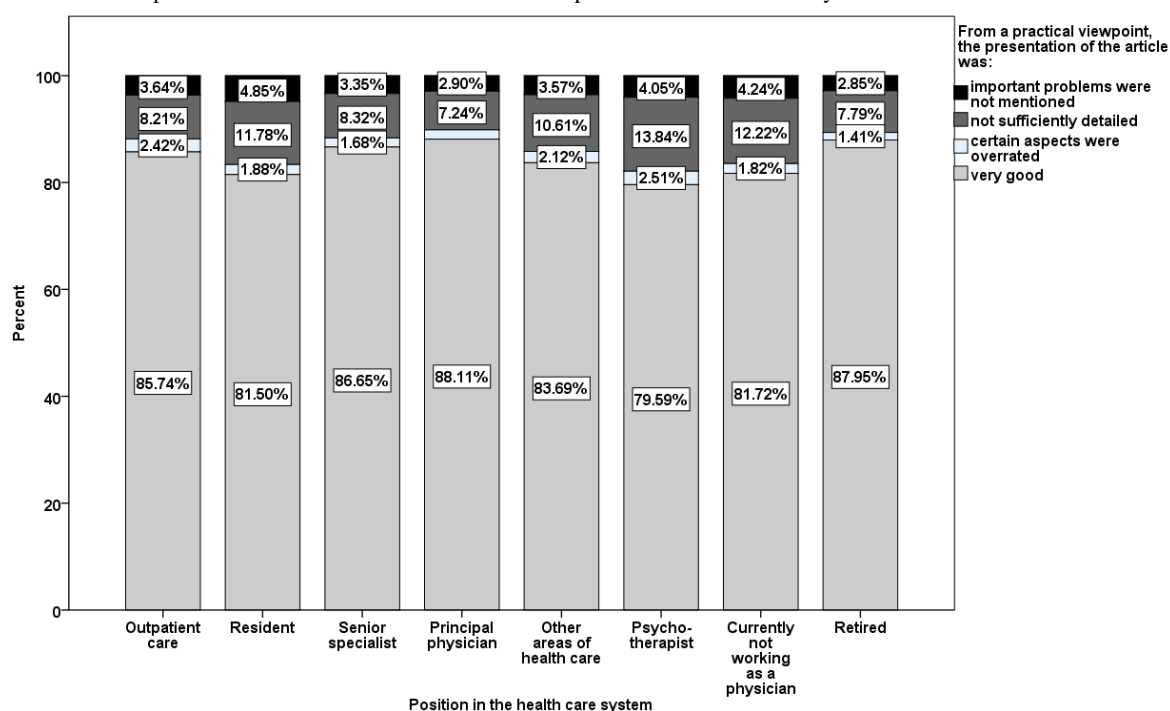
Figure 6. Representation of CME topics in clinical practice of participants.

Figure 7. Presence of established clinical strategy, related to position in the health care system (“irrelevant” excluded).**Figure 8.** Balance of presentation of CME article content related to position in the health care system.

Discussion

This study yielded the following important results:

1. Interdisciplinary CME articles in *Deutsches Ärzteblatt* attract a significant percentage of all German physicians and were positively evaluated by the participants.
2. The time of day and day of week when participants worked on the articles as well as judgments on content indicate that CME in print media meets the needs of self-directed learning.
3. Three groups of participants in interdisciplinary CME articles in *Deutsches Ärzteblatt* could be identified, indicating distinct target audiences of print CME.

Participants and Participations

The high participation rate in the CME articles of *Deutsches Ärzteblatt* supports the idea that print CME is an important pillar among CME programs. With regard to *Deutsches Ärzteblatt*, in any given year, approximately 15% of all licensed physicians take part. Moreover, most participants take part repeatedly, with one-third participating between 11 and 50 times over the years. Therefore, print CME contributes considerably to the sum of CME credits awarded to physicians obliged to prove continuing education. Of note, not only physicians obliged to collect CME credits take part: in our analysis, we found a cluster almost entirely consisting of residents, a group of physicians that under German law is exempt from collecting CME credits because residents are considered to be in a constant process of learning. Although the cluster of residents is the smallest cluster in our study, roughly one in four participants were grouped there,

indicating that continuing education for physicians is not entirely driven by legal obligations.

Among a broad range of participants, we were able to identify two other well-defined groups, with age and gender being key characteristics. One cluster consisted of male physicians mainly in outpatient care, and a second cluster comprised of female physicians occupied in various fields (outpatient care, other areas of health care). It is obvious that the clusters are not representative of the entirety of German physicians; for example, older and younger physicians are underrepresented among our participants.

Evaluations and Knowledge Tests

In general, the participants' evaluations demonstrated a high degree of satisfaction with the articles; almost 80% rated them as good or very good. For 70% of all participations, important aspects of the topic were considered to have been very well presented. In a similar vein, in 96% of all participations readers found the content understandable even to physicians not belonging to the specialty mainly addressed in the article. It is reassuring that in the majority of participations the topic of the article occurred at least rarely, and often regularly and frequently, in participants' everyday work and that many participants reported unresolved problems in dealing with the clinical topics. The high marks for satisfaction and understandability show that by a well-defined editorial process it is possible to deliver useful CME articles on a long-term basis.

The knowledge test was passed in 99% of participations, and in more than 92% the questions could be answered without use of any other additional material than article content. In Germany, knowledge tests are considered as formative and supporting the main messages of the article rather than reflecting the whole content of the CME article; they are not intended to be exams in the strict sense [9]. Nevertheless, knowledge tests add an element of quality assurance unique (in Germany) to CME in print media.

Meeting the Learning Needs of Physicians

Our data further demonstrate that CME in print media meets the needs of self-directed learning not only with regard to content, but also to day and time of "attendance." Articles were read on all days of the week, most frequently on Sunday. Further, almost 40% of participations occur before 09:00 or after 18:00, which clearly distinguishes them from live meetings. These data suggest that flexibility regarding location and hours is one of the main advantages of the CME program in print media. Therefore, if compatibility of family and work remains important or becomes even more important in medicine, print CME certainly has an important place among the variety of CME programs.

Among hospital physicians, articles are read less frequently by residents than by senior specialists and principal physicians, reflecting the fact that only specialists are obliged by law to document their CME activities. This is further supported by the fact that 46% of all participants worked in outpatient care, whereas hospital physicians were underrepresented compared to their nationwide representation [8].

Factors Contributing to a Successful Print Continuing Medical Education Program

Our experience with CME at the *Deutsches Ärzteblatt* suggests that the following factors are important for a successful print CME program: topic diversity, comprehensibility, active intellectual involvement (hands on approach), regularity, easy-to-reach hotline and maintenance, and well-crafted multiple-choice questions that are fair and achievable. The commitment needed on the part of the journal is substantial: the time and effort involved in producing a CME article in *Deutsches Ärzteblatt* far exceeds that of a normal narrative review article. Producing CME articles encompasses, for example, the selection of authors and introducing authors into the specifics of CME articles (eg, formulating learning objectives, take-home messages, or multiple-choice questions). For educational material, both editorial and peer review need to be particularly thorough, and all multiple-choice questions have to be pretested. In total, we estimate that, compared to an average review article, CME articles take twice as much time and effort to publish. As a result, we devote approximately one-sixth of our staff resources to CME alone.

Limitations

This investigation is limited to participants of the *Deutsches Ärzteblatt* CME scheme from 2004 to 2014. The characteristics of nonparticipants and the reasons for nonparticipation lie outside our scope. It would be interesting to see whether there are certain groups of physicians that cannot not be convinced to take part in print CME and what could be done to accommodate their needs. One obvious disadvantage of print CME is its lack of interaction between experts and participants. Therefore, in the future we are planning to offer an online question-and-answer format. Another problem of print CME regards the test format; although other forms of knowledge evaluation, such as interviews, are superior to multiple-choice questions in many ways, the sheer number of participants renders the multiple-choice question format inevitable. Two publications have summarized the relevant factors for good multiple-choice questions, demonstrated an improvement from 2006 to 2012 in *Deutsches Ärzteblatt* and showed the journal to be in a leading position among German journals offering CME [10,11]. Further, the consequences of print CME for everyday clinical practice remain unclear because no follow-up data concerning the application of gained knowledge or abilities are available. This aspect should be investigated in future studies.

Conclusions

In summary, this investigation into the largest print CME program in German-speaking countries indicates that delivery of interdisciplinary CME articles is feasible and attracts a broad range of physicians with regard to medical specialty, age, gender, and position in the health care system. Nevertheless, we have identified certain clusters of physicians for whom self-directed learning by reading CME articles may be especially attractive.

Conflicts of Interest

Christopher Baethge is employed by Deutscher Ärzteverlag, the publisher of *Deutsches Ärzteblatt*.

Multimedia Appendix 1

Example of knowledge test, *Deutsches Ärzteblatt International* 2016; 113: 616–26.

[[PDF File \(Adobe PDF File\), 34KB](#) - [jmir_v19i4e49_app1.pdf](#)]

Multimedia Appendix 2

Explanation of German health care terms in English.

[[PDF File \(Adobe PDF File\), 10KB](#) - [jmir_v19i4e49_app2.pdf](#)]

References

1. Bundesärztekammer. 2004. (Muster-)Satzungsregelung Fortbildung und Fortbildungszertifikat URL: <http://www.bundesaerztekammer.de/aerztetag/beschlussprotokolle-ab-1996/107-daet-2004/top-v-fortbildung/muster-satzungsregelung-fortbildung-und-fortbildungszertifikat/> [accessed 2017-03-13] [WebCite Cache ID 6ovdNdZN0]
2. Accreditation Council for Continuing Medical Education (ACCME) 2015 Annual Report. Chicago, IL: Accreditation Council for Continuing Medical Education; 2016. URL: http://www.accme.org/sites/default/files/630_20160719_2015_Annual_Report.pdf [accessed 2017-03-13] [WebCite Cache ID 6ovdqwkui]
3. Schäfer, CB. Charakterisierung des Teilnehmerspektrums von zertifizierter Fortbildung in Printmedien [dissertation]. Köln, Germany: Universität zu Köln; 2011.
4. Marohl R. Geschlechtsspezifische Aspekte in der Nutzung zertifizierter Fortbildungen in Printmedien [dissertation]. Köln, Germany: Universität zu Köln; 2011.
5. Baethge C, Marx C. Continuing medical education in deutsches arzteblatt - an analysis. *Dtsch Arztebl Int* 2008 Mar;105(11):204-206 [FREE Full text] [doi: [10.3238/arztebl.2008.0204](https://doi.org/10.3238/arztebl.2008.0204)] [Medline: [19629199](https://pubmed.ncbi.nlm.nih.gov/19629199/)]
6. Baethge C, Marx C, Griebenow R, Leuschner H. CME in a general medical journal: three-year evaluation of CME in *Deutsches Ärzteblatt*. *Teach Learn Med* 2010 Jul;22(3):229-232. [doi: [10.1080/10401334.2010.488486](https://doi.org/10.1080/10401334.2010.488486)] [Medline: [20563947](https://pubmed.ncbi.nlm.nih.gov/20563947/)]
7. Christ H, Baethge C, Lehmacher W, Loesche P, Griebenow R. Interdisciplinary CME: is the need evident? Results of the evaluation of CME articles in the *Journal of the German Medical Association*. *J Eur CME* 2015 Oct 21;4:1-6. [doi: [10.3402/jecme.v4.28331](https://doi.org/10.3402/jecme.v4.28331)]
8. Bundesärztekammer. Die ärztliche Versorgung in Deutschland URL: <http://www.bundesaerztekammer.de/ueber-uns/aerztstatistik/> [accessed 2017-03-13] [WebCite Cache ID 6ovowU8Au]
9. Nobbe H, Lösche P, Griebenow R. [Structure of text and multiple-choice questions in continuing medical education (CME) in two specialist journals]. *Med Klin (Munich)* 2008 Jan 15;103(1):14-19. [doi: [10.1007/s00063-008-1001-7](https://doi.org/10.1007/s00063-008-1001-7)] [Medline: [18219458](https://pubmed.ncbi.nlm.nih.gov/18219458/)]
10. Kühne-Eversmann L, Nussbaum C, Reincke M, Fischer MR. [CME activities of medical journals: quality of multiple-choice questions as evaluation tool. Using the example of the German medical journals *Deutsches Ärzteblatt*, *Deutsche Medizinische Wochenschrift*, and *Der Internist*]. *Med Klin (Munich)* 2007 Dec 15;102(12):993-1001. [doi: [10.1007/s00063-007-1123-3](https://doi.org/10.1007/s00063-007-1123-3)] [Medline: [18075719](https://pubmed.ncbi.nlm.nih.gov/18075719/)]
11. Drossard S, Kühne-Eversmann L, Bauer D, Fischer MR: Strukturqualität von MC-Fragen in der ärztlichen Fortbildung-Welchen Einfluss hat die Einführung von Qualitätsmaßnahmen? Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Graz, 26.-28. Düsseldorf: German Medical Science GMS Publishing House; 2013.

Abbreviations

CME: continuing medical education

Edited by G Eysenbach; submitted 31.05.16; peer-reviewed by J Apolinário-Hagen, C Auschra; comments to author 18.07.16; revised version received 16.11.16; accepted 13.12.16; published 03.04.17.

Please cite as:

Christ H, Franklin J, Griebenow R, Baethge C

An Analysis of 2.3 Million Participations in the Continuing Medical Education Program of a General Medical Journal: Suitability, User Characteristics, and Evaluation by Readers

J Med Internet Res 2017;19(4):e49

URL: <http://www.jmir.org/2017/4/e49/>

doi: [10.2196/jmir.6052](https://doi.org/10.2196/jmir.6052)

PMID: [28373156](https://pubmed.ncbi.nlm.nih.gov/28373156/)

©Hildegard Christ, Jeremy Franklin, Reinhard Griebenow, Christopher Baethge. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 03.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Pedestrian Inattention Blindness While Playing Pokémon Go as an Emerging Health-Risk Behavior: A Case Report

Stefania Barbieri^{1,2,3*}, MD; Gianna Vettore^{1*}, MD; Vincenzo Pietrantonio^{1*}, MD; Rossella Snenghi^{4*}, MD; Alberto Tredese^{1*}, MD; Mauro Bergamini², MD; Sara Previato², MD; Armando Stefanati², MD; Rosa Maria Gaudio³, MD; Paolo Feltracco^{1*}, MD

¹Department of Urgent and Emergency Care, University of Padova, Padova, Italy

²Preventive Medicine and Risk Assessment, University of Ferrara, Ferrara, Italy

³Forensic Medicine and Toxicology, University of Ferrara, Ferrara, Italy

⁴Department of Legal Medicine, University of Padova, Padova, Italy

*these authors contributed equally

Corresponding Author:

Stefania Barbieri, MD

Department of Urgent and Emergency Care

University of Padova

Via Giustiniani 2, Padova

Padova,

Italy

Phone: 39 0498074414

Fax: 39 0498212347

Email: stefibarbieri118@gmail.com

Abstract

Background: Cases of trauma resulting from the use of mobile phones while driving motor vehicles have become quite common in recent years. Road injuries incurred by people playing video games on mobile phones (or other media devices) while walking have also become a cause for concern. Pokémon Go has been the world's most popular game since it was launched in July 2016, with more than 15 million players trying to catch all Pokémon available in the game; however, the case detailed here is the first reported accident in the medical literature caused by a pedestrian distracted by the game while crossing a street.

Objective: We aim to provide additional information on the innovative nature of distractions that generate risks in road-users, and to explore the underreporting of pedestrian-motor vehicle collisions due to mobile device usage.

Methods: We included in this case report a 25-year-old male who suddenly crossed a road while playing Pokémon Go and was hit by a van, reporting several injuries and being assisted by the Emergency Medical Service of our hospital (Padova, Italy). The patient's history, the circumstances in which the collision happened, imaging data, and clinical course information were recorded per our hospital's privacy policy.

Results: The patient hit by the van was playing Pokémon Go on his mobile phone while crossing a street, despite red traffic lights, which he did not notice due to of the distraction induced by the game.

Conclusions: Mobile videogames that imply movement (ie, walking, running, cycling) to play are an effective way to improve physical activity practice, especially in adolescents and young adults. Nevertheless, cases like the one presented here point out that these games could pose a significant risk to users who play while walking, cycling, or driving in unsafe areas such as city streets, because players become distracted and may ignore surrounding hazards. Comprehensive, multilevel interventions are needed to reduce accidents caused by distraction, and to stress findings on the positive and negative effects of video games, which are becoming a source of public health concern. Health care providers should be aware of their chief role in these possible prevention strategies, based on their direct interactions with road incident victims.

(*J Med Internet Res* 2017;19(4):e86) doi:[10.2196/jmir.6596](https://doi.org/10.2196/jmir.6596)

KEYWORDS

Pokémon Go; videogames; pedestrians; multiple trauma; traffic accidents; motor vehicle collisions; road injuries

Introduction

Cases of trauma, and even death, resulting from the use of mobile phones while driving motor vehicles have become quite common in recent years. Road injuries incurred by people playing video games on mobile phones (or other media devices) while walking have also become a cause for concern [1-5]. Several authors have reported the combination of distracted walking or driving as a frequent cause of road traffic incidents; mobile phones are a well-known cause of distraction [3]. However, only a very small number of injured individuals confess to have been using the mobile phone while walking or driving. Despite widespread and tighter regulations covering the use of mobile phones while driving, mobile phone distraction rates remain high in young drivers, and often result in fatal incidents [3,5,6]. The use of mobile phones (and other media devices) is not regulated for pedestrians, and considering virtual reality games are particularly popular among young people, emergency departments will probably see a rise in trauma cases caused by the use of video games while walking. Nasar and Troyer showed that pedestrians, like drivers, reported injuries resulting from distraction due to the use of mobile phones or digital devices [3]. In a study promoted by the Department of Health Promotion, Bungum et al [7] described the behavior of 866 individuals: approximately 20% of walkers were distracted as they crossed streets because they were talking on a mobile phone, eating, drinking, smoking, or they were walking inside the crosswalk while a red light was flashing against them. These circumstances were critical in many of the incidents reported [7]. As reported by several authors, mobile phone-related injuries among pedestrians lead to increased morbidity due to distracted attention and unsafe behavior during their use, and users risk serious injury, falls, collisions with obstacles, or even death [8,9].

Trauma in emergency departments may present health care providers with a myriad of unforeseen clinical scenarios, and the anamnestic data entered must reflect the modern context. Pokémon Go and other virtual-reality applications for mobile phones and media devices meant to be used while walking or biking have the advantage of providing entertainment in outdoor activities for children and young people, thus favoring physical exercise [3,7,8,10]. However, the indiscriminate use of mobile phones and other virtual-reality games can also pose a health

hazard, as pedestrians, bikers, and drivers distracted by the games have a higher risk of causing motor vehicle collisions, or becoming their victims. More than 270,000 pedestrians die on the world's roads each year, accounting for 22% of all road traffic deaths (1.24 million total) [11]. In addition, in 2013 more than 150,000 pedestrians were treated in emergency departments for nonfatal crash-related injuries [12]. Some reports describe situations in which an accident occurred when car drivers or pedestrians were using hand-held mobile phones. In 2014, Basch et al [1] observed walking behaviors in New York pedestrians and found that one in four pedestrians (>3500 individuals observed) were distracted by mobile electronic devices while crossing roads during the *walk* (28.8%) and *do not walk* (26.3%) signals.

Pokémon Go has been the world's most popular game since it was launched in July 2016, with more than 15 million players trying to catch all Pokémon available in the game; however, the case reported here is the first accident described in the medical literature caused by a pedestrian distracted by the game while crossing a street. In a recent car collision in Vienna, the driver stopped in the middle of the street while playing Pokémon Go. When questioned by a police officer, the driver admitted that he and his passengers were playing the game [13]. Pokémon Go may become particularly dangerous if the player attempts to gain levels and *Gym Prestige* by completing the *Pokedex*, which involves capturing the rarest Pokémon and requires the most effort. This goal is potentially the most distracting activity in the game, although it does represent a strong motivation to increase walking speed and distance (Figure 1).

Here we briefly report details of a serious motor vehicle accident involving a young pedestrian, who was distracted while playing Pokémon Go as he was crossing a residential road with a traffic light signaling against him. The purpose of this paper is to provide additional information on the innovative nature of distractions that generate risks in road-users, and to explore the underreporting of pedestrian-motor vehicle collisions due to mobile device usage.

After obtaining appropriate patient consent, a thorough review of the incident scene, the patient's hospital records (prehospital and emergency room sources), and official police-reported incident data were examined.

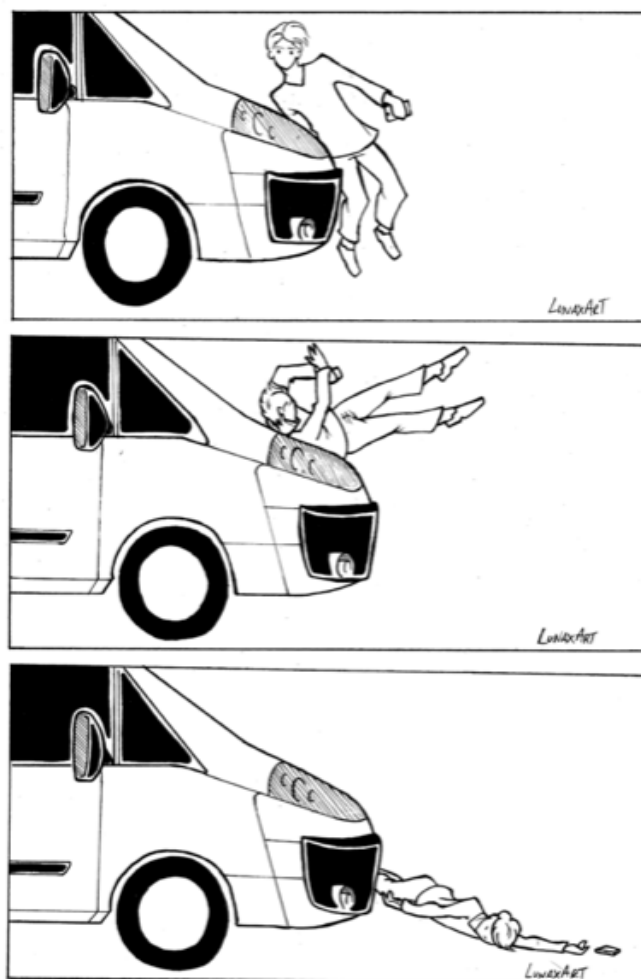
Figure 1. Pokémon Go is a virtual reality smartphone-based videogame.



Methods

This paper describes a case report a 25-year-old male who suddenly crossed a road while playing Pokémon Go and was

hit by a van (Figure 2), reporting several injuries and being assisted by the Emergency Medical Service of our hospital (Padova, Italy). The patient's history, the circumstances in which the collision happened, imaging data, and clinical course information were recorded per our hospital's privacy policy.

Figure 2. The pedestrian-vehicle collision sequence.

Results

The patient had been walking fast, on his own, on the sidewalk of a residential street when he unexpectedly, “jumped into the road at the intersection without attempting to look in any direction,” as described by the van driver; witnesses to the incident provided similar accounts (Figure 3). The patient collided with the front of a 3-ton van driven by a man who did not expect the event. According to the driver’s expectation of the young man’s movements immediately before the collision, the driver could not have foreseen his intention, as the youth’s eyes were focused on his smartphone. The speed of the van was compatible with legal restrictions and with the conditions of the street, and the driver avoided a more severe impact by slamming on his brakes.

Later analysis of the incident showed that the young man’s violation of the red light had caused the accident, that the van was travelling at a reasonable speed, and that the violence of impact was compatible with a short man-vehicle distance (high-energy trauma). According to the young man’s statement, he was distracted; he was intensely involved in the game, and neither looked at the traffic light nor heard the noise of a rapidly approaching vehicle.

According to the existing literature, pedestrian simulation models are based on flow dynamics, traffic conditions, vehicle type, road geometry, characteristics of pedestrian behavior, and police reports. Table 1 lists the details as a report covering some aspects of the context of this particular road incident, and its potential contributing factors.

Table 1. Characteristics of the collision.

Parameter	Details
Pedestrian	
	Male
	25 years old
	No alcohol or drugs
	University student
	Wore glasses
	Pedestrian travel speed (calculated 4.6 kilometers/hour)
	Pedestrian kinematics followed a typical <i>wrap</i> trajectory
	Injury type: lower and upper extremities, thorax
	Distraction due to playing Pokémon Go: no perception of crossing device
Environment	
	12:00 AM; high-volume traffic
	Residential area (speed limit 50 kilometers/hour)
	Pedestrian crosswalk
	Red traffic light for pedestrian
Vehicle	
	Van
	Number of vehicles involved in a collision=1
Driver	
	Male
	50 years old
	No alcohol or drug abuse
Dynamic	
	Vehicle speed <40 kilometers/hour
	Distance pedestrian-van: 10 meters
	Lateral collision
	Contact point pedestrian-van: (1) primary injuries, and (2) secondary injuries
	Injured body regions: lateral collision
	Pedestrian impact kinematics: the young man was thrown onto the van bonnet and projected towards the windscreen; he landed on the van roof, and then fell to the ground

The severity of the young man's injuries was related to many factors, including vehicle speed and the angle at which the impact occurred. The patient sustained blunt closed thoracic and lower limb trauma. In the prehospital setting the patient complained of intense pain in many areas (posterior head and neck, right arm, right leg, abdomen) but was hemodynamically stable and almost completely conscious. The patient was transferred directly from the scene of the accident to the hospital and immediately subjected to computerized tomography scanning, which revealed multiple high-energy traumas due to major injuries. In accordance with *Advanced Trauma Life Support* guidelines, a complete primary survey was carried out,

including diagnostic chest X-rays and pelvic X-rays. The initial *Focused Assessment with Sonography for Trauma* examination was followed by a secondary survey with a whole-body computerized tomography scan. The powerful impact of the van had caused multiple fractures to the ribs, and a left-sided pneumothorax was identified (Figure 4), together with fractures of the left femur and right tibia. No brain damage was detected. A pleural drainage system was inserted, and the lower limb fractures were stabilized. In-hospital stay was lengthy (10 days), but no major respiratory problems or infectious diseases occurred, and repair to the lower limb fractures was uneventful.

Figure 3. Circumstances and setting of the incident.

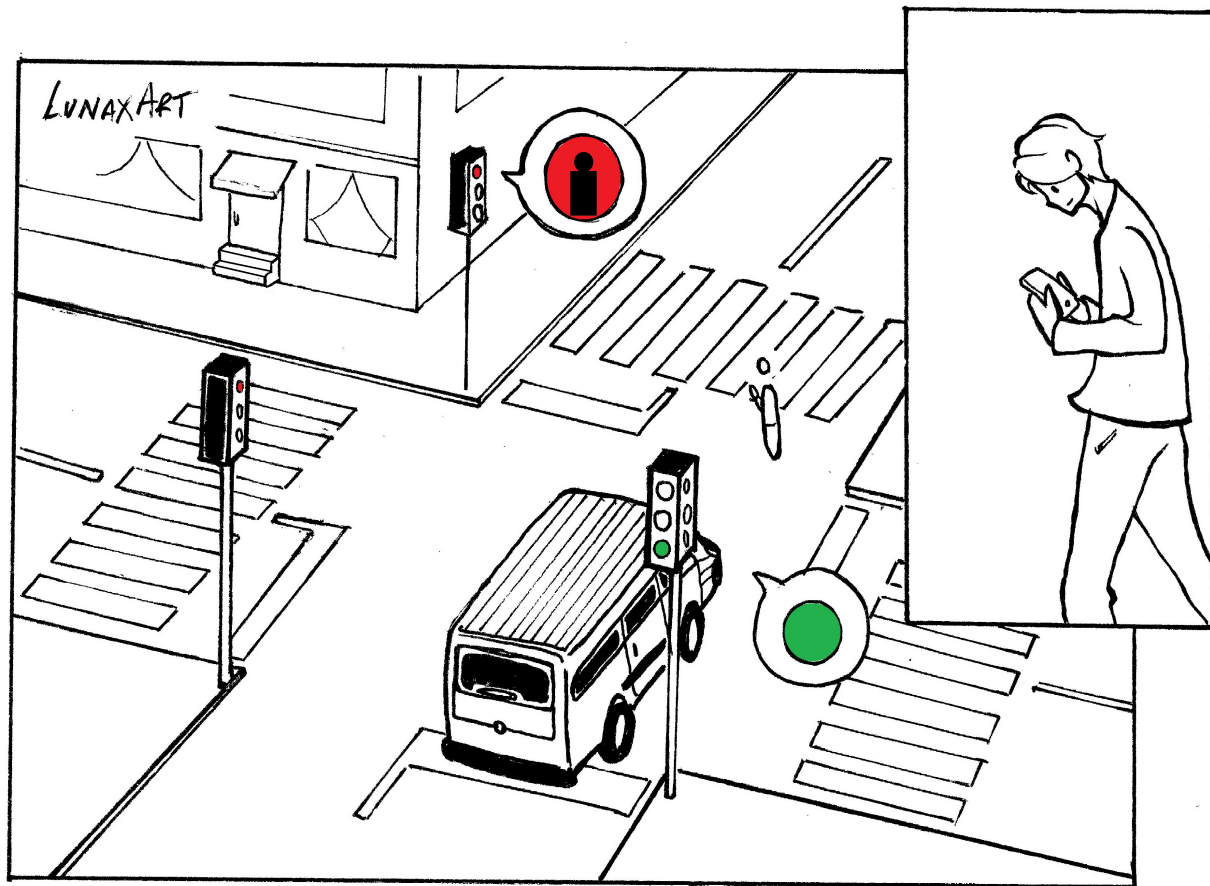
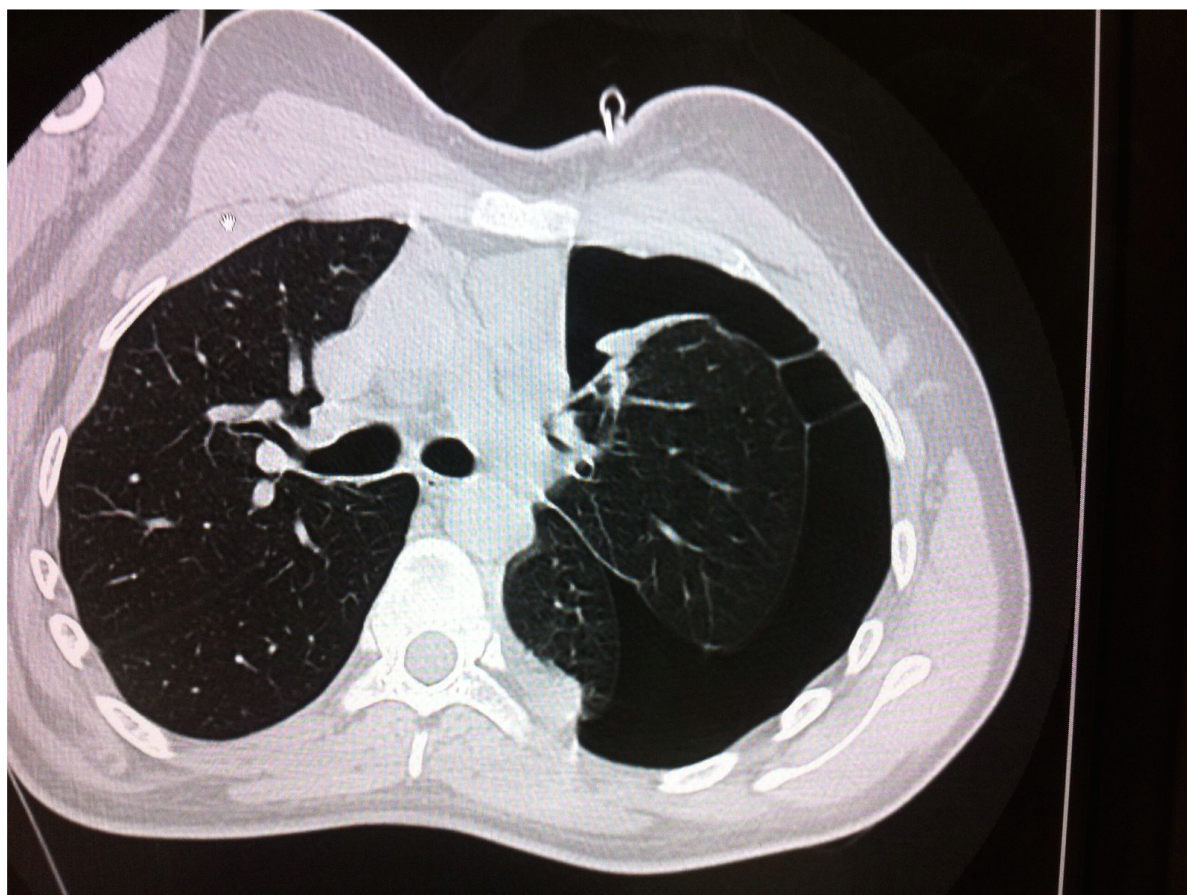


Figure 4. Left-sided pneumothorax due to the collision.



Discussion

Walking and physical activity are important factors in reducing noncommunicable diseases, including cardiovascular disease, cancers, obesity, type 2 diabetes, and impaired quality of life. Health professionals have noted the potential benefits of Pokémon Go smartphone games, which have reached millions of people and improved outdoor physical activity and social interactions. To increase the chances of catching Pokémon characters and Pokémon eggs, participants are encouraged to walk through city streets as an extra activity, viewing the world through their mobile phones and using their cameras for 2, 5, or 10 kilometers (while incubating an egg for a new character to be born) [10].

The purpose of this paper is to provide additional and specific information on the underreporting of pedestrian-motor vehicle collisions. Knowledge of the sequence of events during the impact of the distracted pedestrian and the vehicle has been crucial for proper reconstruction of the incident, along with vehicle speed at the moment of the impact, how far the person was thrown after impact, and when the driver braked. The responsibilities of emergency medicine physicians consists of: (1) preventing severe morbidity due to injury patterns manifested in pedestrian trauma; (2) describing the incident scene and reconstructing pedestrian-vehicle collisions; and (3) identifying

the mechanism(s) of injury, the initial and final positions of the pedestrian, and the sequence of events leading to trauma.

In the accident described here, the dynamic response of the pedestrian resulted in the characteristic injuries of pedestrian-road vehicle collisions in terms of severity of injury, body posture, direction of force, impact injuries and subsequent contact with the road surface, direction of acting force, and when exactly the incident happened. The interview was conducted to analyze and report data to determine all causes of the incident which, in this particular case, was due to playing a videogame in hazardous conditions. Although many studies and prevention programs have been implemented to curtail pedestrian injuries, this case is important for a better understanding of the reasons why this accident occurred, and how to plan for future preventive actions.

This case report confirms the high risk of injury to pedestrians using mobile phones, especially if playing Pokémon Go, which is a very new form of distraction. This is the first case study concerning a road incident with injuries to a pedestrian in which the essential cause was *Pokémon Go* game search. Such cases are helpful for pediatricians and physicians to recognize this phenomenon and its risks, and proper knowledge is important for the prevention and early detection of adolescents' at-risk behaviors. The risks associated with distracted walking, biking, or driving can be analyzed from several viewpoints when

considering the relationship between distraction and inattention as a prevalent causative factor of road incidents.

Conclusions

Mobile apps that integrate game play with physical exercise, such as Pokémon Go, lead to substantial physical activity increases and reach sedentary populations [10]. In the case presented here, the need to capture virtual Pokémon creatures in real locations represents a strong motivation to increase one's walking distance and speed, but safety while walking must still be considered. The positive and negative consequences of gaming indicate that, although facilitating improved daily physical activity, gaming may result in low vigilance and secondary task distraction (a well-known phenomenon that contributes to road accidents), especially in young people.

Available data clearly indicate that distraction while moving around city streets could potentially represent a serious public health problem. Interventions should be implemented following research-based recommendations, and improved knowledge of

the context is an important consideration (in this case, behaviors concerning road traffic safety). The case presented here emphasizes the difficulties involved in identifying the circumstances of a pedestrian trauma and shows the importance of implementing common strategies to reduce road users' risks. Comprehensive, multilevel interventions and further research projects are needed to reduce incidents caused by distraction, and to more thoroughly study the positive and negative effects of video games.

Health care providers should be aware of their chief role in these possible prevention strategies, based on their direct interactions with road incidents victims. Pokémon Go was released very recently, and no follow-up studies have examined the long-term effectiveness of encouraging players to increase physical activity levels while following Pokémon creatures. Therefore, we still do not know whether the benefits of playing the game outweigh the risks, and further research over a longer observation period is needed.

Acknowledgments

The authors gratefully acknowledge the support of Milanopulo G. and Sguotti G. for their contributions to the manuscript and technical assistance throughout the process.

Conflicts of Interest

None declared.

References

1. Basch CH, Ethan D, Rajan S, Basch CE. Technology-related distracted walking behaviours in Manhattan's most dangerous intersections. *Inj Prev* 2014 Oct;20(5):343-346. [doi: [10.1136/injuryprev-2013-041063](https://doi.org/10.1136/injuryprev-2013-041063)] [Medline: [24667654](https://pubmed.ncbi.nlm.nih.gov/24667654/)]
2. Richtel M. The New York Times. 2010 Jan 10. Forget gum. Walking and using phone is risky URL: <http://www.nytimes.com/2010/01/17/technology/17distracted.html> [accessed 2017-03-15] [WebCite Cache ID 6oz1C6kAv]
3. Nasar J, Troyer D. Pedestrian injuries due to mobile phone use in public places. *Accid Anal Prev* 2013 Aug;57:91-95. [doi: [10.1016/j.aap.2013.03.021](https://doi.org/10.1016/j.aap.2013.03.021)] [Medline: [23644536](https://pubmed.ncbi.nlm.nih.gov/23644536/)]
4. Neider MB, McCarley JS, Crowell JA, Kaczmarek H, Kramer AF. Pedestrians, vehicles, and cell phones. *Accid Anal Prev* 2010 Mar;42(2):589-594. [doi: [10.1016/j.aap.2009.10.004](https://doi.org/10.1016/j.aap.2009.10.004)] [Medline: [20159083](https://pubmed.ncbi.nlm.nih.gov/20159083/)]
5. Stavrinou D, Byington KW, Schwebel DC. Effect of cell phone distraction on pediatric pedestrian injury risk. *Pediatrics* 2009 Feb;123(2):e179-e185. [doi: [10.1542/peds.2008-1382](https://doi.org/10.1542/peds.2008-1382)] [Medline: [19171568](https://pubmed.ncbi.nlm.nih.gov/19171568/)]
6. Hatfield J, Murphy S. The effects of mobile phone use on pedestrian crossing behaviour at signalized and unsignalized intersections. *Accid Anal Prev* 2007 Jan;39(1):197-205. [doi: [10.1016/j.aap.2006.07.001](https://doi.org/10.1016/j.aap.2006.07.001)] [Medline: [16919588](https://pubmed.ncbi.nlm.nih.gov/16919588/)]
7. Bungum T, Day C, Henry L. The association of distraction and caution displayed by pedestrians at a lighted crosswalk. *J Community Health* 2005 Aug;30(4):269-279. [Medline: [15989209](https://pubmed.ncbi.nlm.nih.gov/15989209/)]
8. Edell A, Jung J, Solomon J, Palu RN. Distracted pedestrian sustains orbital fracture while on cell phone. *Clin Ophthalmol* 2013;7:671-673 [FREE Full text] [doi: [10.2147/OPHTH.S41257](https://doi.org/10.2147/OPHTH.S41257)] [Medline: [23579597](https://pubmed.ncbi.nlm.nih.gov/23579597/)]
9. Violano P, Roney L, Bechtel K. The incidence of pedestrian distraction at urban intersections after implementation of a Streets Smarts campaign. *Inj Epidemiol* 2015 Dec;2(1):18 [FREE Full text] [doi: [10.1186/s40621-015-0050-7](https://doi.org/10.1186/s40621-015-0050-7)] [Medline: [27747750](https://pubmed.ncbi.nlm.nih.gov/27747750/)]
10. Althoff T, White R, Horvitz E. Influence of Pokémon Go on physical activity: study and implications. *J Med Internet Res* 2016 Dec 06;18(12):e315 [FREE Full text] [doi: [10.2196/jmir.6759](https://doi.org/10.2196/jmir.6759)] [Medline: [27923778](https://pubmed.ncbi.nlm.nih.gov/27923778/)]
11. Sminkey L. World Health Organization. Geneva: WHO Media Centre; 2013 May 02. More than 270,000 pedestrians killed on roads each year URL: http://www.who.int/mediacentre/news/notes/2013/make_walking_safe_20130502/en/ [accessed 2017-03-15] [WebCite Cache ID 6oz22zHrG]
12. Centers for Disease Control and Prevention. WISQARS (Web-based Injury Statistics Query and Reporting System). Atlanta, GA: US: Department of Health and Human Services, CDC; 2015; 2015. Pedestrian safety URL: https://www.cdc.gov/motorvehiclesafety/pedestrian_safety/ [accessed 2017-03-15] [WebCite Cache ID 6oz2FrKQT]

13. Barton MA. Vienna Patch, 2016. Vienna: Vienna Patch; 2016 Jul 15. Pokemon Go in Vienna: driver stopped in the middle of the road playing popular game URL: <http://patch.com/virginia/vienna/pokemon-go-vienna-driver-stopped-middle-road-playing-popular-game> [accessed 2017-03-15] [[WebCite Cache ID 6oz2OJKIN](#)]

Edited by G Eysenbach; submitted 25.09.16; peer-reviewed by T Baranowski, A Quistberg; comments to author 28.01.17; revised version received 03.03.17; accepted 04.03.17; published 01.04.17.

Please cite as:

Barbieri S, Vettore G, Pietrantonio V, Snenghi R, Tredese A, Bergamini M, Previato S, Stefanati A, Gaudio RM, Feltracco P. Pedestrian Inattention Blindness While Playing Pokémon Go as an Emerging Health-Risk Behavior: A Case Report. *J Med Internet Res* 2017;19(4):e86

URL: <http://www.jmir.org/2017/4/e86/>

doi: [10.2196/jmir.6596](https://doi.org/10.2196/jmir.6596)

PMID: [28365563](https://pubmed.ncbi.nlm.nih.gov/28365563/)

©Stefania Barbieri, Gianna Vettore, Vincenzo Pietrantonio, Rossella Snenghi, Alberto Tredese, Mauro Bergamini, Sara Previato, Armando Stefanati, Rosa Maria Gaudio, Paolo Feltracco. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Toward Game-Based Digital Mental Health Interventions: Player Habits and Preferences

Regan Lee Mandryk^{1*}, MSc, PhD; Max Valentin Birk^{1*}, Dipl-Psych

Interaction Lab, Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada

* all authors contributed equally

Corresponding Author:

Regan Lee Mandryk, MSc, PhD

Interaction Lab

Department of Computer Science

University of Saskatchewan

176 Thorvaldson Bldg

110 Science Place

Saskatoon, SK, S7N5C9

Canada

Phone: 1 3069664888

Fax: 1 3069664884

Email: regan@cs.usask.ca

Abstract

Background: Designers of digital interventions for mental health often leverage interactions from games because the intrinsic motivation that results from game-based interventions may increase participation and translate into improved treatment efficacy. However, there are outstanding questions about the suitability (eg, are desktop or mobile interventions more appropriate?) and intervention potential (eg, do people with depression activate enough to play?) of games for mental health.

Objective: In this paper, we aimed to describe the presently unknown relationship between gaming activity and indicators of well-being so that designers make informed choices when designing game-based interventions for mental health.

Methods: We gathered validated scales of well-being (Beck's Depression Inventory [BDI-II], Patient Health Questionnaire [PHQ-9], trait anxiety [TA], and basic psychological needs satisfaction [BPNS]), play importance (control over game behavior: control; gamer identity: identity), and play behavior (play frequency, platform preferences, and genre preferences) in a Web-based survey (N=491).

Results: The majority of our participants played games a few times a week (45.3%, 222/490) or daily (34.3%, 168/490). In terms of depression, play frequency was associated with PHQ-9 ($P=.003$); PHQ-9 scores were higher for those who played daily than for those who played a few times a week or less. Similarly, for BDI-II ($P=.01$), scores were higher for those who played daily than for those who played once a week or less. Genre preferences were not associated with PHQ-9 ($P=.32$) or BDI-II ($P=.68$); however, platform preference (ie, mobile, desktop, or console) was associated with PHQ-9 ($P=.04$); desktop-only players had higher PHQ-9 scores than those who used all platforms. Platform preference was not associated with BDI-II ($P=.18$). In terms of anxiety, TA was not associated with frequency ($P=.23$), platform preference ($P=.07$), or genre preference ($P=.99$). In terms of needs satisfaction, BPNS was not associated with frequency ($P=.25$) or genre preference ($P=.53$), but it was associated with platform preference ($P=.01$); desktop-only players had lower needs satisfaction than those who used all platforms. As expected, play frequency was associated with identity ($P<.001$) and control ($P<.001$); those who played more had identified more as a gamer and had less control over their gameplay. Genre preference was associated with identity ($P<.001$) and control ($P<.001$); those who played most common genres had higher control over their play and identified most as gamers. Platform preference was not associated with control ($P=.80$), but was with identity ($P=.001$); those who played on all devices identified more as a gamer than those who played on mobiles or consoles only.

Conclusions: Our results suggest that games are a suitable approach for mental health interventions as they are played broadly by people across a range of indicators of mental health. We further unpack the platform preferences and genre preferences of players with varying levels of well-being.

(*J Med Internet Res* 2017;19(4):e128) doi:[10.2196/jmir.6906](https://doi.org/10.2196/jmir.6906)

KEYWORDS

computer games; mental health; depression; anxiety

Introduction**Motivating Digital Interventions for Mental Health**

The prevalence of mental illness is on the rise [1]—the National Institute of Mental Health reports that 18% of adults in the United States had a diagnosed mental illness in 2014 [2]; however, 57% of adults with mental illness in the United States do not receive treatment [3]. Untreated mental illness has serious consequences [4]. The cost of depression and anxiety alone is estimated at US \$1 trillion per year [5]. In addition to these financial costs, people experience costs to their well-being that range from a lower quality of life [6] to a loss of life [7]—mental illness is the most important risk factor for suicide as more than 90% of people who commit suicide have a mental or addictive disorder [8,9].

There are several reasons why treatment of mental illness has not evolved to meet the growing demand. First, health care systems cannot handle the burden—there are not enough trained professionals to provide treatment to those in need [10]. Second, access to treatment is lower for people who live in underpopulated areas. For example, in Nunavut (a territory in northern Canada), the suicide rate is 10 times higher than in the rest of Canada [11,12], prompting Nunavut's chief coroner to suggest that a state of emergency regarding mental illness in the north should be declared [12]. Third, access to treatment is lower for people who live in low-income areas [10]; one in five American adults with mental illness report that the cost of treatment is a barrier for them [3]. And fourth, because mental illness ranges from mild to disabling impairment [1], there are subclinical populations with mild impairment who would benefit from treatment but lack access to a system that is already heavily burdened from treating those with more severe problems.

These limitations in access to treatment cannot be addressed solely by growing the existing mental health system [13]; rather, the solution requires a variety of approaches, including a fundamental shift in delivery mechanisms that will necessarily involve digital solutions [14-16]. There are already several examples of digital systems for interventions in the domain of mental health. For example, previous research has shown that mobile phone apps can be used to treat a variety of mental health disorders such as depression or anxiety [17-19]. Self-help resources such as mobile mental health apps are available for individuals experiencing mild to moderate anxiety and depression [20]. The US Department of Veteran Affairs recommends the mobile app “PTSD coach” (US Department of Veterans Affairs, 2011) that allows people with post-traumatic stress disorder (PTSD) to get access to information about PTSD, track symptoms, and includes small tasks to help people handle their symptoms of stress [21]. The self-help app “Koko” (Koko Inc, 2015) takes a different approach by connecting people in need of emotional support with people who are willing to respond to such a demand. Koko was evaluated in a randomized control trial (RCT) and was shown to be effective [22]. Computerized cognitive behavioral therapy (CBT) has also

shown promising results in combating anxiety [23-25], PTSD [21,26,27], and depression [28,29].

Results such as these demonstrate that technology-based mental health interventions offer promise for use in self-help or as an adjunct to clinical treatment [30]; however, there are also several problems that arise when moving digital mental health interventions out of the clinic and into the world. For example, digital solutions can involve interaction with a therapist or formal setting [27,31]; however, many digital solutions are intended to be used without aid from a therapist, which can result in a loss of accountability for a participant [32]. Moving out of the clinic also necessarily results in a loss of external regulation [33] for patients, as there is no longer a psychotherapist but instead a digital app guiding a participant through an exercise. Designers of digital self-help mental health interventions must compensate for these types of problems, as the efficacy of treatment cannot be sacrificed for the wider reach that digital solutions provide.

One solution is to introduce external regulation into digital interventions. For example, we could design systems that require check-ins or provide tangible benefits for daily participation [34,35]. Another solution is to highlight and reinforce the positive benefits to mental health that will eventually occur through sustained participation in the intervention [35,36]. The form of motivation used by these examples, that is, engaging in an activity (training) because it leads to a desirable but separable outcome (benefits to mental health, tangible rewards), is called *extrinsic motivation* [33,37]. Rather than including methods of external regulation or increasing extrinsic motivation, we could alternatively make the treatment activity engaging enough so that people participate because they actually enjoy the treatment itself, and not just because they await the beneficial outcome to their mental health that will result from sustained participation. This form of motivation, that is, engaging in an activity because it is inherently interesting, is referred to as *intrinsic motivation* [33]. Intrinsically motivated people have been shown to be more willing to invest effort into a task and also derive more enjoyment from it [38]. However, the question remains on how treatment activities for improved mental health can be made inherently enjoyable, rather than being undertaken because participants desire the benefits to mental health that sustained participation will eventually provide.

Game-Based Digital Interventions for Mental Health

One common approach to increasing intrinsic motivation with a digital system is to increase the inherent enjoyment of the activity itself by leveraging the motivational pull of digital games [39]. People play a lot of games; recent estimates suggest that more money is spent purchasing games (US \$92 billion) than music (US \$18 billion) and movies (US \$62 billion) combined [40]. It was found that 4 out of 5 American households own a device that is used to play video games and 115 million Americans play games [41]. Internationally, the global game market is expected to exceed US \$102 billion by

2017 [42]. Although people sometimes assume that it is highly immersive console and desktop games that drive the game industry, 35% of those same revenues are expected to be generated through mobile phones and tablets, which people use to play games that are more casual in nature [42]. With so much time and money being spent on digital game play, researchers have started to question what it is about games that make them so compelling to play. One leading explanation for game enjoyment arises out of self-determination theory (SDT) [38] and explains that playing digital games satisfies our psychological needs of competence, autonomy, and relatedness during play itself [39], leading to enjoyment of the experience. Digital game researchers have started to deconstruct how needs satisfaction during play leads to game enjoyment for different players [39,43,44] and in varying gameplay scenarios [45-47]. Others have explored how to leverage these game design elements to motivate people in nongame contexts—a process called gamification [48,49]. Gamified systems have been evaluated in contexts such as learning [47] and behavior change [50,51], and there are also already several examples of gamified interventions in the context of health [52-58]. One of the most successful examples of gamified health interventions is Re-Mission [54], a game designed for children with cancer in which they could shoot cancer cells, manage nausea and constipation, and overcome bacterial infections. A large international study demonstrated that children who played Re-Mission exhibited better adherence to a treatment protocol, improved self-efficacy, and greater cancer-related knowledge than a group who played another computer game [59]. More recently, the application of gamification to treatment has extended beyond addressing physical health into the domain of improving mental health.

For example, a common technique used in CBT is attentional retraining of a participant's attention away from maladaptive cognitive processes [60]; attention-bias modification training (ABMT) has been shown to be an effective technique to shift a participant's attention away from negative stimuli [61]. Dennis and O'Toole [52] showed the efficacy of a mobile phone ABMT game in reducing the player's response to threatening stimuli. In another example, SPARX is a fantasy role-playing game (RPG) for managing depression that is also based on CBT and was shown to be as effective as therapist-administered CBT for treating adolescents with depression [55]. Similarly, CBT principles were translated into biofeedback-supported mini-games in the serious game, Dojo. In a RCT conducted through the school system, adolescents with anxiety who played Dojo showed decreases in anxiety-related symptoms [57], although the study's commercial control game showed similar effects. Rather than replacing CBT, some game-based approaches instead augment traditional CBT, for example, adolescents with anger-management problems were better able to improve their trait anger when the biofeedback game RAGE-Control was added to a traditional CBT protocol [53]. Finally, the biofeedback game, MindLight, was tested with children with anxiety problems and showed decreases in symptoms over a 3-week trial, although the control game used in the study showed similar effects [58]. A recent meta-analysis of game-based interventions for depression suggests a moderate effect size of game interventions at posttreatment [62].

Furthermore, the study also revealed that games based on entertainment, virtual reality exposure therapy, and exercise showed stronger results than those based on psychoeducation and training.

In an even broader context, games offer an opportunity for improving mental health at a large-scale as it is known that people use games to recover from noxious moods—a process known as mood management. Mood repair through escapism is included in validated scales on both motivations for play [63] and game addiction [64]. In addition, a recent study using gamer forums supports the idea that playing games helps players regulate their mood [65], and children have also noted that a major motivation for play is because games help them relax and forget their problems [66]. On the basis of this idea that games help promote mood equilibrium, several researchers have started to experimentally examine how games help people recover from the stress of life. Using mood management theory, Reinecke [67] used survey data to demonstrate that people use games for recovery from stressful and boring life events. Furthermore, an experiment confirmed that a game that requires interaction facilitates recovery [68]. Similarly, Bowman and Tamborini examined task demand in video games for recovery and showed that greater task demand in a game is better for mood repair, so long as the demand is not too high [69], and that participants choose games with the appropriate level of demand for recovery [70]. Finally, several researchers have suggested that digital games provide an opportunity for improved emotion regulation [71], for example, by rewarding players who successfully downregulate negative affect [72].

Contribution of This Paper

It is clear that there are potential benefits of motivating participation in digital interventions for mental health by introducing game-based elements; however, there are several outstanding questions about the whether or not games are really appropriate in this context and how great their intervention potential truly is in this domain. For example, we may argue that games have a high intervention potential because people who have low need satisfaction experience greater compulsion to play games [73,74] and are also more likely to experience less control over their play behavior [64]; and because low need satisfaction is related to depression, we may assume that people with depression pass significant time playing games. However, people who have moderate to severe depression find it difficult to activate behaviors [75,76]; simply getting out of bed can be a major challenge. Thus, it is unclear whether people with depression are in fact activating enough to play games. Furthermore, games are argued to have high intervention potential because they are regularly played on mobile phones or other mobile devices, which are platforms that have been targeted for digital interventions in mental health [52,77]. However, are gamers with mental health issues part of the growing group who play games on a mobile platform [41,42] or are they more likely to play in the comfort of their homes on a desktop or console platform? Finally, most game-based interventions use simple game mechanics that are found in casual games, that is, “games with a low barrier to entry that can be enjoyed in short increments” (p.9; [78]), rather than those found in more complex or immersive games, such as RPGs (eg,

Mass Effect) or first-person shooter (FPS) games (eg, Call of Duty). If players with depression avoid casual games instead preferring the types of complex and immersive games that are generally played on dedicated gaming consoles, they may eschew playing a game with simple mechanics on a mobile phone and thus may not be motivated to play a game-based mental health intervention delivered on that platform.

Taken together, these questions highlight a problem in our understanding of how games can be used to create digital interventions in mental health: *as researchers, we don't know how to design, personalize, or market games for mental health interventions, because we don't know the gaming habits, platform preferences, or genre preferences of people who have poor mental health, low life-satisfaction, or who suffer from depression or anxiety.*

To investigate the intervention potential of digital games, we gathered survey data from 491 online participants who completed validated scales related to their well-being, including on their levels of depression (Beck's Depression Inventory [BDI-II], Patient Health Questionnaire [PHQ-9], their trait anxiety [TA]), and their satisfaction of their basic psychological needs satisfaction (BPNS). We also asked questions about the importance of play in their lives, including about the control they felt over their gaming behavior and how much they identified as a gamer. Finally, we asked questions about play behavior, including their frequency of play, platform preferences, and genre preferences. Our results suggest that games are a suitable approach for mental health interventions as they are played broadly by people across all indicators of mental health. Throughout the remainder of this paper, we further unpack the play frequency, platform preferences, and genre preferences of players with varying levels of well-being.

Methods

Research Questions

Our study was designed to answer several research questions about the intervention potential and suitability of games for mental health interventions. Specifically:

RQ1: Are people with depression or anxiety activating enough to play games? How frequently do they play?

RQ2: Are gamers with depression or anxiety part of the growing group who play games on a mobile platform or are they more likely to play in the comfort of their homes on a desktop or console platform?

RQ3: Do players with depression or anxiety play casual games, or do they prefer more complex and immersive genres of games?

Participants, Deployment Platform, and Procedure

We recruited 491 participants through Amazon Mechanical Turk (AMT), which acts as a broker between requesters who offer Human Intelligence Tasks (HITs) and paid workers who complete them. Participants received compensation of US \$2.5 for their participation, which took approximately 15 min. Ethical approval was obtained from the University of Saskatchewan Behavioral Research Ethics Board, and participants were asked to provide informed consent at the beginning of the task. To

comply with ethical guidelines, the task was only available to workers from the United States who were older than 18 years. Additionally, only workers with an approval rate above 90% were offered the task as a means of quality control.

Measures

We gathered indicators of well-being, indicators of the importance of play, and information regarding gaming behaviors.

Indicators of Well-Being

Degree of depression was measured using the BDI-II [79] and the PHQ-9 [80]. Both are well-established measures that are commonly used as screening instruments in a clinical context.

The BDI-II is a psychometric test that measures depression using 21 questions that address categories of behavior that are associated with depression, including pessimism, past failure, self-dislike, suicidal thoughts, crying, agitation, and loss of interest in sex. For each question, participants were asked how they have been feeling over the last two weeks, including today. Responses range from not experiencing a feeling or displaying a behavior, for example, "I do not feel sad," or to an extreme expression of the same feeling or behavior, for example, "I am so sad or unhappy that I can't stand it." Responses are then mapped to a severity rating between 0 and 3. Responses are added into a total score (Cronbach alpha=.947); the BDI-II also provides threshold values that indicate categories of depression.

The PHQ-9 asks participants to rate the frequency of negative experiences, over the last two weeks (eg, "Little interest or pleasure in doing things," "Feeling bad about yourself or that you're a failure or have let yourself or your family down") on a scale from "Not at all" to "Nearly every day." Each rating was assigned a value from 0 to 3. All items were added to create a single score reflecting an overall level of depression (Cronbach alpha=.903).

TA was measured using the 20-item trait scale from the State-Trait Anxiety Scale (STAI; [81]). Participants rated their general distress using statements that people have used before to describe themselves, for example, "I'm calm," using a 4-point Likert scale from "Not at all" to "Very much" (Cronbach alpha=.951).

Need satisfaction was measured using the BPNS scale [82]. BPNS includes subscales for the basic satisfaction of competence (Cronbach alpha=.698), autonomy (Cronbach alpha=.858), and relatedness (Cronbach alpha=.865), as three ongoing needs that people need satisfied to optimally develop and function. Participants rated their agreement to statements, "People I know tell me I am good at what I do" on a 7-point Likert scale from "Strongly Disagree" to "Strongly Agree."

Indicators of Importance of Play

To measure how much importance gaming has in a person's life, we measured *control over play behavior* (CPB) using the game addiction scale, which assesses excessive, compulsive, and generally problematic use of videogames, and includes subscales for salience (Cronbach alpha=.775), tolerance (Cronbach alpha=.750), mood modification (Cronbach

alpha=.840), relapse (Cronbach alpha=.799), withdrawal (Cronbach alpha=.869), conflict (Cronbach alpha=.827), and problems (Cronbach alpha=.770). Participants rate the frequency for which they show a behavior, for example, “Did you play longer than intended,” on a 5-point scale from “never” to “very often.” A higher score indicates greater loss of CPB. Although we were not interested in assessing addictive videogame playing on a clinical level, the scale assesses the degree of uncontrollable excessive and compulsive use of digital games that results in social or emotional problems. Only at the extreme end of the scale we would consider the degree of playing as pathological. Our goal was to use the full range of the scale to assess the degree of control over obsessive or harmful game play.

Additionally, we measured *gamer identity*, asking participants to rate how much they self-identify as a gamer on a scale from “Not at all” to “Gamer,” on a single-item scale. Correlations with time played ($\rho=.425$, using Spearman rank correlation), and CPB ($r=.417$), suggest that the item measures the intended construct. We have previously used this single item alongside a 6-section (60 questions) scale based on self-attributes [83] and found that the 1-item measure correlated at $r=.735$ with the longer version (see [Multimedia Appendix 1](#)).

Measures of Play Behavior

We gathered subjective measures of play habits and preferences.

Frequency of play was measured by asking participants how often on average they play games. Responses were restricted to the following categories: “Every day,” “A few times per week,” “Once per week,” “A few times per month,” “Once a month,” “A few times per year,” “Once per year,” “Not at all.” We grouped players into 3 groups according to play frequency:

daily players ($n=168$), a few times per week ($n=222$), and once a week or less ($n=100$).

Participants indicated their *genre preferences* using a check-all-that-apply question. We asked participants whether or not they enjoy playing the following genres: “action games,” “platform games,” “FPS,” “Beat ‘em up,” “adventure games,” “RPG,” “massively multiplayer online role-playing games (MMORPG),” “simulations,” “vehicle simulations,” “strategy games,” “music games,” “sport games,” “multiplayer online battle arena (MOBA),” and “casual games.” Our data shows that some genre categories, for example, music games, had very few true answers, whereas other groups, for example, FPS, had many members. To cluster participants into groups that represent genre preferences, we performed K-means clustering [84] on the 15 genre variables that contained true or false responses. We found the best differentiation at 4 clusters. One cluster included players who enjoy many of the most common genres ($n=102$); one was for players who mainly enjoy single-player games ($n=117$); one cluster represented players who enjoyed only FPS and action games ($n=150$); and one cluster was for players who enjoyed only puzzle and casual games ($n=121$); see [Figure 1](#).

Participants indicated their *platform preferences* using a check-all-that-apply question. We asked whether participants enjoy playing on “mobile,” “desktop,” or “console” systems. Using K-means clustering on the three Boolean variables, players were separated into 4 categories of platform preference; there was one category for each individual platform (mobile only: $n=35$; console only: $n=38$; desktop only: $n=214$) and a combined category for players who use all gaming platforms ($n=203$).

Figure 1. Clusters of genre preferences from the K-means clustering.

	Cluster 1 <i>most genres</i>	Cluster 2 <i>single-player</i>	Cluster 3 <i>FPS or action</i>	Cluster 4 <i>casual or puzzle</i>
Action	X	X	X	
Adventure	X	X		
Beat ‘em Up		X		
Casual	X	X		X
FPS		X	X	
MMORPG		X		
MOBA		X		
Music				
Platform		X		
Puzzle	X			X
RPG	X	X		
Simulation	X			
Sport				
Strategy	X	X		
Vehicle Simulation				

Statistical Analysis

Data Exclusion

Although AMT has been shown to be reliable as a recruitment tool for research [85] in the domain of human-computer interaction [86] and mental health [15,87], we followed recommendations for excluding participants from the analysis if they showed indications of having not completed the questionnaires with care and attention. We calculated variance for each participant within each survey subscale and removed participants ($n=1$) from subsequent analyses who demonstrated response variance greater than three standard deviations above mean variance (over the sample) on three or more questionnaire subscales. Having high variance within a subscale is indicative of not paying attention to the survey questions and the reverse-coded items. After the outlier participant was excluded, 490 participants remained (58.2% female (285/490), mean age 34.4 years, SD 11.3) in all of our analyses.

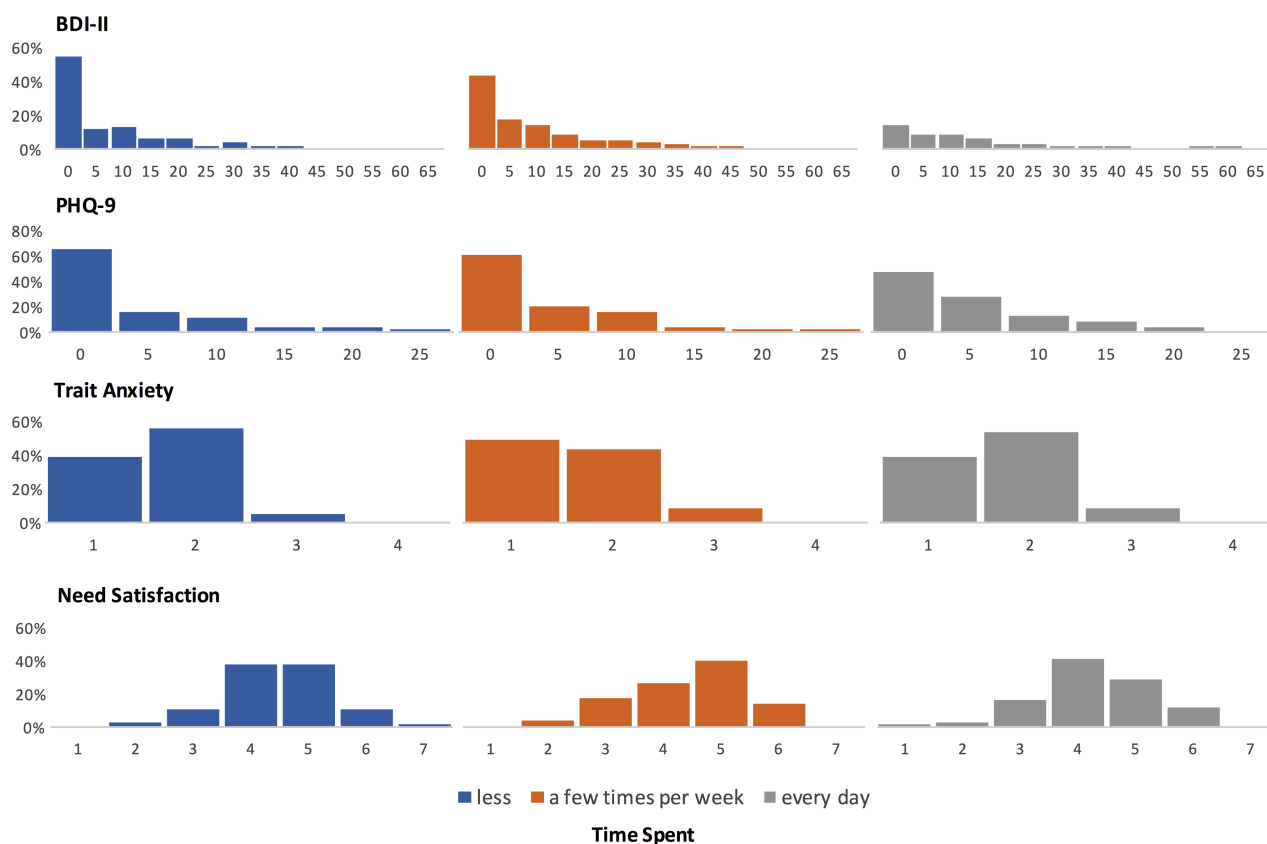
Data Analyses

BDI-II, PHQ-9, and TA data were summed into a single score reflecting level of depression, and TA respectively. As shown in Figure 2, data for BDI-II and PHQ-9 followed a Pareto distribution, thus we log-transformed data for both scales to increase robustness of the subsequent analyses [84]. For BPNS,

we calculated the mean across the three subscales of satisfaction of competence, autonomy, and relatedness, into a single score representing need satisfaction (Cronbach $\alpha=.925$). For CPB, we calculated the average across all subscales; however, we removed the mood modification subscale, because it showed low factor loadings in the original construction of the scale [64] and because playing games as a method of dissipating noxious moods is a common behavior that allows people to recover from stress or boredom [67,88] and represents a beneficial coping strategy rather than a pathological behavior [67] (Cronbach $\alpha=.947$).

To investigate the relationship between mental health and play behavior, we performed multivariate analysis of variance (MANOVA), with play behavior variables (play frequency, platform preference cluster, genre preference cluster) as factors, and indicators of well-being (BDI-II, PHQ-9, STAI trait, BPNS) and importance of play (control, identity) as dependent measures, controlling for sex and age. It is important to note that the predictor variables are not randomly assigned, but reflect the choices of participants. Therefore, our results cannot be interpreted as a causal relationship between the factor and the measure (as is common with ANOVA interpretation), but rather as an indicator of a statistically significant relationship between a categorical variable and a continuous variable. All analyses were performed using SPSS 24 (IBM Corp).

Figure 2. Bar charts of the constructs of well-being (Beck's Depression Inventory [BDI-II], Patient Health Questionnaire [PHQ-9], Trait Anxiety, and Need Satisfaction) displayed for each level of time spent (less, a few times per week, every day). For each construct, the chart shows the count per bin on the x-axis, represented as a percentage of participants in that level of time spent.



Results

Sample Distribution

Figure 3 shows the distribution of responses to the question on the frequency of play. As shown, the vast majority of respondents played games at least a few times per week or every day. Furthermore, as shown in Figure 3, there was no systematic relationship between gender and game play frequency (a Kruskal-Wallis test of game play frequency on gender was nonsignificant: $\chi^2_2=3.8$, $P=.15$).

The age of our sample ranged between 18 and 73 years (mean 34.1, SD 9.8); furthermore, as Figure 4 shows, there was no systematic relationship between age and game play frequency (a Kruskal-Wallis test of game play frequency on age was nonsignificant: $\chi^2_2=1.2$, $P=.55$).

Figure 2 shows the distributions of responses for the indicators of well-being for the three frequencies of play groupings. Figure 5 shows the means and standard deviations for the indicators of well-being and importance of play separately for the frequency of play groups, platform clusters, and genre clusters.

Figure 3. Stacked bar chart of frequency of play and gender.

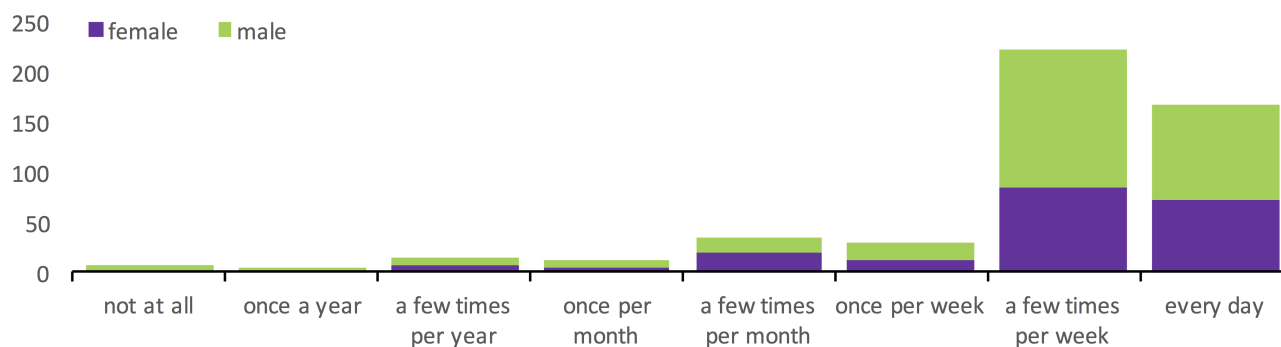


Figure 4. Bar charts of the constructs of age, loss of control, and identity displayed for each level of time spent (less, a few times per week, every day). For each construct, the chart shows the count per bin on the x-axis, represented as a percentage of participants in that level of time spent.

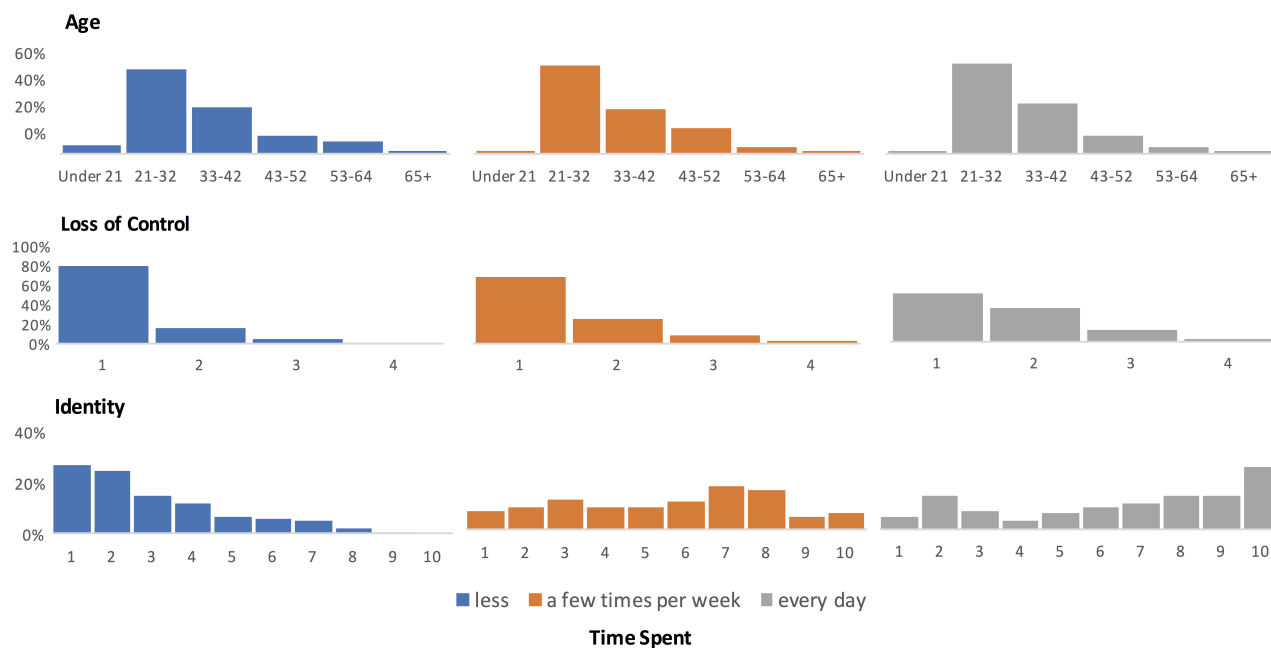


Figure 5. Means and standard deviations (SD) for the indicators of well-being and importance of play for: (Top) frequency of play groups, (Middle) platform clusters, and (Bottom) genre clusters.

	Time Spent					
	less		a few times per week		every day	
	mean	SD	mean	SD	mean	SD
BDI-II	1.58	1.03	1.78	0.99	2.07	0.92
PHQ-9	1.23	1.22	1.35	1.24	1.63	1.12
Trait Anxiety	2.07	0.57	2.03	0.61	2.14	0.57
Need Satisfaction	4.97	0.91	4.95	1.02	4.79	0.99
Control over Play Behavior	1.49	0.58	1.79	0.67	2.13	0.79
Gamer Identity	3.01	1.99	5.58	2.61	6.67	3.00

	Game Platform							
	Cluster 1 desktop		Cluster 2 all		Cluster 3 console		Cluster 4 mobile	
	mean	SD	mean	SD	mean	SD	mean	SD
BDI-II	1.96	1.19	1.72	1.19	1.95	1.13	1.67	1.31
PHQ-9	1.56	0.96	1.31	0.99	1.41	0.97	1.27	1.10
Trait Anxiety	2.14	0.59	2.00	0.60	2.15	0.49	2.09	0.56
Need Satisfaction	4.73	0.98	5.06	0.98	4.80	0.93	5.14	0.94
Control over Play Behavior	1.86	0.76	1.86	0.69	1.91	0.80	1.63	0.69
Gamer Identity	5.34	2.92	6.00	2.81	4.71	3.16	3.43	2.67

	Game Genre							
	Cluster 1 most genres		Cluster 2 single-player		Cluster 3 FPS or action		Cluster 4 casual or puzzle	
	mean	SD	mean	SD	mean	SD	mean	SD
BDI-II	1.88	1.19	1.98	1.15	1.80	1.24	1.73	1.23
PHQ-9	1.53	1.01	1.51	0.92	1.37	1.00	1.31	0.99
Trait Anxiety	2.09	0.61	2.08	0.58	2.07	0.57	2.06	0.60
Need Satisfaction	4.97	0.95	4.82	0.98	4.77	0.95	5.07	1.05
Control over Play Behavior	1.84	0.63	2.11	0.80	1.86	0.77	1.61	0.63
Gamer Identity	5.61	2.90	7.58	2.36	5.39	2.87	3.50	2.17

Correlations Among Dependent Measures

There were significant correlations between the dependent measures. As expected, CPB was correlated with gamer identity ($r=.42$). Furthermore, BDI-II was positively correlated with PHQ-9 ($r=.88$; they both measure mental health) and TA ($r=.70$), and negatively correlated with need satisfaction ($r=-.66$). PHQ-9 was positively correlated with TA ($r=.71$), and negatively correlated with need satisfaction ($r=-.67$). TA was negatively correlated with need satisfaction ($r=-.79$).

Importance of Play

Frequency of play showed a significant relationship with CPB ($F_{2,485}=26.48$, $P<.001$, $\eta^2_p=.098$), and gamer identity ($F_{2,485}=61.00$, $P<.001$, $\eta^2_p=.20$). Bonferroni-corrected post hoc tests showed that frequency of play differed between all levels of CPB (all $P<.01$). The more people play, the more loss of control they experience; and on all levels of gamer identity (all $P<.01$), the more people play, the more they identify as a gamer.

In terms of genre preference, we found significant relationships with CPB ($F_{2,484}=6.35$, $P<.001$, $\eta^2_p=.038$), and gamer identity

($F_{2,484}=30.56$, $P<.001$, $\eta^2_p=.159$). For CPB, Bonferroni-corrected post hoc tests revealed that playing across common genres was related to a loss of CPB compared with playing (1) action or FPS-games or (2) casual or puzzle-games. There was no difference between single player games and the other three categories (all $P>.05$). Bonferroni-corrected post hoc tests for gamer identity showed the highest identity ratings for those who played all common genres, compared with the other three clusters (all $P<.001$). The single player and action or FPS clusters also showed higher identity ratings compared with the puzzle or casual-cluster (both $P<.001$), which was the genre cluster with the lowest overall ratings of gamer identity.

Platform preference showed a significant effect on gamer identity ($F_{2,484}=5.69$, $P<.001$, $\eta^2_p=.03$); but not on CPB ($P=.80$). Bonferroni-corrected post hoc analysis revealed that people who play on all devices identify more as a gamer, compared with people who play on mobile ($P<.001$), or console ($P<.001$) only. There was no difference between all devices and desktop ($P=.09$).

Indicators of Well-Being

Measures of Depression

Frequency of play showed a significant relationship with PHQ-9 ($F_{2,485}=5.78, P=.003, \eta^2_p=.023$), and BDI-II ($F_{2,485}=5.35, P=.01, \eta^2_p=.022$). For PHQ-9, Bonferroni-corrected post hoc tests revealed that people who play every day score significantly higher on the PHQ-9 compared with those who play less than a few times per week ($P<.001$), or play a few times per week ($P<.001$). For BDI-II, Bonferroni-corrected post hoc tests revealed that people who play every day score significantly higher than people who play less than a few times per week ($P<.001$). There was no significant relationship between genre preference and PHQ-9 ($P=.32$) or BDI-II ($P=.68$).

Furthermore, when we consider just those players who are classified as having significant depression ($n=35$) according to the BDI-II [79], we find that 12 play every day, 16 play a few times per week, and 7 play once per week or less.

Platform preferences showed a significant relationship with PHQ-9 ($F_{2,485}=2.79, P=.04, \eta^2_p=.017$), but not BDI-II ($P=.18$). Bonferroni-corrected post hoc tests revealed that people who used only a desktop computer scored higher on the PHQ-9 ($P=.03$) than people who use all devices.

Measures of Anxiety

TA shows no significant relationship with frequency of play ($P=.23$), platform preference ($P=.07$), or genre preference ($P=.99$).

Measures of Need Satisfaction

Need satisfaction showed no significant relationship with frequency of play ($P=.25$) or genre preference ($P=.53$). For platform preference, however, we see a significant relationship ($F_{2,485}=4.16, P=.01, \eta^2_p=.025$) with need satisfaction. Bonferroni-corrected post hoc tests revealed that people who play only on desktop computers score significantly lower on need satisfaction than those who play on all devices ($P=.004$). The mobile and console clusters showed no differences compared with the other clusters.

Discussion

Principal Findings and Comparison With Prior Work

In summary, our results for importance of play showed that people who play frequently identify more as a gamer and feel less in control over their play behavior. It is not a surprise that investing more in an activity makes this activity a bigger part of our life and is integrated more into our identity [83]. Additionally, we show that people who play across many genres—as opposed to a small set of closely related genres—identify more as gamers and experience less control over their play behavior. Furthermore, we show that desktop players identify more as gamers than players who use only mobile or console devices.

In terms of the effects on indicators of well-being, we showed that playing every day was associated with higher scores on the

BDI-II and PHQ-9, but did not affect TA or need satisfaction. It is not surprising that playing every day is associated with indicators of poorer mental health, as research on pathological gaming has found that low need satisfaction is associated with the experience of game addiction [64]. In addition, research on disordered patterns of play has found that people with low basic satisfaction of relatedness feel more compelled to play (ie, plan to play more in the future) but get less enjoyment out of the play experience [73].

Furthermore, desktop-only players scored higher on the PHQ-9 and lower on need satisfaction; however, genre preference was not significantly related to any indicator of well-being. This last result bodes well for intervention designers who wish to incorporate game-based elements into their digital interventions, as it appears that they are not limited in terms of the appeal of different game mechanics or genres and can use what works best for their particular intervention design. However, the finding that playing on desktop computers alone (as opposed to consoles, mobile devices, or all three) is associated with indicators of poorer mental health suggests that intervention designers might want to account for this preference when deciding on a platform for delivery.

Implications for the Design of Digital Interventions for Mental Health

In terms of the research questions that we set out to answer:

RQ1. Are people with depression or anxiety activating enough to play games? How frequently do they play?

The majority of our participants played games a few times a week or daily. In addition, PHQ-9 scores were higher for those who played daily than for those who played a few times a week or less, and BDI-II scores were higher for those who played daily than for those who played once a week or less.

Our sample suggests that people with depression are activating enough to play games at least a few times per week, and often daily.

RQ2. Are gamers with depression or anxiety part of the growing group who play games on a mobile platform or are they more likely to play in the comfort of their homes on a desktop or console platform?

Depression and low need satisfaction were both significantly associated with desktop-only play.

Although players with mental health issues use all types of platforms for playing games, there seems to be an increased prevalence of desktop-only play for people with depression that designers may want to consider when creating game-based interventions for mental health.

RQ3. Do players with depression or anxiety play casual games, or do they prefer more complex and immersive genres of games?

Genre preferences were not associated with any indicators of mental health.

Our results suggest that designers can choose from a range of game mechanics and genres in creating game-based

interventions for mental health, and should not feel limited to a specific genre or type of game.

The Benefits of Games for Mental Health Interventions

Our results suggest that games may be a valuable approach for the design of digital mental health interventions. Here, we discuss three main advantages of game-based intervention design for mental health based on our results and previous work [89].

Motivational Pull

We motivated the contribution of this paper by arguing that leveraging the motivational pull of games [39] may increase participation in a digital mental health intervention. For efficacious treatment, patients need to adhere to the intervention repeatedly over the long-term, and previous work has shown that participants who engage in training under their own volition exhibit better adherence, which may translate into greater efficacy [90]. In our previous work, we have sought out ways to use interaction design to increase intrinsic motivation with digital apps. We initially showed that how we see our own personality in a game (game-self) affects the motivation to play [91]. We then leveraged this finding to show that identification with an in-game avatar in a simple game leads to increased engagement with and invested effort in a boring task, also translating into significant differences in motivated behavior as measured by time spent in a free-choice task [92]. Additionally, we demonstrated that these motivational benefits lasted in a daily training task over the medium term (1 week) [34]. Together, this work demonstrates how we can use games to foster intrinsic motivation with apps and foreshadows how game-design could be used to facilitate longer-term engagement in interventions for the treatment of mental health issues by fostering intrinsic motivation.

Broad Appeal

Although people from different demographics may be more susceptible to mental health issues (eg, adolescents [93], lower socioeconomic status [3]), the prevalence of depression, anxiety, or personality disorders are not restricted to a specific demographic [94]; people of various ages, genders, socioeconomic statuses, and cultures are affected. Similarly, although games may have a reputation of appealing to the young male demographic, the appeal of digital games extends across a range of demographics. Our own sample in this study ($n=490$) shows how age and gender are not related to frequency of play. The Entertainment Software Association (ESA) [41] reports that although 26% of gamers are aged less than 18 years, 27% are over 50 years. Additionally, 44% of gamers are female—in fact, females over 18 years represent a significantly larger part of the game-playing population (33%) than boys less than 18 years (15%) [41].

Our previous research on computer games spans various ages of players, ranging from a focus on children with developmental disabilities [95] to the elderly who live in institutionalized care [96]. Although we sometimes focus our research on committed players who self-identify as gamers [97] or play specific games [98,99], our use of crowdsourcing platforms (such as AMT used in this study) to gather data and conduct experiments allows us to access a broad and representative sample of people who enjoy

playing a diverse range of digital games [34,43,92]. In creating game-based interventions for mental health issues, designers do not need to be concerned that their solutions will only appeal to a narrow group of people with specific gaming interests. Our results from this study suggest that the motivational pull of games described in the previous section applies broadly across the range of demographic groups who can benefit from apps that aim to improve mental health.

Accessibility

Because mental health issues are prevalent across demographic groups, it is important to ensure that people have access to treatment, independent of their work schedule, geographic location, or the capacity of the health care system. Games are generally accessible, independent of time of the day or location (thanks to the increased prevalence of mobile games). Because digital content can be delivered to any place that has Internet access, games are accessible in most geographical locations, which gives them advantages over traditional psychotherapy for countries with distributed and remote populations, such as in Nunavut—the territory in northern Canada that has a suicide rate 10 times greater than the rest of the country [11,12]. Although games are accessible in remote locations, they also hold advantages for addressing mental health in populated regions in which the capacity for treatments is exhausted. The shortage of available treatment resources can result in greater waiting time for patients [100] and undiagnosed mental issues in early stages [101], which may translate into tangible (eg, loss of work hours) and intangible (eg, loss of life quality) outcomes. Game-based interventions may help to bridge this gap. Finally, the convenience of in-home therapy offered by game-based approaches (as opposed to visiting a clinic) may expand the reach of mental health interventions, making treatment more accessible for those who need it [95].

Limitations

Our results demonstrate several important relationships between the habits and preferences of players and indicators of their mental health. However, there are several limitations that should be addressed through future work. First, the survey uses self-report measures of play behavior and is subject to all the biases present in self-report. Although the indicators of mental health also use self-report, the measures that we used are standard in terms of diagnosing depression (ie, BDI-II) and are recommended by the World Health Organization for assessing mental health (ie, the PHQ-9). Second, the survey was only available to be taken by the US residents, and thus our results are not generalizable beyond this particular context. Third, our results provide information on the play habits and preferences of people with varying levels of depression, anxiety, and need satisfaction to provide guidance to developers of digital interventions for mental health. Because we do not randomly assign people to experimental conditions, there is no intention in this paper of claiming causal links between game play habits and mental health. Research on pathological gaming [102-104] addresses the relationship between play and mental health from that perspective; our goal was to establish that people with depression and anxiety activate enough to play games from a variety of genres, and on a range of platforms. Finally, we only

examined how much people played and what their play preferences were in terms of platform and genre. It is possible that there are other factors regarding the context of play that are also important to understand, for example, Lemola et al [103] showed how habitual computer game playing at night (between 10 pm and 6 am) is associated with an increase in depression scores, even after controlling for the total time played, suggesting that the timing of play is also relevant in this context. In future work, we will investigate the patterns of play in addition to the habits and preferences of players.

Conclusions

In this paper, we use data from a Web-based survey (N=491) to describe the relationships between gameplay habits, gameplay preferences, and indicators of mental health by answering research questions about the differential gameplay habits of

people with varying levels of depression, anxiety, and need satisfaction. In general, we reveal that the vast majority of people in our sample play games a few times a week or more, and that playing daily is associated with indicators of depression. Desktop play was also associated with higher indicators of depression and lower need satisfaction. As expected, those who played more had identified more as a gamer and had less control over their gameplay.

Our results suggest that games are a suitable approach for mental health interventions as they are played broadly by people across a range of indicators of mental health, have strong motivational pull, and are accessible to players from a broad range of demographics. Our contribution is of interest to the community as we establish that games are an approach with great potential to add to the growing literature on digital interventions for improving mental health.

Acknowledgments

We thank NSERC and SWaGUR for funding, members of the Interaction Lab for support (in particular Colby Johansson for his efforts in creating BOF), and our participants for contributing data.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of one-item measure of gamer identity.

[PDF File (Adobe PDF File), 42KB - [jmir_v19i4e128_app1.pdf](#)]

References

1. Substance Abuse and Mental Health Association Service Administration. 2015 Sep. Behavioral health trends in the United States: results from the 2014 National Survey on Drug Use and Health URL: <https://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf> [accessed 2017-03-20] [WebCite Cache ID 6p6UjY9Kq]
2. National Institute of Mental Health. Any anxiety disorder among adults URL: <https://www.nimh.nih.gov/health/statistics/prevalence/any-anxiety-disorder-among-adults.shtml> [accessed 2017-02-24] [WebCite Cache ID 6oWHsmmQB]
3. Mental Health America. The State of Mental Health in America URL: <http://www.mentalhealthamerica.net/issues/state-mental-health-america> [accessed 2017-03-20] [WebCite Cache ID 6p6XTispC]
4. Friedli L. World Health Organization. 2009. Mental health, resilience and inequalities URL: http://www.euro.who.int/data/assets/pdf_file/0012/100821/E92227.pdf [accessed 2017-03-20] [WebCite Cache ID 6p6Vg4LRr]
5. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry* 2016 May;3(5):415-424 [FREE Full text] [doi: [10.1016/S2215-0366\(16\)30024-4](https://doi.org/10.1016/S2215-0366(16)30024-4)] [Medline: [27083119](#)]
6. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Qual Life Res* 2004 Mar;13(2):299-310. [Medline: [15085902](#)]
7. Nock MK, Hwang I, Sampson N, Kessler RC, Angermeyer M, Beautrais A, et al. Cross-national analysis of the associations among mental disorders and suicidal behavior: findings from the WHO World Mental Health Surveys. *PLoS Med* 2009 Aug;6(8):e1000123 [FREE Full text] [doi: [10.1371/journal.pmed.1000123](https://doi.org/10.1371/journal.pmed.1000123)] [Medline: [19668361](#)]
8. Mościcki EK. Epidemiology of completed and attempted suicide: toward a framework for prevention. *Clin Neurosci Res* 2001 Nov;1(5):310-323. [doi: [10.1016/S1566-2772\(01\)00032-9](https://doi.org/10.1016/S1566-2772(01)00032-9)]
9. Weir E, Wallington T. Suicide: the hidden epidemic. *CMAJ* 2001 Sep 04;165(5):634, 636 [FREE Full text] [Medline: [11563217](#)]
10. Mental Health Atlas. Geneva, Switzerland: World Health Organization; Dec 17, 2015.
11. Canadian Mental Health Association. 2016. Suicide statistics URL: http://toronto.cmha.ca/mental_health/suicide-statistics/ [accessed 2017-02-24] [WebCite Cache ID 6oWHQrfmZ]

12. Elyse S. Canadian Broadcasting Corporation. Nunavut's suicide inquest prompts pleas for action, but progress 'slow' URL: <http://www.cbc.ca/news/canada/north/nunavut-s-suicide-inquest-prompts-pleas-for-action-but-progress-slow-1.3244367> [accessed 2017-02-24] [WebCite Cache ID 6oWIAzC0t]
13. Cuijpers P, Riper H, Andersson G. Internet-based treatment of depression. *Curr Opin Psychol* 2015 Aug;4:131-135. [doi: [10.1016/j.copsyc.2014.12.026](https://doi.org/10.1016/j.copsyc.2014.12.026)]
14. Calvo R, Dinakar K, Picard R, Maes P. Computing in mental health. In: *Proceedings of the SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems*. . ACM.: ACM; 2016 May 07 Presented at: SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems; May 07-12, 2016; San Jose, CA p. 3438-3445. [doi: [10.1145/2851581.2856463](https://doi.org/10.1145/2851581.2856463)]
15. Gillan CM, Daw ND. Taking psychiatry research online. *Neuron* 2016 Jul 06;91(1):19-23. [doi: [10.1016/j.neuron.2016.06.002](https://doi.org/10.1016/j.neuron.2016.06.002)] [Medline: [27387647](https://pubmed.ncbi.nlm.nih.gov/27387647/)]
16. Regalado A. Technology Review. Why America's top mental health researcher joined alphabet URL: <https://www.technologyreview.com/s/541446/why-americas-top-mental-health-researcher-joined-alphabet/> [accessed 2017-02-24] [WebCite Cache ID 6oWI53FkZ]
17. Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of IntelliCare: a publicly available suite of mental health and well-being apps. *Internet Interv* 2016 May;4(2):152-158 [FREE Full text] [doi: [10.1016/j.invent.2016.06.003](https://doi.org/10.1016/j.invent.2016.06.003)] [Medline: [27398319](https://pubmed.ncbi.nlm.nih.gov/27398319/)]
18. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. *BMJ Open* 2014;4(1):e003440 [FREE Full text] [doi: [10.1136/bmjopen-2013-003440](https://doi.org/10.1136/bmjopen-2013-003440)] [Medline: [24413342](https://pubmed.ncbi.nlm.nih.gov/24413342/)]
19. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res* 2015;17(7):e175 [FREE Full text] [doi: [10.2196/jmir.4273](https://doi.org/10.2196/jmir.4273)] [Medline: [26180009](https://pubmed.ncbi.nlm.nih.gov/26180009/)]
20. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013;15(11):e247 [FREE Full text] [doi: [10.2196/jmir.2791](https://doi.org/10.2196/jmir.2791)] [Medline: [24240579](https://pubmed.ncbi.nlm.nih.gov/24240579/)]
21. Kuhn E, Greene C, Hoffman J, Nguyen T, Wald L, Schmidt J, et al. Preliminary evaluation of PTSD Coach, a smartphone app for post-traumatic stress symptoms. *Mil Med* 2014 Jan;179(1):12-18. [doi: [10.7205/MILMED-D-13-00271](https://doi.org/10.7205/MILMED-D-13-00271)] [Medline: [24402979](https://pubmed.ncbi.nlm.nih.gov/24402979/)]
22. Morris RR, Schueller SM, Picard RW. Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. *J Med Internet Res* 2015;17(3):e72 [FREE Full text] [doi: [10.2196/jmir.4167](https://doi.org/10.2196/jmir.4167)] [Medline: [25835472](https://pubmed.ncbi.nlm.nih.gov/25835472/)]
23. Andersson G, Carlbring P, Holmström A, Sparthán E, Furmark T, Nilsson-Ihrfelt E, et al. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. *J Consult Clin Psychol* 2006 Aug;74(4):677-686. [doi: [10.1037/0022-006X.74.4.677](https://doi.org/10.1037/0022-006X.74.4.677)] [Medline: [16881775](https://pubmed.ncbi.nlm.nih.gov/16881775/)]
24. Halmetoja CO, Malmquist A, Carlbring P, Andersson G. Experiences of internet-delivered cognitive behavior therapy for social anxiety disorder four years later: a qualitative study. *Internet Interv* 2014 Jul;1(3):158-163. [doi: [10.1016/j.invent.2014.08.001](https://doi.org/10.1016/j.invent.2014.08.001)]
25. Mathiasen K, Riper H, Ehlers LH, Valentin JB, Rosenberg NK. Internet-based CBT for social phobia and panic disorder in a specialised anxiety clinic in routine care: Results of a pilot randomised controlled trial. *Internet Interv* 2016 May;4:92-98. [doi: [10.1016/j.invent.2016.03.001](https://doi.org/10.1016/j.invent.2016.03.001)]
26. Ivarsson D, Blom M, Hesser H, Carlbring P, Enderby P, Nordberg R, et al. Guided internet-delivered cognitive behavior therapy for post-traumatic stress disorder: a randomized controlled trial. *Internet Interv* 2014 Mar;1(1):33-40. [doi: [10.1016/j.invent.2014.03.002](https://doi.org/10.1016/j.invent.2014.03.002)]
27. Knaevelsrud C, Maercker A. Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: a randomized controlled clinical trial. *BMC Psychiatry* 2007;7:13 [FREE Full text] [doi: [10.1186/1471-244X-7-13](https://doi.org/10.1186/1471-244X-7-13)] [Medline: [17442125](https://pubmed.ncbi.nlm.nih.gov/17442125/)]
28. Kayrouz R, Dear BF, Johnston L, Gandy M, Fogliati VJ, Sheehan J, et al. A feasibility open trial of guided internet-delivered cognitive behavioural therapy for anxiety and depression amongst Arab Australians. *Internet Interv* 2015 Mar;2(1):32-38. [doi: [10.1016/j.invent.2014.12.001](https://doi.org/10.1016/j.invent.2014.12.001)]
29. Meyer B, Bierbrodt J, Schröder J, Berger T, Beevers CG, Weiss M, et al. Effects of an internet intervention (Deprexis) on severe depression symptoms: randomized controlled trial. *Internet Interv* 2015 Mar;2(1):48-59. [doi: [10.1016/j.invent.2014.12.003](https://doi.org/10.1016/j.invent.2014.12.003)]
30. Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. *Med J Aust* 2010 Jun 7;192(11 Suppl):S4-11. [Medline: [20528707](https://pubmed.ncbi.nlm.nih.gov/20528707/)]
31. Cook JE, Doyle C. Working alliance in online therapy as compared to face-to-face therapy: preliminary results. *Cyberpsychol Behav* 2002 Apr;5(2):95-105. [Medline: [12025884](https://pubmed.ncbi.nlm.nih.gov/12025884/)]

32. Barazzone N, Cavanagh K, Richards DA. Computerized cognitive behavioural therapy and the therapeutic alliance: a qualitative enquiry. *Br J Clin Psychol* 2012 Nov;51(4):396-417. [doi: [10.1111/j.2044-8260.2012.02035.x](https://doi.org/10.1111/j.2044-8260.2012.02035.x)] [Medline: [23078210](https://pubmed.ncbi.nlm.nih.gov/23078210/)]
33. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp Educ Psychol* 2000 Jan;25(1):54-67. [doi: [10.1006/ceps.1999.1020](https://doi.org/10.1006/ceps.1999.1020)] [Medline: [10620381](https://pubmed.ncbi.nlm.nih.gov/10620381/)]
34. Birk M, Mandryk R, Atkins C. The motivational push of games: the interplay of intrinsic motivation external rewards in games for training. In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*.: ACM; 2016 Oct 16 Presented at: Annual Symposium on Computer-Human Interaction in Play; 2016; Austin, TX p. 291-303. [doi: [10.1145/2967934.2968091](https://doi.org/10.1145/2967934.2968091)]
35. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res* 2012;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](https://doi.org/10.2196/jmir.2104)] [Medline: [23151820](https://pubmed.ncbi.nlm.nih.gov/23151820/)]
36. Robinson E, Titov N, Andrews G, McIntyre K, Schwencke G, Solley K. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010;5(6):e10942 [FREE Full text] [doi: [10.1371/journal.pone.0010942](https://doi.org/10.1371/journal.pone.0010942)] [Medline: [20532167](https://pubmed.ncbi.nlm.nih.gov/20532167/)]
37. Deci EL, Koestner R, Ryan RM. Extrinsic rewards and intrinsic motivation in education: reconsidered once again. *Rev Educ Res* 2001 Jan 01;71(1):1-27. [doi: [10.3102/00346543071001001](https://doi.org/10.3102/00346543071001001)]
38. 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](https://doi.org/10.1207/s15327965pli1104_01)]
39. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. *Motiv Emot* 2006 Nov 29;30(4):344-360. [doi: [10.1007/s11031-006-9051-8](https://doi.org/10.1007/s11031-006-9051-8)]
40. Dring C. The Market for Computer & Video Games. More money is spent on games than movies and music combined, says IHS (web resource) URL: <http://www.mcvuk.com/news/read/more-money-is-spent-on-games-than-movies-and-music-combined-says-ihs/0151059> [accessed 2017-02-24] [WebCite Cache ID 6oWIPvWBN]
41. Entertainment Software Association. 2015. Essential facts URL: <http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf> [accessed 2017-02-24] [WebCite Cache ID 6oWIIsIjq]
42. Entertainment Software Association. 2014. Essential facts URL: http://www.theesa.com/wp-content/uploads/2014/10/ESA_EF_2014.pdf [accessed 2017-02-24] [WebCite Cache ID 6oWINISiy]
43. Birk MV, Mandryk R, Miller M, Gerling K. How self-esteem shapes our interactions with play technologies. In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 2015 Oct 05 Presented at: Annual Symposium on Computer-Human Interaction in Play; 2015 Oct 05-07; London, England p. 35-45. [doi: [10.1145/2793107.2793111](https://doi.org/10.1145/2793107.2793111)]
44. Johnson D, Gardner J, Sweetser P. Motivations for videogame play: predictors of time spent playing. *Comput Human Behav* 2016 Oct;63:805-812. [doi: [10.1016/j.chb.2016.06.028](https://doi.org/10.1016/j.chb.2016.06.028)]
45. Deterding S. Contextual autonomy support in video game play: a grounded theory. In: *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*.: ACM; 2016 May 12 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 3931-3943. [doi: [10.1145/2858036.2858395](https://doi.org/10.1145/2858036.2858395)]
46. Gerling K, Birk M, Mandryk R, Doucette A. The effects of graphical fidelity on player experience. In: *Proceedings of the International Conference on Making Sense of Converging Media*.: ACM; 2013 Oct 01 Presented at: International Conference on Making Sense of Converging Media; October 01-04, 2013; Tampere, Finland p. 229-237. [doi: [10.1145/2523429.2523473](https://doi.org/10.1145/2523429.2523473)]
47. Iacovides I, Cox AL, McAndrew P, Aczel J, Scanlon E. Game-play breakdowns and breakthroughs: exploring the relationship between action, understanding, and involvement. *Hum-Comput Int* 2014 Dec 22;30(3-4):202-231. [doi: [10.1080/07370024.2014.987347](https://doi.org/10.1080/07370024.2014.987347)]
48. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. 2011 Sep 28 Presented at: International Academic MindTrek Conference; September 28-30, 2011; Tampere, Finland p. 9-15. [doi: [10.1145/2181037.2181040](https://doi.org/10.1145/2181037.2181040)]
49. Nicholson S. A recipe for meaningful gamification. In: Reiners T, Wood LC, editors. *Gamification in education and business*. New York, NY: Springer International Publishing; 2014.
50. Fogg BJ. Persuasive technology. *Ubiquity* 2002 Dec 01;2002(December):2. [doi: [10.1145/764008.763957](https://doi.org/10.1145/764008.763957)]
51. Orji R, Vassileva J, Mandryk RL. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. *User Model User-Adap Inter* 2014 Jul 14;24(5):453-498. [doi: [10.1007/s11257-014-9149-8](https://doi.org/10.1007/s11257-014-9149-8)]
52. Dennis TA, O'Toole L. Mental health on the go: effects of a gamified attention bias modification mobile application in trait anxious adults. *Clin Psychol Sci* 2014 Sep 01;2(5):576-590 [FREE Full text] [doi: [10.1177/2167702614522228](https://doi.org/10.1177/2167702614522228)] [Medline: [26029490](https://pubmed.ncbi.nlm.nih.gov/26029490/)]
53. Kahn J, Ducharme P, Rotenberg A, Gonzalez-Heydrich J. “RAGE-Control”: a game to build emotional strength. *Games Health J* 2013 Feb;2(1):53-57. [doi: [10.1089/g4h.2013.0007](https://doi.org/10.1089/g4h.2013.0007)] [Medline: [26196556](https://pubmed.ncbi.nlm.nih.gov/26196556/)]

54. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. *Pediatrics* 2008 Aug;122(2):e305-e317 [[FREE Full text](#)] [doi: [10.1542/peds.2007-3134](#)] [Medline: [18676516](#)]
55. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen Mathijs F G. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ* 2012;344:e2598 [[FREE Full text](#)] [Medline: [22517917](#)]
56. Ritterfeld U, Cody M, Vorderer P. *Serious Games: Mechanisms and Effects*. New York: Routledge; 2009.
57. Scholten H, Malmberg M, Lobel A, Engels RC, Granic I. A randomized controlled trial to test the effectiveness of an immersive 3d video game for anxiety prevention among adolescents. *PLoS One* 2016;11(1):e0147763 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0147763](#)] [Medline: [26816292](#)]
58. Schoneveld E, Malmberg M, Lichtwarck-Aschoff A, Verheijen G, Engels R, Granic I. A neurofeedback video game (MindLight) to prevent anxiety in children: A randomized controlled trial. *Comput Human Behav* 2016 Oct;63:321-333. [doi: [10.1016/j.chb.2016.05.005](#)]
59. Kato PM. Video games in health care: closing the gap. *Rev Gen Psycho* 2010;14(2):113-121. [doi: [10.1037/a0019441](#)]
60. Hazen RA, Vasey MW, Schmidt NB. Attentional retraining: a randomized clinical trial for pathological worry. *J Psychiatr Res* 2009 Mar;43(6):627-633. [doi: [10.1016/j.jpsychires.2008.07.004](#)] [Medline: [18722627](#)]
61. Hakamata Y, Lissek S, Bar-Haim Y, Britton JC, Fox NA, Leibenluft E, et al. Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biol Psychiatry* 2010 Dec 1;68(11):982-990 [[FREE Full text](#)] [doi: [10.1016/j.biopsych.2010.07.021](#)] [Medline: [20887977](#)]
62. Li J, Theng Y, Foo S. Game-based digital interventions for depression therapy: a systematic review and meta-analysis. *Cyberpsychol Behav Soc Netw* 2014 Aug;17(8):519-527. [doi: [10.1089/cyber.2013.0481](#)] [Medline: [24810933](#)]
63. De Grove F, Cauberghe V, Van Looy J. Development and validation of an instrument for measuring individual motives for playing digital games. *Media Psychol* 2014 Jun 25;19(1):101-125. [doi: [10.1080/15213269.2014.902318](#)]
64. Lemmens JS, Valkenburg PM, Peter J. Development and validation of a game addiction scale for adolescents. *Media Psychol* 2009 Mar 05;12(1):77-95. [doi: [10.1080/15213260802669458](#)]
65. Hilgard J, Engelhardt CR, Bartholow BD. Individual differences in motives, preferences, and pathology in video games: the gaming attitudes, motives, and experiences scales (GAMES). *Front Psychol* 2013;4:608 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2013.00608](#)] [Medline: [24058355](#)]
66. Ferguson CJ, Olson CK. Friends, fun, frustration and fantasy: child motivations for video game play. *Motiv Emot* 2012 Mar 14;37(1):154-164. [doi: [10.1007/s11031-012-9284-7](#)]
67. Reinecke L. Games and recovery. *J Media Psychol* 2009 Jan;21(3):126-142. [doi: [10.1027/1864-1105.21.3.126](#)]
68. Reinecke L, Klatt J, Krämer NC. Entertaining media use and the satisfaction of recovery needs: recovery outcomes associated with the use of interactive and noninteractive entertaining media. *Media Psychol* 2011 May 31;14(2):192-215. [doi: [10.1080/15213269.2011.573466](#)]
69. Bowman ND, Tamborini R. "In the Mood to Game": selective exposure and mood management processes in computer game play. *New Media & Soc* 2015 Mar;17(3):375-393. [doi: [10.1177/1461444813504274](#)]
70. Bowman ND, Tamborini R. Task demand and mood repair: The intervention potential of computer games. *New Media & Soc* 2012 Dec;14(8):1339-1357. [doi: [10.1177/1461444812450426](#)]
71. Johnson D, Wyeth P, Sweetser P. The people-game-play model for understanding videogames' impact on wellbeing. In: *Proceedings of the IEEE International Games Innovation Conference (IGIC)*.: IEEE; 2013 Sep 23 Presented at: IEEE International Games Innovation Conference (IGIC); September 23-25, 2013; Vancouver, British Columbia, Canada. [doi: [10.1109/IGIC.2013.6659143](#)]
72. Granic I, Lobel A, Engels RC. The benefits of playing video games. *Am Psychol* 2014 Jan;69(1):66-78. [doi: [10.1037/a0034857](#)] [Medline: [24295515](#)]
73. Przybylski AK, Weinstein N, Ryan RM, Rigby CS. Having to versus wanting to play: background and consequences of harmonious versus obsessive engagement in video games. *Cyberpsychol Behav* 2009 Oct;12(5):485-492. [doi: [10.1089/cpb.2009.0083](#)] [Medline: [19772442](#)]
74. Vallerand RJ, Losier GF. An integrative analysis of intrinsic and extrinsic motivation in sport. *J Appl Sport Psychol* 1999 Mar;11(1):142-169. [doi: [10.1080/10413209908402956](#)]
75. Cuijpers P, van Straten A, Warmerdam L. Behavioral activation treatments of depression: a meta-analysis. *Clin Psychol Rev* 2007 Apr;27(3):318-326. [doi: [10.1016/j.cpr.2006.11.001](#)] [Medline: [17184887](#)]
76. MacPhillamy DJ, Lewinsohn PM. Depression as a function of levels of desired and obtained pleasure. *J Abnorm Psychol* 1974 Dec;83(6):651-657. [Medline: [4155711](#)]
77. Meinschmidt G, Lee J, Stalujanis E, Belardi A, Oh M, Jung EK, et al. Smartphone-based psychotherapeutic micro-interventions to improve mood in a real-world setting. *Front Psychol* 2016;7:1112 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2016.01112](#)] [Medline: [27516747](#)]
78. Rohrl D. International Game Developers Association. 2008. Casual games white paper URL: https://www.igda.org/resource/collection/BCB11E9B-13E6-40D0-B390-952B5E11D35A/IGDA_Casual_Games_White_Paper_2008.pdf [accessed 2017-03-20] [[WebCite Cache ID 6oWHfcU05](#)]

79. Beck A, Steer T, Brown G. Beck depression inventory-II. San Antonio 1996;78(2):490-498.
80. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. *J Gen Intern Med* 2001 Sep;16(9):606-613. [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)]
81. Spielberger C. State-Trait anxiety inventory. In: Weiner EB, Craighead WE, editors. *Corsini Encyclopedia of Psychology*. Hoboken, NJ: John Wiley & Sons, Inc; Jan 30, 2010.
82. Gagne M. Autonomy support and need satisfaction in the motivation and well-being of gymnasts. *J Appl Sport Psychol* 2003 Oct;15(4):372-390. [doi: [10.1080/714044203](https://doi.org/10.1080/714044203)]
83. Pelham BW, Swann WB. From self-conceptions to self-worth: on the sources and structure of global self-esteem. *J Pers Soc Psychol* 1989 Oct;57(4):672-680. [Medline: [2795437](https://pubmed.ncbi.nlm.nih.gov/2795437/)]
84. Field A. *Discovering Statistics Using IBM SPSS Statistics*. Thousand Oaks, CA: Sage Publications; 2013.
85. Buhrmester M, Kwang T, Gosling SD. Amazon's mechanical turk: a new source of inexpensive, yet high-quality, data? *Perspect Psychol Sci* 2011 Jan;6(1):3-5. [doi: [10.1177/1745691610393980](https://doi.org/10.1177/1745691610393980)] [Medline: [26162106](https://pubmed.ncbi.nlm.nih.gov/26162106/)]
86. Kittur A, Chi E, Suh B. Crowdsourcing user studies with mechanical turk. In: *Proceedings of the SIGCHI conference on human factors in computing systems*.: ACM; 2008 Apr 05 Presented at: SIGCHI conference on human factors in computing systems; April 05-10, 2008; Florence, Italy p. 453-456. [doi: [10.1145/1357054.1357127](https://doi.org/10.1145/1357054.1357127)]
87. Shapiro DN, Chandler J, Mueller PA. Using mechanical turk to study clinical populations. *Clinical Psychological Science* 2013 Jan 31;1(2):213-220. [doi: [10.1177/2167702612469015](https://doi.org/10.1177/2167702612469015)]
88. Reinecke L. Games at work: the recreational use of computer games during working hours. *Cyberpsychol Behav* 2009 Aug;12(4):461-465. [doi: [10.1089/cpb.2009.0010](https://doi.org/10.1089/cpb.2009.0010)] [Medline: [19619038](https://pubmed.ncbi.nlm.nih.gov/19619038/)]
89. Birk M, Mandryk R. SIGCHI'16 Workshop on Computing and Mental Health. 2016 May 08. The Benefits of Digital Games for the Assessment Treatment of Mental Health URL: http://mentalhealth.media.mit.edu/wp-content/uploads/sites/46/2016/04/chi2016_mentalHealth_birk_mandryk_camera_ready.pdf [accessed 2017-03-19] [WebCite Cache ID 6p6U1vGdc]
90. Ryan RM, Lynch MF, Vansteenkiste M, Deci EL. motivation and autonomy in counseling, psychotherapy, and behavior change: a look at theory and practice 1ψ7. *Couns Psychol* 2010 Feb 10;39(2):193-260. [doi: [10.1177/0011000009359313](https://doi.org/10.1177/0011000009359313)]
91. Birk M, Mandryk R. Control your game-self: effects of controller type on enjoyment, motivation, and personality in game. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.: ACM; 2013 Apr 27 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27 - May 02, 2013; Paris, France p. 685-694. [doi: [10.1145/2470654.2470752](https://doi.org/10.1145/2470654.2470752)]
92. Birk M, Atkins C, Bowey J, Mandryk R. Fostering Intrinsic Motivation through Avatar Identification in Digital Games. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.: ACM; 2016 May 07 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 07-12, 2016; San Jose, CA p. 7-12. [doi: [10.1145/2858036.2858062](https://doi.org/10.1145/2858036.2858062)]
93. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Arch Gen Psychiatry* 2003 Aug;60(8):837-844. [doi: [10.1001/archpsyc.60.8.837](https://doi.org/10.1001/archpsyc.60.8.837)] [Medline: [12912767](https://pubmed.ncbi.nlm.nih.gov/12912767/)]
94. Marcus M, Yasamy MT, van OM, Chisholm D, Saxena S. World Health Organization. 2012. Depression: a global public health concern URL: http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf [accessed 2017-03-19] [WebCite Cache ID 6p6XIEssZ]
95. Mandryk R, Dielschneider S, Kalyn M. Games as neurofeedback training for children with FASD. In: *Proceedings of the International Conference on Interaction Design Children*.: ACM; 2013 Jun 24 Presented at: Proceedings of the International Conference on Interaction Design Children; June 24-27, 2013; New York, NY p. 165-172. [doi: [10.1145/2485760.2485762](https://doi.org/10.1145/2485760.2485762)]
96. Gerling K, Mandryk R, Linehan C. Long-term use of motion-based video games in care home settings. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.: ACM; 2015 Apr 18 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 18-23, 2015; Seoul, Korea p. 18-23. [doi: [10.1145/2702123.2702125](https://doi.org/10.1145/2702123.2702125)]
97. Johanson C, Mandryk R. Scaffolding player location awareness through audio cues in first-person shooters. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.: ACM; 2016 May 07 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 07-12, 2016; San Jose, CA p. 3450-3461. [doi: [10.1145/2858036.2858172](https://doi.org/10.1145/2858036.2858172)]
98. Livingston I, Gutwin C, Mandryk R, Birk M. How players value their characters in world of warcraft. In: *Proceedings of the conference on Computer supported cooperative work & social computing*.: ACM; 2014 Feb 15 Presented at: Conference on Computer supported cooperative work & social computing; February 15-19, 2014; Baltimore, MD p. 1333-1343. [doi: [10.1145/2531602.2531661](https://doi.org/10.1145/2531602.2531661)]
99. Birk M, Toker D, Mandryk R, Conati C. modeling motivation in a social network game using player-centric traits and personality traits. In: *Proceedings of the 23rd International Conference on User Modeling, Adaptation, and Personalization*.: Springer International Publishing; 2015 Presented at: International Conference on User Modeling, Adaptation, and Personalization; July 3, 2015; Dublin, Ireland p. 18-30. [doi: [10.1007/978-3-319-20267-9_2](https://doi.org/10.1007/978-3-319-20267-9_2)]
100. Canadian Institute for Health Information. Health care in Canada: a focus on wait times URL: https://secure.cihi.ca/free_products/HCIC2012-FullReport-ENweb.pdf [accessed 2017-02-24] [WebCite Cache ID 6oWHWaY19]
101. Kessler D, Bennewith O, Lewis G, Sharp D. Detection of depression and anxiety in primary care: follow up study. *BMJ* 2002 Nov 02;325(7371):1016-1017 [FREE Full text] [Medline: [12411363](https://pubmed.ncbi.nlm.nih.gov/12411363/)]

102. Anderson CA, Shibuya A, Ihori N, Swing EL, Bushman BJ, Sakamoto A, et al. Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. *Psychol Bull* 2010 Mar;136(2):151-173. [doi: [10.1037/a0018251](https://doi.org/10.1037/a0018251)] [Medline: [20192553](https://pubmed.ncbi.nlm.nih.gov/20192553/)]
103. Lemola S, Brand S, Vogler N, Perkinson-Gloor N, Allemann M, Grob A. Habitual computer game playing at night is related to depressive symptoms. *Pers Individ Dif* 2011 Jul;51(2):117-122. [doi: [10.1016/j.paid.2011.03.024](https://doi.org/10.1016/j.paid.2011.03.024)]
104. Van Rooij AJ, Schoenmakers TM, Vermulst AA, Van den Eijnden RJ, Van de Mheen D. Online video game addiction: identification of addicted adolescent gamers. *Addiction* 2011 Jan;106(1):205-212. [doi: [10.1111/j.1360-0443.2010.03104.x](https://doi.org/10.1111/j.1360-0443.2010.03104.x)] [Medline: [20840209](https://pubmed.ncbi.nlm.nih.gov/20840209/)]

Abbreviations

ABMT: attention-bias modification training
AMT: Amazon Mechanical Turk
ANOVA: analysis of variance
BDI-II: Beck's Depression Inventory
BPNS: basic psychological needs satisfaction
CBT: cognitive behavioral therapy
CPB: control over play behavior
ESA: Entertainment Software Association
FPS: first-person shooter
HIT: Human Intelligence Task
MANOVA: multivariate analysis of variance
MMORPG: massively multiplayer online role-playing game
MOBA: Multiplayer online battle arena
PHQ-9: Patient Health Questionnaire
PTSD: post-traumatic stress disorder
RCT: randomized controlled trial
RPG: role-playing game
SDT: self-determination theory
STAI: State-Trait Anxiety Scale
TA: trait anxiety
US: United States

Edited by R Calvo, J Torous; submitted 28.10.16; peer-reviewed by P Bhoopathi, J Li; comments to author 06.02.17; accepted 18.02.17; published 20.04.17.

Please cite as:

Mandryk RL, Birk MV

Toward Game-Based Digital Mental Health Interventions: Player Habits and Preferences

J Med Internet Res 2017;19(4):e128

URL: <http://www.jmir.org/2017/4/e128/>

doi: [10.2196/jmir.6906](https://doi.org/10.2196/jmir.6906)

PMID: [28428161](https://pubmed.ncbi.nlm.nih.gov/28428161/)

©Regan Lee Mandryk, Max Valentin Birk. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 20.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Analyzing and Predicting User Participations in Online Health Communities: A Social Support Perspective

Xi Wang¹, MSc; Kang Zhao^{1,2}, PhD; Nick Street^{1,2}, PhD

¹Interdisciplinary Graduate Program in Informatics, The University of Iowa, Iowa City, IA, United States

²Department of Management Sciences, The University of Iowa, Iowa City, IA, United States

Corresponding Author:

Kang Zhao, PhD

Department of Management Sciences

The University of Iowa

S224 PBB

Iowa City, IA, 52242

United States

Phone: 1 3193353831

Fax: 1 3193350297

Email: kang-zhao@uiowa.edu

Abstract

Background: Online health communities (OHCs) have become a major source of social support for people with health problems. Members of OHCs interact online with similar peers to seek, receive, and provide different types of social support, such as informational support, emotional support, and companionship. As active participations in an OHC are beneficial to both the OHC and its users, it is important to understand factors related to users' participations and predict user churn for user retention efforts.

Objective: This study aimed to analyze OHC users' Web-based interactions, reveal which types of social support activities are related to users' participation, and predict whether and when a user will churn from the OHC.

Methods: We collected a large-scale dataset from a popular OHC for cancer survivors. We used text mining techniques to decide what kinds of social support each post contained. We illustrated how we built text classifiers for 5 different social support categories: seeking informational support (SIS), providing informational support (PIS), seeking emotional support (SES), providing emotional support (PES), and companionship (COM). We conducted survival analysis to identify types of social support related to users' continued participation. Using supervised machine learning methods, we developed a predictive model for user churn.

Results: Users' behaviors to PIS, SES, and COM had hazard ratios significantly lower than 1 (0.948, 0.972, and 0.919, respectively) and were indicative of continued participations in the OHC. The churn prediction model based on social support activities offers accurate predictions on whether and when a user will leave the OHC.

Conclusions: Detecting different types of social support activities via text mining contributes to better understanding and prediction of users' participations in an OHC. The outcome of this study can help the management and design of a sustainable OHC via more proactive and effective user retention strategies.

(*J Med Internet Res* 2017;19(4):e130) doi:[10.2196/jmir.6834](https://doi.org/10.2196/jmir.6834)

KEYWORDS

social support; machine learning; community networks; patient engagement; prediction

Introduction

Overview

Nowadays more and more people use the Internet to satisfy their health-related needs. According to a study by the Pew Research Center, 80% of adult Internet users in the United States use the Internet for health-related purposes. Among them, 34% read health-related experiences or comments from others [1]. Online

health communities (OHCs) offer a venue for people to interact with peers facing similar health problems. Modern OHCs have incorporated many ways for communication and health management, such as listserv, discussion forums, private messaging, chat rooms, blogs, friend subscriptions, health tracking tools, and so on. OHCs range from general-purpose communities, such as MedHelp and PatientsLikeMe, to those dedicated to a specific health issue, such as diabetes and

smoking. Many OHCs host their own websites, whereas some are built on existing social networking services, such as Facebook. Many studies have revealed the advantages and disadvantages of OHCs compared with traditional offline support groups [2]. Although OHCs may face problems such as sporadic membership of active users, inaccurate information, deception, and insincerity of strangers [3-5], they also offer tremendous benefits such as broad reach, availability, and anonymity. Consequently, OHCs have gained popularity in recent years, and it is estimated that 5% of all Internet users participated in an OHC [6].

Studies of user behaviors in OHCs are valuable in several ways. First, outcomes of these studies can inform better management and design of a successful OHC, which can help to promote new treatments and healthy lifestyles and reveal adverse drug effects [7-9]. Like other online communities, successful OHCs would like to encourage users' participations and prevent users' churn (ie, leaving a community), because one of the keys for the success of an online community is active participations from and relationship building among its members [10,11]. In contrast, poor participations and transient membership can lead to the termination or failure of an online community [12]. Specifically, OHCs will not be sustainable if its users keep leaving because most of the social support can only be provided by active members of the OHC.

Second, a successful and sustainable OHC will provide more benefits to individual users. This is because a user's continued participation in an OHC can be helpful and therapeutic [13-18]. On one hand, receiving such support can be empowering [19] and help patients adjust to the stress of living with and fighting against their diseases [20,21]. The support they receive online can also improve their offline life and health management [22]. On the other hand, besides receiving support from others, staying in an OHC and providing support to others can be beneficial to providers as well [20]. There is actually a positive relationship between posting frequency and psychosocial well-being [23]. In other words, a user's continued participation in an OHC can help herself or himself as well as others. Admittedly, for some individuals who have received satisfactory support from an OHC or recovered from the disease, leaving the OHC may not be a bad thing for themselves. However, even though user-generated information about a disease will still be available on the Internet to new OHC members, most of the psychosocial benefits for individual users cannot be achieved if the exodus of experienced users in the OHC keeps happening, leaving new members stranded [23]. In fact, providing assistance for new members from experienced members and reminding members to participate continuously are also key factors for the success of online communities [12]. Therefore, better understanding and accurate prediction of users' participations in OHCs can help to build and sustain a successful OHC through improved community design, management, and user retention.

As social support is a pillar of OHCs, a natural question to ask would be: when it comes to users' participations, are a user's Web-based activities in different types of social support related to her or his continued participation in an OHC? If so, can we predict whether and when a user will churn from an OHC based on these social support activities? Despite the large amount of

research on social support in OHCs, few studies have answered this question systematically by examining users' seeking, receiving, and provision of various types of social support from large-scale datasets. An explanatory model [24] suggested that receiving more emotional support is associated with users' longer stay in an OHC. However, the types of social support investigated were limited and only the receiving of support was considered, while we mentioned earlier that providing social support is also important and beneficial. Analyzing large-scale data from a real-world OHC with various data analytics techniques, including text mining, survival analysis, and predictive modeling, our research explained as well as predicted users' continued participations in OHCs from the perspective of online social support.

Background and Research Goals

Social Support and OHCs

According to Shumaker and Brownell [25], social support refers to the "exchange of resources between at least two individuals perceived by the provider or the recipient to be intended to enhance the well-being of the recipient." Based on the nature of exchanged "resources," community psychology researchers have identified different types of social support [26,27]. In this research, we adopted the four types of social support proposed in [28,29]: informational support, emotional support, companionship, and instrumental support. Informational support is the transmission of information, suggestion, or guidance to the community users [30]. The content of such a post in an OHC is usually related to advice, referrals, education, and personal experience with the disease or health problem. Example topics include side effects of a drug, ways to deal with a symptom, experience with a physician, or medical insurance problems. Emotional support, as its name suggests, contains the expression of understanding, encouragement, empathy, affection, affirming, validation, sympathy, caring, and concern. Companionship, also known as network support, consists of chatting, humor, teasing, as well as discussions of offline activities and daily life that are not necessarily related to one's health problems. Thus, they are sometimes referred to as "off-topic" discussions. Examples include sharing jokes, birthday wishes, holiday plans, or Web-based scrabble games. Instrumental support, or tangible support, refers to offline support activities in the physical world, such as transporting others to hospitals, assistance in grocery shopping, and so forth. Empirical studies suggested that informational support, emotional support, and companionship are common in many OHCs, but instrumental support is rare, as such support is limited by geographical proximity [31,32]. Also, the further exchange and arrangement of instrumental support may often occur via private or offline communication channels (eg, setting a time for grocery shopping via cell phones). To simplify our automated social support classification, we did not consider instrumental support in this study.

The emergence of OHCs provides new opportunities to study social support at unprecedented scales and granularities. Traditional studies on offline support communities studies relied heavily on data collected through ethnographical observations, interviews, questionnaires, or surveys [14,33-36]. However, research using these data collection methods faces 3 challenges.

First, the scale of the data is limited because observations and interviews are labor intensive and time consuming. Second, results may be biased due to the realities of sampling community members. For example, members who are active in or satisfied with their communities may be more likely to respond to questionnaires or surveys. Third, survey and interview methods typically have coarse temporal granularity and rely on members' recall of past events and associated feelings. This makes it very difficult to accurately track community members' activities during an extended time period.

By contrast, OHCs not only enable but also record asynchronous and distributed social interactions among individuals, making the "big data" available for computational analysis. Such detailed data of users' online interactions (eg, the amount, content, and time of interactions) contain valuable information on users' behaviors. To study social support at such a large scale and fine granularity, we need to reveal the nature of social support embedded in users' contributions in an automated way. Hence, our first research goal was about mining large-scale text data contributed by OHC users to detect different types of social support activities.

Goal #1: Detect the seeking and provision of different types of social support from unstructured text of large-scale distributed interactions among OHC users.

Online Community Participations

According to Preece [37], an online community is a group of people who are connected through the Internet and interact over time around a shared purpose, interest, or need. The success of online communities depends largely on sustained participations and voluntary contributions from users [38]. Researchers have revealed factors related to user participation in online communities, such as open-source software development [39], Wikipedia [40], and Question & Answering communities [41,42].

Different from other types of online communities, seeking and obtaining various types of social support is a key reason people participate in an OHC [43]. On one hand, OHC users have a common identity as the patient of a disease, and information about the disease will be discussed and exchanged very often. On the other hand, the exchange of emotional support and participations in companionship, often in the form of seemingly off-topic discussions, can help OHC users get to know each other personally as they share things beyond health and the common disease. To understand which types of social support are more indicative of user engagement in the community, our second research goal was to run an explanatory model to connect different types of social support with user continued participation.

Goal #2: Develop an explanatory model to explore whether users' activities in seeking, providing, and receiving different types of social support are related to their continued participations in an OHC.

Churn Predictions

In addition to building an explanatory model to understand factors related to users' continued participations, another key to sustain an online community is to predict user churn, so that the community can intervene when a user is about to churn and try to retain her or him. Implications for churn prediction are not limited to online communities, but also to other online and offline businesses, such as telecommunication [44], retail [45], Internet access service [46], and online gaming [47]. These models have leveraged different types of data about customers and the market, including those related to money, contracts, demographics, usage, products, complaints, competitions, and social networks [48-50].

When it comes to online communities, traditional churn prediction faces challenges as well as opportunities. On one hand, many of the features commonly used for churn prediction in for-profit business are not available or make no sense. For instance, users' demographic data (eg, residential address, income, and ethnicity) is usually unavailable or inaccurate in online communities. Also, because many online communities are based on voluntary participations and do not charge any fee, monetary and contractual issues become largely irrelevant. On the other hand, online communities provide more detailed data about users' behaviors for predictive analytics [51]. While previous churn prediction studies have leveraged structured data of users' activities, few have examined the unstructured content of users' interactions or contributions. In contrast, in many online communities, including OHCs, large amount of such content is publicly available from the Web. Previous research on online social networks and social media has suggested that content analysis can be helpful in areas such as personalized recommendation [52], community discovery [53], and influential user identification [54]. We believe analyzing unstructured text posted by online community users from a social support perspective should contribute to accurate churn prediction in OHCs.

Moreover, many churn predictions for traditional business are limited to snapshot data—a model is learned from data for customers, who were active during a specific period (ie, the training period, usually a couple of months to half a year), based on which customers churned in the subsequent testing period (often a few months). For an online community, data for a user's complete "life span" in the community can be available for analysis. Such complete data can provide valuable information because those who churn after the first week may behave differently from those who churn after a month. Thus, our last research goal is about building a predictive model using data of users' social support activities.

Goal #3: Leverage data about users' Web-based social support activities over time to build a predictive model to forecast whether and when a user will churn from an OHC.

Methods

In this research, we used the data from a very popular peer-to-peer OHC (Breastcancer.org) among breast cancer survivors as a case study. We designed a Web crawler to collect

data from its online forum. Our dataset consisted of all the public posts and basic user profile information from October 2002 to August 2013.

Methods of Social Support Detection

As we mentioned earlier, informational support, emotional support, and companionship are the three major types of social supports in OHCs. Thus for each post from an OHC, we need to determine whether it was seeking informational support (SIS), providing informational support (PIS), seeking emotional

support (SES), providing emotional support (PES), or simply about companionship (COM). Note that we did not differentiate the seeking and provision of companionship because the nature of companionship was about participation and sharing. By getting involved in activities or discussions about companionship through posting, one was seeking and providing support at the same time. It was also possible that a post could belong to more than 1 of the aforementioned categories. [Table 1](#) lists example posts for each category and a post that belongs to 2 categories.

Table 1. Example posts for different types of social support.

Social support category	Examples
Companionship (COM)	(1) <i>Kelly Have a wonderful time in Florida, enjoy the sun and fun. Heather</i> (2) <i>I'm loving her new CD. Didn't recognize any of the songs at first, but there are a few now that I find myself singing the rest of the day.</i> (3) <i>This game has the poster making a new 2 word phrase starting with the second word of the last post Example: Post : Hand out Next poster: Outcast Next poster: Cast Iron Next poster: Iron Age Now let's begin the game~ Age Old</i>
Seeking informational support (SIS)	<i>Where do you buy digestive enzymes and what are they called?</i>
Seeking emotional support (SES)	<i>I feel like everyone else's lives are going forward, they have plans, hopes, aspirations because they feel. I am one of those not yet out of the woods. I was also someone who could never get cancer. I was a good person, exercised, ate well. Good people don't get sick. I have taken the step of antidepressants, they mitigate the damage, but do not block the pain or sadness I feel.</i>
Providing informational support (PIS)	<i>I had surgery Aug05 for bc recurrence. B4 surgery I had 33 IMRT rads, prior to that had 4A/C & 4 Taxol. I had bc in 2000 & had 37 rads in same general area. Now, my surgery won't heal. Wound doc says there is adema or something on my sternum (shown on recent MRI). My wound has been draining since it broke open in Sept.</i>
Providing emotional support (PES)	<i>Hope you feel better soon, we are here! Prayers Hugs come from Massachusetts APPLE♥ .</i>
Providing informational support (PIS) and providing emotional support (PES)	<i>I am also the daughter of a 35 yrs BC survivor. Mom is just now going through some more Cancer - alas - they found it in her lung, but it is totally unlikely to be a follow-up of her old BC. I am 45, and was 43 at DX time, my mom was diagnosed at 38... and I am a BRCA2 carrier. Tina, one day at a time. Maybe you'll get good news - it is so hard to wait!!! It is also important to remember that - whatever it is, it is highly treatable, and that YOU WILL SURVIVE too!!! and life goes on after. It will take some time, but it goes on... see my picture? even the hair is back!!! Hugs to all. I am happy you all found your way here, it is a great site for exchanging information, learning and finding support.</i>

Because it is practically impossible to label all 2.8 million posts manually, we used text classification algorithms to decide what kinds of social support each post contained. Text mining techniques have been adopted to analyze large-scale text data from online social networks, including texts from online health communities (similar findings by Ko D-G, Mai F, and Zhe S, unpublished data, 2015). To train a text classification algorithm, we leveraged human annotated data. We randomly selected 1333 posts out of our dataset. After being trained on the definitions and examples of the aforementioned 5 categories of social supports (SIS, PIS, SES, PES, and COM), 5 human annotators were asked to read each post and decide whether the post belongs to one or more categories of social support (See [Multimedia Appendix 1](#) for the training instruction). To control the quality of human annotations, we also added to the pool 10 posts that had been annotated by domain experts. For each post, we only accepted results from annotators whose performance on the 10 quality-control posts was among the top 3. The results from the other 2 annotators were discarded. Then, a majority vote among the top 3 annotators was used to determine whether

a post was related to a category of social support. [Multimedia Appendix 1](#) lists the outcome of the annotation.

Users in OHCs may have different writing styles or linguistic preferences to express themselves. To capture these characteristics, we examined each post and extracted various types of features for training the classifier: basic features, lexical features, sentiment features, and topic features. [Multimedia Appendix 2](#) includes more details about the feature engineering for social support classification.

Methods of Participation Analysis

After detecting the nature of social support in each post, we conducted survival analysis to study how different types of social support activities were related to users' participations. An individual may enter or exit a community not only based on his or her own expectations and behaviors, but also based on the community's reactions toward this individual [55]. Thus, in addition to users' own posting behaviors, we also examined whether the receiving or exposure to different types of social support would impact a user's participation.

Our survival analysis was based on the Cox proportional-hazards model [56], which assessed the importance of different independent variables on the “survival time” it takes for a specific event to occur (Multimedia Appendix 3 includes more details of the model setup). Specifically, for our analysis, an “event” referred to a user’s cessation of activities in the OHC (ie, churn from the OHC). A user’s survival time was measured from the difference between her last and first posts in the OHC. Similar to a previous study [24], we assumed that a user had churned from this OHC if she had no post during the last 12 weeks in our dataset. For those who were active in the OHC during the last 12 weeks, their survival time was right-censored because they were still participating in this OHC.

Table 2 summarizes independent variables in our model. They reflect users’ own posting behaviors in various social support categories, as well as the amount of social support they received in threaded discussions in direct or indirect ways. A user received support directly when she initiated a thread to seek support and got support from others’ replies to the thread. Meanwhile, social support could also be received indirectly when one replied to a thread started by another user because she might be exposed to support that other users provided to the original poster. In addition to these independent variables,

we also included 3 control variables to reflect users’ overall levels of activities.

The experiment included 19,165 users whose time spans of activities in the OHC exceeded 1 month. Values of control and independent variables were collected based on their behaviors in seeking, providing, and receiving social support in the first month of their participations. To reduce the impact of multi-collinearity, we calculated the correlation coefficients for every pair of variables. We then removed TotalPost and NumThread from the model, as both were strongly correlated with the other control variable InitPost (with correlation coefficients greater than .8). Thus, our model for survival analysis included 1 control variable and 10 independent variables.

Methods of Churn Prediction

If different types of social support activities are indeed related to users’ participations in OHCs, OHC managers can design more effective interventions to retain users. Such interventions can be more targeted when OHC managers know who are likely to leave and when. Therefore, this section proposes a model to predict whether and when a user will churn from an OHC and demonstrates the value of including social support activities over time in such predictions.

Table 2. Control variables and independent variables in the survival analysis.

Variables	Descriptions
TotalPost ^a	The total number of posts a user has published (excluded from the model due to strong correlation with <i>InitPost</i>)
InitPost ^a	The total number of threads a user initiated
NumThread ^a	The number of threads a user contributed to (excluded from the model due to strong correlation with <i>InitPost</i>)
PES ^b	The number of a user’s posts that provided emotional support
PIS ^c	The number of a user’s posts that provided informational support
SES ^d	The number of a user’s posts that sought emotional support
SIS ^e	The number of a user’s posts that sought informational support
COM ^f	The number of a user’s posts that were related to companionship
RIS _D	Direct informational support received—the number of informational support posts a user received after initiating a support-seeking thread.
RES _D	Direct emotional support received—the number of emotional support posts a user received after initiating a support-seeking thread.
RIS _I ^g	Indirect informational support received—the number of informational support posts a user was exposed to in threads that she or he did not initiate but contributed to.
RES _I ^g	Indirect emotional support received—the number of emotional support posts a user was exposed to in threads that she or he did not initiate but contributed to.
RCOM ^g	Companionship received—the number of companionship posts a user was exposed to in threads that she or he did not initiate but contributed to.

^adenotes the three control variables.

^bPES: providing emotional support.

^cPIS: providing informational support.

^dSES: seeking emotional support.

^eSIS: seeking informational support.

^fCOM: companionship.

^gFor *RIS_I*, *RES_I*, and *RCOM*, we assumed that a user read others’ replies that were posted within 7 days before the user’s replies in the same thread.

Basic features for our predictive model are derived from the 13 independent variables we used for survival analysis (Table 2). Because these features aggregated users' activities during the training period, we also measured how users' values on the 13 features varied over time using four types of temporal features. Specifically, for each user, we divided her activities measured by each of the 13 basic features into weeks and used 4 additional metrics to capture how the value of each feature changes over the weeks, including overall slope, Shannon entropy, stability, and temporal variations (TV) as proposed in [57]. In addition to cumulative values for each basic feature during the training period, we also conjectured that a user's intention to churn might be better captured during the last week of her online activities. Thus, we also included values for basic features during the last week of the training period if the training period was longer than 1 week. Each basic feature for the last week also had 4 corresponding features to reflect its temporal patterns (ie, slope, Shannon entropy, stability, and TV), although the unit of time was day instead of week. We also added into the feature set the time difference between a user's registration time and the time of her first post because it might reveal what brought the user to the OHC for the first time. A user who is eager to find some information might have a low gap between the registration time and the time of first posting. More details of features are presented in Multimedia Appendix 4.

In terms of modeling the churn prediction problem, a user was said to churn in her k -th week if her last online activity occurred during her k -th week in the OHC. Similar to our hazard model, users whose last online activities occurred during the last 12 weeks in our dataset were not considered as churned. To predict whether a user would churn in the k -th week of her online activities, we focused on all users who were still active before the k -th week and extracted data based on their k weeks of activities. For example, the dataset for predicting user churn during the third week contains users who were still active in the OHC before their third week of online activities. Data of their behaviors during their first 2 weeks were collected for training. Users who churned in their third week and never came back were labeled as "positive" instances in the dataset.

Previous studies indicate that different predictive models for each time period may not be an efficient solution. If the OHC wanted to know who would churn in their second and third weeks, 2 models were needed. Inspired by [58], we tried to consolidate all predictive models for churn in different weeks into 1 unified model by leveraging a user's social support activities across her complete "online life span." Specifically, we added 1 new feature to the predictive model—time stamp t . An instance in the dataset would reflect a user's historical activities until t . As the unit of t was the same for all users (a

week in our experiment), 1 user could correspond to multiple instances in the dataset. For example, a user who churned in her third week of activities had 3 instances in the dataset—1 instance for her activities and features until the end of her first, second, and third week, respectively. The first 2 instances were labeled as "negative" because the user was still active during these 2 weeks, while the third instance was labeled as a "positive" instance because the user churned in her third week. In other words, the unified model tried to capture the complete life span of a user in the OHC.

To train the unified model, 24,000 users were randomly selected from 47,581 users in the OHC to be included in the training dataset, while others were placed in the hold-out testing dataset. It is worth noting that the unified model with time stamps as a feature greatly increased the amount of training data because a loyal user who had been active for a long time would have many instances in the dataset. However, 24,000 users in the training dataset resulted in 132,341 instances in total. We built the training dataset and trained the model on a high-performance computing cluster. We also made sure that instances for the same user must belong to the same fold in cross-validation.

Results

Results of Social Support Detection

In our dataset from Breastcancer.org, there were more than 2.8 million posts contributed by nearly 50,000 users, including 107,549 initial posts. Figures 1 and 2 show the distribution of the number of published posts and the time span of users' posting activities in this OHC. The 2 plots indicate that users' online behaviors featured highly skewed distributions that are similar to power-law distributions. In other words, many users were not very active in posting, while some users were very productive and stayed for a long time.

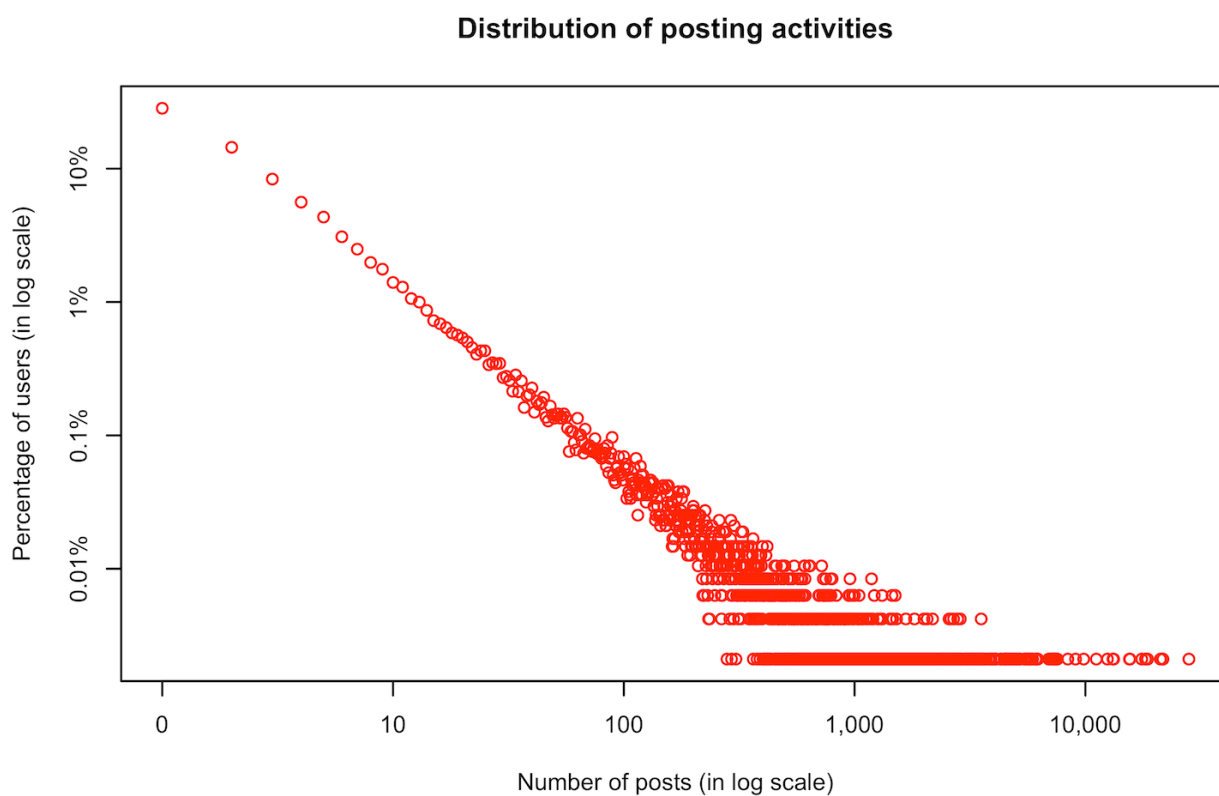
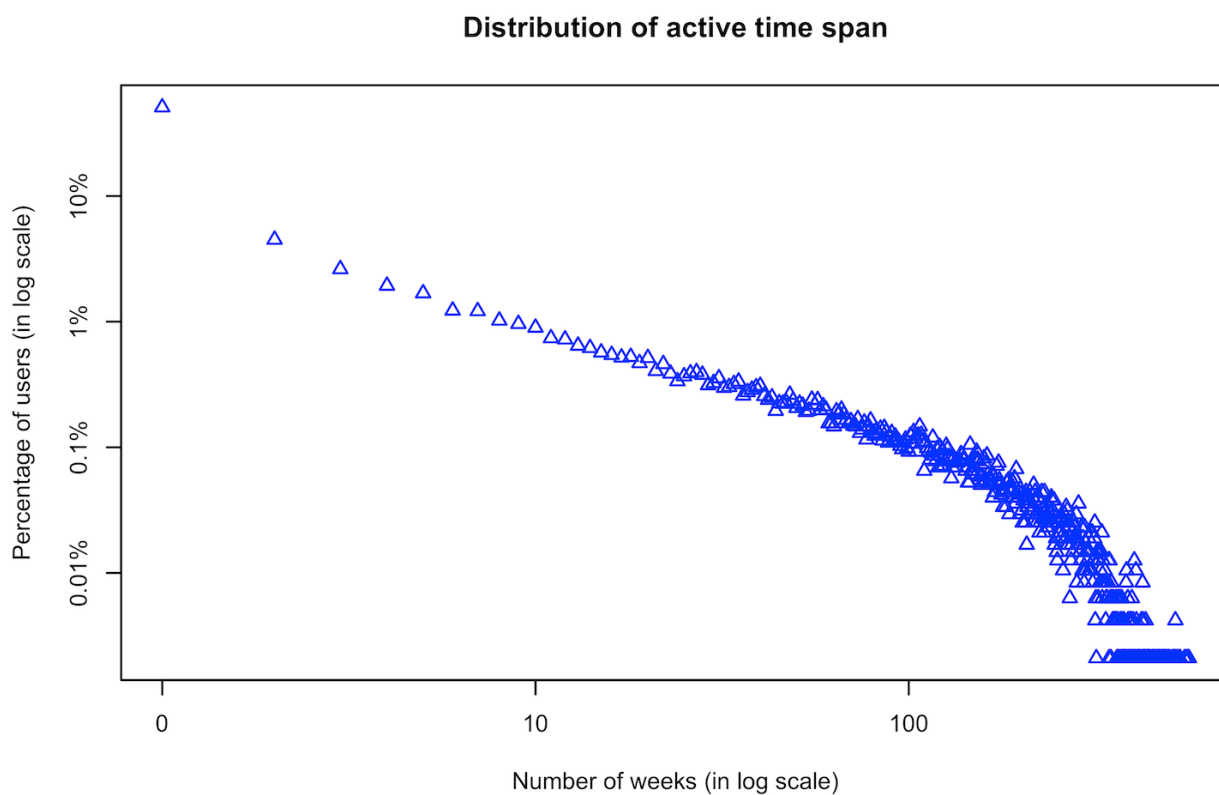
Because we considered 5 categories of social supports and a post may belong to more than 1 category, we built a classifier for each category. For the classification of each category of social support, we applied various classification algorithms on annotated posts and picked the best performing algorithm using 10-fold cross-validation. Because posts seeking emotional support (SES) accounts for only a small proportion among annotated posts (22 out of 1333), we oversampled posts seeking emotional support when building the SES classifier. Among all the classifiers we tried, AdaBoost, with Naïve Bayesian as the weak learner, was chosen to classify COM, PES, PIS, and SIS, while logistic regression was the best choice for SES (Table 3). Overall, our classifiers achieved decent performance with an accuracy rate of more than 0.8 in all 5 classification tasks.

Table 3. Performance of classification algorithms for 5 categories of social support.

Social support	Results	Naïve Bayesian	Logistic regression	Support Vector Machine (polynomial kernel)	Random forest	Decision tree	AdaBoost
COM ^a	Accuracy	.696	.787	.783	.771	.767	.804 ^f
	AUC	.839	.817	.768	.848	.75	.852 ^f
PES ^b	Accuracy	.713	.830	.840 ^f	.830	.81	.817
	AUC	.823	.787	.681	.825 ^f	.687	.817
PIS ^c	Accuracy	.753	.813	.823 ^f	.767	.779	.801
	AUC	.824	.83	.783	.837	.717	.859 ^f
SES ^d	Accuracy	.893	.901	.970 ^f	.967	.963	.963
	AUC	.749	.867 ^f	.656	.851	.671	.668
SIS ^e	Accuracy	.851	.880	.943 ^f	.931	.937	.914
	AUC	.893 ^f	.803	.745	.86	.766	.869

^aCOM: companionship.^bPES: providing emotional support.^cPIS: providing informational support.^dSES: seeking emotional support.^eSIS: seeking informational support.^fdenotes the best performer for each row.**Table 4.** Total numbers of posts in each category of social support.

Social support category	Total number of posts
Companionship (COM)	932,538
Seeking informational support (SIS)	284,027
Seeking emotional support (SES)	227,188
Providing informational support (PIS)	1,034,682
Providing emotional support (PES)	497,096

Figure 1. Log-log plot of users' posting activities in the online health community (OHC).**Figure 2.** Log-log plot for the time span of users' online posting activities.

Results of Participation Analysis

Table 5 shows the results of the participation analysis based on Cox Proportional Hazard models. Variables with hazard ratios lower than 1 contributed positively to the “survival” (ie, continued participation) of users, whereas those with hazard ratio higher than 1 were considered “hazardous” to keep users in this OHC. Three independent variables (PIS, SES, and COM) had hazard ratios that were lower than 1, meaning that users who provided more informational support, sought more emotional support or posted more companionship had longer time spans of activities in the OHC. More specifically, a hazard ratio of 0.919 for companionship meant that a user’s “survival” rate after 1 month was 8.1% higher (100%–91.9%) if her number of companionship posts was one standard deviation higher than the average. In contrast, those who sought or received more informational support (SIS, RIS_D, and RIS_I) often left the OHC earlier. Other variables were not significant in the experiment (eg, PES).

Results of Churn Prediction

We measured the performance of predictive classifiers using standard metrics for classification, including precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC). After comparing the performance of different classification algorithms (Naïve Bayes, logistic regression, and SVM with polynomial kernel) with 10-fold cross-validation on the training set, logistic regression emerged as the best performer for the unified model. As shown in Table 6, the model offers very good performance in predicting churn during the first week. Although recall of the positive class (ie, leaving the OHC) decreased for prediction in later weeks, the precision was still higher than 0.8, and the overall performance measured by AUC was still more than 0.9.

We also plotted 2 hazard curves to visualize the model’s performance (Figure 3): one based on empirical data and the other based on predictions from the unified model. The horizontal axis represented weeks, and the vertical axis referred to the probability of users’ churn in specific weeks. The 2 curves were very close to each other, indicating good predictive performance from our model at the community level.

Table 5. Results from the survival analysis experiment.

Variables	Hazard ratio	P Value
InitPost (control)	.995	.75
PES ^a	1.000	.99
PIS ^b	.948***	.001
SES ^c	.972*	.01
SIS ^d	1.050***	.000
COM ^e	.919***	.000
RIS _D	1.047*	.02
RES _D	.997	.79
RIS _I	1.053*	.02
RES _I	.964	.11
RCOM	.983	.41

^aPES: providing emotional support.

^bPIS: providing informational support.

^cSES: seeking emotional support.

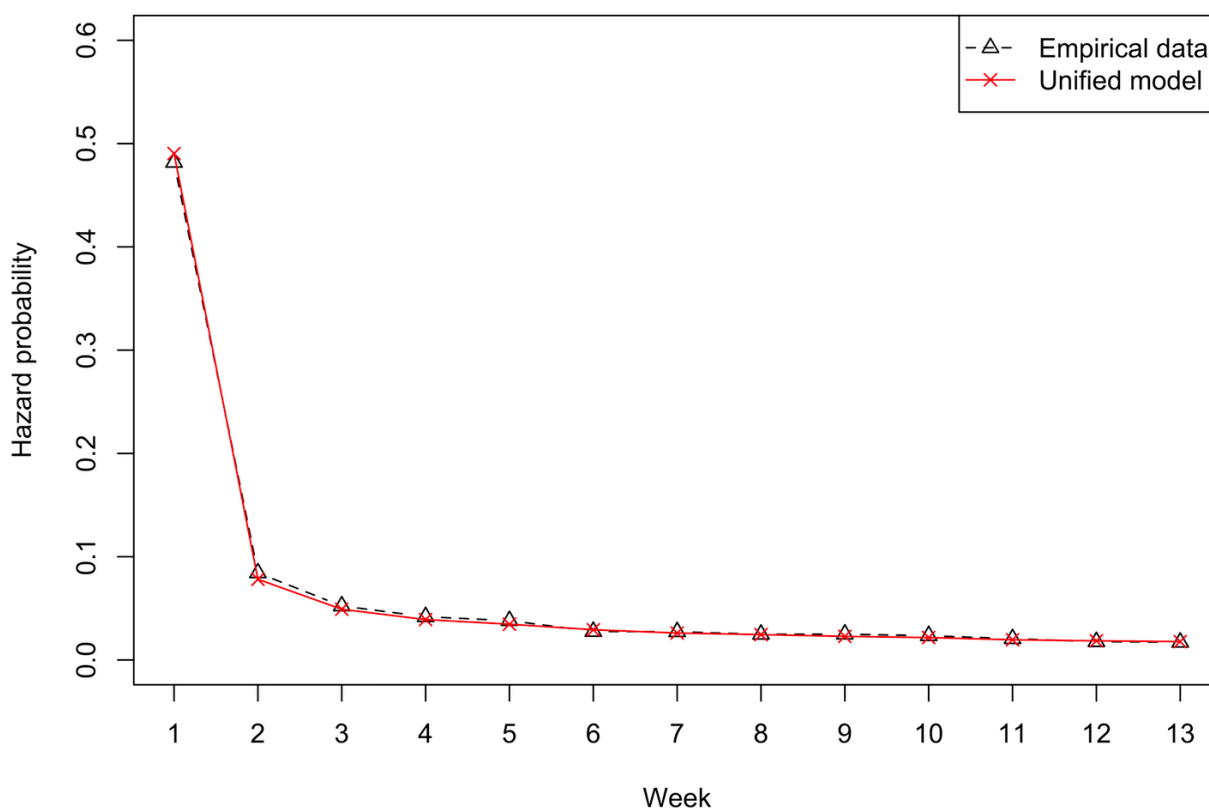
^dSIS: seeking informational support.

^eCOM: companionship.

Table 6. Performance of the unified model on hold-out testing sets in different weeks (precision and recall are for the positive class).

Measures	Churn in the 1st week	Churn in the 3rd week	Churn in the 5th week	Churn in the 13th week
Precision	.950	.872	.880	.838
Recall	.937	.534	.511	.504
F1 score	.943	.662	.647	.629
AUC ^a	.972	.901	.909	.929

^aAUC: area under the receiver operating characteristic (ROC) curve.

Figure 3. Empirical and predicted hazard curves for user participations.

Discussion

Principal Findings

The results of survival analysis showed that seeking or providing various types of social support was related to users' participations in different ways. First, informational support is the most popular social support being sought and provided. This was expected for communities based on common social identities [59] because the large amount of information about a disease and the common identity as patients of the disease are probably why many users come to the OHC in the first place. While providing more informational support is positively correlated with longer participations, seeking and receiving informational support are negatively associated with participations. In other words, those who focus on seeking information may not stay in the long run, even after they receive informational support.

Second, companionship had the lowest hazard ratio. Recall that companionship includes discussions of offline events, sharing daily life stories, birthday wishes, and playing online games. This is a very interesting finding—even though this is an OHC about cancer, discussions of non-cancer-related issues are the key to keeping users engaged in the community. This highlights the importance of building personal bond [59] through off-topic discussions in the form of sharing personal stories about life or having fun together, which can strengthen the connections among users more than informational support. The role of companionship has significant implications for the management of an OHC. Although some OHCs may discourage off-topic

discussions in order to achieve a “cleaner” environment with only relevant content, these discussions turn out to be a good way to bond users and keep them engaged, and OHC managers may want to encourage, or even initiate, more of these activities.

Third, although we expected emotional support to be positively related to user participation as suggested by [24], the results are mixed based on whether emotional support was being sought, provided, or received. The hazard ratio of SES was below 1 in the experiment, which contradicts the effect of SIS and suggests that SES can be a sign of longer participations, especially for those who have been with the OHC for a while. However, providing and receiving emotional support are not significant factors. We conjectured that a fair amount of emotional support in the OHC could be generic and a mere formality (eg, “I will pray for you,” “Love you and Hug”). Such emotional support can still be valuable for those who seek support, but activities in providing and receiving such support are not related to users' continued participations.

Our survival analysis has shown the effects of social support activities on users' engagement. How much do such social support activities contribute to the performance of the churn predictive model? To illustrate which features are more important for the unified model's predictive power, we ranked the 145 features in the model using information gain [60]. Among the top 20 features (Table 7), 11 features among the top 20 were made possible only after our classification of different types of social support. Also, 18 of them were features that reflected the temporal dynamics in users' social support activities, especially the stability during the last week of the training period. Overall, this shows that users' activities in

seeking, providing, and receiving different types of social support, as well as their temporal dynamics in these activities, can greatly enhance churn predictions in OHCs.

In this research, we mined large-scale data to better understand and predict users' continued participations in OHCs. We first detected the seeking and provision of different types of social support from OHC users' posts using text mining techniques. Then, survival analysis revealed that companionship is a significant and positive predictor of users' continued participations. Not limiting the potential of the large-scale data to descriptive analytics, we also developed a churn prediction model with high accuracies. Our work serves as an example that highlights the power of data analytics in exploring complex human behaviors.

From a managerial perspective, the outcome of our study can provide OHC managers with suggestions on how to sustain users' participations and decision support to retain users through interventions (eg, post recommendations and email reminders). A sustainable and successful OHC will eventually benefit its users. From a methodological perspective, this study was the

first to use text mining to differentiate the seeking and providing of various types of social support from large-scale OHC data, and demonstrated how such detection of social support activities could help to understand and predict users' engagement in OHCs.

This study has practical implications. Traditionally, an OHC will send reminder emails to a user who has been inactive for a while, hoping to raise the user's interests in coming back. With the help of our churn prediction model, an OHC can find at an early stage whether a user is about to leave. Then, it can intervene proactively and try to retain the user via email reminders. More importantly, instead of including a generic reminder or some random recent posts from the community, such emails can be designed based on the results of our survival analysis. For example, because companionship is a key predictor of users' continued participations, including some of these companionship posts (eg, birthday wishes, holiday plans, and online scrabble games) in reminder emails may be more effective to keep users engaged than having random posts or just informational posts.

Table 7. Top 20 features by information gain for the full unified model.

Rank	Feature
1	Stability of the total number of threads a user initiated during the last week of the training period
2	Stability of the number of threads a user participated during the last week of the training period
3	Stability of the number of SIS ^a a user posted during the last week of the training period
4	Stability of the number of SES ^b a user posted during the last week of the training period
5	Stability of the total number of posts from a user during the last week of the training period
6	Stability of the number of PIS ^c a user posted during the last week of the training period
7	Stability of the number of PES ^d posts a user received directly during the last week of the training period
8	Stability of the number of PES a user posted during the last week of the training period
9	Stability of the number of COM ^e a user posted during the last week of the training period
10	Stability of the number of PIS posts a user received directly during the last week of the training period
11	Stability of the number of COM a user was exposed to during the last week of the training period
12	Stability of the number of PES posts a user received indirectly during the last week of the training period
13	Stability of the number of PIS posts a user received indirectly during the last week of the training period
14	Total number of posts from a user during the last week of the training period
15	The number of threads a user participated in during the last week of the training period
16	Stability of the number of threads a user participated in across weeks
17	Stability of the total number of posts from a user across weeks
18	Entropy of the total number of posts from a user during the last week of the training period
19	Stability of the number of PIS posts a user received indirectly across weeks
20	Stability of the total number of threads a user initiated across weeks

^aSIS: seeking informational support.

^bSES: seeking emotional support.

^cPIS: providing informational support.

^dPES: providing emotional support.

^eCOM: companionship.

Limitations

This study also had limitations. First, for the 3 independent variables for indirect support received (RIS_i , RES_i , and $RCOM$), we assumed that a user received indirect support when she replied to a thread initiated by another user and read other users' replies to the thread. This approach of capturing indirect support received could be inaccurate: on one hand, we might underestimate the amount of support because we limited our calculation to threads a user replied to, while a user could get indirect support by reading a thread without posting a reply. On the other hand, our approach might also overestimate such indirect support because when posting to a long thread, a user might not have time to read all the previous replies. This limitation can be addressed by analyzing users' click streams, but such data were not available for this study and can be difficult to obtain for many studies of OHCs. The lack of clickstream data also prevented us from analyzing lurking behaviors, which might also provide social support to lurkers [35]. Having users' clickstream data will also help us better define each user's temporal span of online activities. Second, our survival analysis only reveals the correlation between users' social support activities and their participations, without showing any causality. Although randomized experiments are better choices to infer causality, there might be ethical concerns to run such experiments in OHCs (eg, keeping certain users away from social support). Alternative approaches are needed to identify

causal relationships. Last but not least, our study was based on data from one OHC for breast cancer. In OHCs for other diseases, especially acute diseases (eg, flu), social support activities and users' engagement patterns may differ. Although the specific results we found for this breast cancer OHC may not be applicable for all OHCs, the framework of methods we used to classify social support and analyze users' continued participation based on social support activities can be applied to other OHCs.

Future Work

There are several interesting directions for future research. Detecting users' health status from their posts will be an interesting endeavor, as it not only can help understand why a user leaves an OHC, but also can potentially improve the recommendation and retrieval of Web-based information. We are also interested in improving the unified predictive model, which is easier for OHCs to use. One possible way is to rebalance instances in the unified model's dataset because the current dataset features way more negative instances than positive. It would also be interesting to explore whether users' engagement behaviors change over time, especially when accessing the Web using mobile devices is becoming more popular in recent years. We would also like to collaborate with OHC operators to evaluate the effectiveness of interventions aiming at keeping users engaged.

Acknowledgments

Kang Zhao's work has been partially supported by the National Natural Science Foundation of China (Award #: 71572013).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Post tagging guidelines and outcomes.

[PDF File (Adobe PDF File), 62KB - [jmir_v19i4e130_app1.pdf](#)]

Multimedia Appendix 2

Features engineering for social support classification.

[PDF File (Adobe PDF File), 36KB - [jmir_v19i4e130_app2.pdf](#)]

Multimedia Appendix 3

Hazard model.

[PDF File (Adobe PDF File), 39KB - [jmir_v19i4e130_app3.pdf](#)]

Multimedia Appendix 4

Features engineering for the churn predictive model.

[PDF File (Adobe PDF File), 45KB - [jmir_v19i4e130_app4.pdf](#)]

References

1. Fox S. Pew Research Center. 2011. The social life of health information, 2011 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2017-03-04] [WebCite Cache ID 6oidDcGPg]

2. White M, Dorman SM. Receiving social support online: implications for health education. *Health Educ Res* 2001 Dec;16(6):693-707. [Medline: [11780708](#)]
3. Caplan SE, Turner JS. Bringing theory to research on computer-mediated comforting communication. *Comput Human Behav* 2007 Mar;23(2):985-998 [FREE Full text] [doi: [10.1016/j.chb.2005.08.003](#)]
4. Wright K, Bell S, Wright KB, Bell SB. Health-related support groups on the internet: linking empirical findings to social support and computer-mediated communication theory. *J Health Psychol* 2003 Jan;8(1):39-54. [doi: [10.1177/1359105303008001429](#)] [Medline: [22113899](#)]
5. Wright K. Social support within an on-line cancer community: an assessment of emotional support, perceptions of advantages and disadvantages, and motives for using the community from a communication perspective. *J Appl Commun Res* 2002 Jan;30(3):195-209. [doi: [10.1080/00909880216586](#)]
6. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. *J Med Internet Res* 2009 Aug;11(4):e48-e47 [FREE Full text] [doi: [10.2196/jmir.1249](#)]
7. Barak A, Boniel-Nissim M, Suler J. Fostering empowerment in online support groups. *Comput Human Behav* 2008 Sep;24(5):1867-1883. [doi: [10.1016/j.chb.2008.02.004](#)]
8. Yang C, Yang H, Jiang L, Zhang M. Social media mining for drug safety signal detection. In: Proceedings of the 2012 international workshop on Smart health and wellbeing. 2012 Oct 29 Presented at: CIKM Conference on Information and Knowledge Management; October 29, 2012; Maui, HI p. 33-40 URL: <http://dl.acm.org/citation.cfm?doid=2389707.2389714> [doi: [10.1145/2389707.2389714](#)]
9. Yang C. Patient centered healthcare informatics. *IEEE Intelligent Informatics Bulletin* 2014 Dec;15(1):1-5 [FREE Full text]
10. Kraut R, Resnick P, Kiesler S, Burke M, Chen Y, Kittur N, et al. Building Successful Online Communities: Evidence-Based Social Design. Cambridge, MA: The MIT Press; Mar 2012.
11. Young C. Community management that works: how to build and sustain a thriving online health community. *J Med Internet Res* 2013 Jun 11;15(6):e119 [FREE Full text] [doi: [10.2196/jmir.2501](#)] [Medline: [23759312](#)]
12. Iriberri A, Leroy G. A life-cycle perspective on online community success. *ACM Comput Surv* 2009 Feb 01;41(2):1-29. [doi: [10.1145/1459352.1459356](#)]
13. Bouma G, Admiraal JM, de Vries E, Schröder CP, Walenkamp AM, Reyners AK. Internet-based support programs to alleviate psychosocial and physical symptoms in cancer patients: a literature analysis. *Crit Rev Oncol Hematol* 2015 Jul;95(1):26-37. [doi: [10.1016/j.critrevonc.2015.01.011](#)] [Medline: [25701515](#)]
14. Campbell HS, Phaneuf MR, Deane K. Cancer peer support programs-do they work? *Patient Educ Couns* 2004 Oct;55(1):3-15. [doi: [10.1016/j.pec.2003.10.001](#)] [Medline: [15476984](#)]
15. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Primary care health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ* 2004 May 13;328:1166. [doi: [10.1136/bmj.328.7449.1166](#)]
16. Hoey LM, Ieropoli SC, White VM, Jefford M. Systematic review of peer-support programs for people with cancer. *Patient Educ Couns* 2008 Mar;70(3):315-337. [doi: [10.1016/j.pec.2007.11.016](#)] [Medline: [18191527](#)]
17. Idriss SZ, Kvedar JC, Watson AJ. The role of online support communities: benefits of expanded social networks to patients with psoriasis. *Arch Dermatol* 2009 Jan;145(1):46-51. [doi: [10.1001/archdermatol.2008.529](#)] [Medline: [19153342](#)]
18. Zhang S, Bantum E, Owen J, Elhadad N. Does sustained participation in an online health community affect sentiment? *AMIA Annu Symp Proc* 2014;2014:1970-1979 [FREE Full text] [Medline: [25954470](#)]
19. Burrows R, Nettleton S, Pleace N, Loader B, Muncer S. VIRTUAL COMMUNITY CARE? SOCIAL POLICY AND THE EMERGENCE OF COMPUTER MEDIATED SOCIAL SUPPORT. *Information, Communication & Society* 2000 Jan;3(1):95-121. [doi: [10.1080/136911800359446](#)]
20. Dunkel-Schetter C. Social support and cancer: findings based on patient interviews and their implications. *J Soc Issues* 1984;40(4):77-98. [doi: [10.1111/j.1540-4560.1984.tb01108.x](#)]
21. Qiu B, Zhao K, Mitra P, Wu D, Caragea C, Yen J. Get online support, feel better -- sentiment analysis and dynamics in an online cancer survivor community. 2011 Oct Presented at: Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing; 2011; Boston, MA p. 274-281. [doi: [10.1109/PASSAT/SocialCom.2011.127](#)]
22. Maloney-Krichmar D, Preece J. A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Trans Comput-Hum* 2005 Jun 01;12(2):201-232. [doi: [10.1145/1067860.1067864](#)]
23. Rodgers S, Chen Q. Internet community group participation: Psychosocial benefits for women with breast cancer. *J Comput Mediat Commun* 2005;10(4):00. [doi: [10.1111/j.1083-6101.2005.tb00268.x](#)]
24. Wang Y, Kraut R, Levine J. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. USA: ACM; 2012 Feb Presented at: ACM 2012 conference on Computer Supported Cooperative Work; February 11-15, 2012; Seattle, Washington, DC p. 833-842. [doi: [10.1145/2145204.2145329](#)]
25. Shumaker S, Brownell A. Toward a theory of social support: closing conceptual gaps. *J Soc Issues* 1984;40(4):11-36. [doi: [10.1111/j.1540-4560.1984.tb01105.x](#)]
26. House J. Work Stress and Social Support. Reading, MA: Addison-Wesley Educational Publishers Inc; 1981.

27. Barrera M, Ainlay S. The structure of social support: a conceptual and empirical analysis. *J Community Psychol* 1983 Apr;11(2):133-143. [Medline: [10299305](#)]
28. Bambina A. *Online Social Support: The Interplay of Social Networks and Computer-Mediated Communication*. Youngstown, NY: Cambria Press; 2007.
29. Keating DM. Spirituality and support: a descriptive analysis of online social support for depression. *J Relig Health* 2013 Sep;52(3):1014-1028. [doi: [10.1007/s10943-012-9577-x](#)] [Medline: [22322336](#)]
30. Krause N. Social support, stress, and well-being among older adults. *J Gerontol* 1986 Jul;41(4):512-519. [Medline: [3722737](#)]
31. Coulson N, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. *Patient Educ Couns* 2007 Oct;68(2):173-178. [doi: [10.1016/j.pec.2007.06.002](#)] [Medline: [17629440](#)]
32. Zhang M, Yang CC, Gong X. Social support and exchange patterns in an online smoking cessation intervention program. In: *Proceedings of the 2013 IEEE International Conference on Healthcare Informatics*. 2013 Sep Presented at: 2013 IEEE International Conference on Healthcare Informatics; September 9, 2013; Philadelphia, PA p. 219-228. [doi: [10.1109/ICHI.2013.37](#)]
33. Lieberman M. The role of insightful disclosure in outcomes for women in peer-directed breast cancer groups: a replication study. *Psychooncology* 2007 Oct;16(10):961-964. [doi: [10.1002/pon.1181](#)] [Medline: [17328007](#)]
34. Gorlick A, Bantum E, Owen J. Internet-based interventions for cancer-related distress: exploring the experiences of those whose needs are not met. *Psychooncology* 2014 Apr;23(4):452-458. [doi: [10.1002/pon.3443](#)]
35. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. *J Med Internet Res* 2011 Dec 29;13(4):e122 [FREE Full text] [doi: [10.2196/jmir.1696](#)] [Medline: [22204869](#)]
36. Hambly K. Activity profile of members of an online health community after articular cartilage repair of the knee. *Sports Health* 2011 May;3(3):275-282 [FREE Full text] [doi: [10.1177/1941738111402151](#)] [Medline: [23016018](#)]
37. Preece J. *Online communities: designing usability, supporting sociability*. Chichester: John Wiley; 2000.
38. Burke M, Marlow C, Lento T. Feed me: motivating newcomer contribution in social network sites. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Proc 27th Int Conf Hum Factors Comput Syst CHI09 1518847: ACM; 2009 Apr Presented at: Conference on Human Factors in Computing Systems; April 04-09, 2009; Boston, MA p. 945-954. [doi: [10.1145/1518701.1518847](#)]
39. Lakhani KR, Wolf R. Why hackers do what they do: understanding motivation and effort in free/open source software projects. In: Feller J, Fitzgerald B, Hissam S, Lakhani KR, editors. *Perspectives on free and open source software*. Cambridge, MA: MIT Press; 2005.
40. Nov O. What motivates Wikipedians? *Commun ACM* 2007 Nov 01;50(11):60-64. [doi: [10.1145/1297797.1297798](#)]
41. Khansa L, Ma X, Liginlal D, Kim S. Understanding members' active participation in online question-and-answer communities: a theory and empirical analysis. *J Manage Inform Syst* 2015 Aug 28;32(2):162-203. [doi: [10.1080/07421222.2015.1063293](#)]
42. Yang J, Wei X, Ackerman M, Adamic L. Activity lifespan: an analysis of user survival patterns in online knowledge sharing communities. In: *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*. 2010 Presented at: 4th International AAAI Conference on Weblogs and Social Media; May 23-26, 2010; Washington, DC p. 186-193.
43. Kim E, Han JY, Moon TJ, Shaw B, Shah DV, McTavish FM. The process and effect of supportive message expression and reception in online breast cancer support groups. *Psychooncology* 2012;21(5):531-540. [doi: [10.1002/pon.1942](#)]
44. Wei C, Chiu I. Turning telecommunications call details to churn prediction: a data mining approach. *Expert Syst Appl* 2002 Aug;23(2):103-112. [doi: [10.1016/S0957-4174\(02\)00030-1](#)]
45. Buckinx W, Van den Poel D. Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *Eur J Oper Res* 2005 Jul;164(1):252-268. [doi: [10.1016/j.ejor.2003.12.010](#)]
46. Huang B, Kechadi M, Buckley B. Customer churn prediction for broadband internet services. In: *Proceedings of the 11th International Conference on Data Warehousing and Knowledge Discovery*. Heidelberg: Springer-Verlag Berlin; 2009 Sep Presented at: 11th International Conference on Data Warehousing and Knowledge Discovery; August 31-September 02, 2009; Linz, Austria p. 229-243 URL: http://link.springer.com/chapter/10.1007/978-3-642-03730-6_19 [doi: [10.1007/978-3-642-03730-6_19](#)]
47. Kawale J, Pal A, Srivastava J. Churn prediction in MMORPGs: a social influence based approach. In: *Proceedings of the 2009 International Conference on Computational Science and Engineering*. 2009 Aug Presented at: 2009 International Conference on Computational Science and Engineering; August 29-31, 2009; Vancouver, BC, Canada p. 423-428. [doi: [10.1109/CSE.2009.80](#)]
48. Backiel A, Baesens B, Claeskens G. Mining telecommunication networks to enhance customer lifetime predictions. In: *Artificial Intelligence and Soft Computing*. Switzerland: Springer International Publishing; 2014 Presented at: 13th International Conference ICAISC 2014; June 1-5, 2014; Zakopane, Poland p. 15-26 URL: http://link.springer.com/chapter/10.1007/978-3-319-07176-3_2 [doi: [10.1007/978-3-319-07176-3_2](#)]
49. Berson A, Smith S, Thearling K. *Building Data Mining Applications for CRM*. London: McGraw-Hill; 1999.

50. Zhang X, Zhu J, Xu S, Wan Y. Predicting customer churn through interpersonal influence. *Knowl-Based Syst* 2012 Apr 01;28(17):97-104. [doi: [10.1016/j.knosys.2011.12.005](https://doi.org/10.1016/j.knosys.2011.12.005)]
51. Shmueli G, Koppius O. Predictive analytics in information systems research. *MIS Quarterly* 2011 Sep;35(3):553-572.
52. Barbieri N, Bonchi F, Manco G. Who to follow and why: link prediction with explanations. In: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. USA: ACM; 2014 Presented at: 20th ACM SIGKDD international conference on Knowledge discovery and data mining; August 24-27, 2014; New York, NY p. 1266-1275. [doi: [10.1145/2623330.2623733](https://doi.org/10.1145/2623330.2623733)]
53. Sachan M, Contractor D, Faruque T, Subramaniam L. Using content and interactions for discovering communities in social networks. In: *Proceedings of the 21st international conference on World Wide Web*. USA: ACM; 2012 Presented at: 21st international conference on World Wide Web; April 16-20, 2012; Lyon, France p. 331-340. [doi: [10.1145/2187836.2187882](https://doi.org/10.1145/2187836.2187882)]
54. Zhao K, Yen J, Greer G, Qiu B, Mitra P, Portier K. Finding influential users of online health communities: a new metric based on sentiment influence. *J Am Med Inform Assoc* 2014 Oct;21(e2):e212-e218 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2013-002282](https://doi.org/10.1136/amiajnl-2013-002282)] [Medline: [24449805](https://pubmed.ncbi.nlm.nih.gov/24449805/)]
55. Levine J, Moreland R. Group socialization: theory and research. *Eur Rev Soc Psychol* 1994 Jan;5(1):305-336. [doi: [10.1080/14792779543000093](https://doi.org/10.1080/14792779543000093)]
56. Cox D. Regression Models and Life-Tables. *J R Stat Soc Series B (Methodological)* 1972;34(2):187-220 [[FREE Full text](#)]
57. Zhao K, Kumar A. Who blogs what: understanding the publishing behavior of bloggers. *World Wide Web* 2012 May 17;16(5-6):621-644. [doi: [10.1007/s11280-012-0167-3](https://doi.org/10.1007/s11280-012-0167-3)]
58. Street N, Mangasarian O, Wolberg W. An inductive learning approach to prognostic prediction. In: *Proceedings of the Twelfth International Conference on Machine Learning*. 1995 Presented at: 12th International Conference on Machine Learning; 1995; San Francisco, CA p. 522-530.
59. Ren Y, Kraut R, Kiesler S. Applying common identity and bond theory to design of online communities. *Organ Stud* 2007 Mar;28(3):377-408. [doi: [10.1177/0170840607076007](https://doi.org/10.1177/0170840607076007)]
60. Quinlan J. Induction of decision trees. *Mach Learn* 1986;1(1):81-106. [doi: [10.1023/A:1022643204877](https://doi.org/10.1023/A:1022643204877)]

Abbreviations

AUC: area under the receiver operating characteristic curve

COM: companionship

OHC: online health community

PES: providing emotional support

PIS: providing informational support

SES: seeking emotional support

SIS: seeking informational support

SVM: support vector machine

TV: temporal variations

Edited by G Eysenbach; submitted 17.10.16; peer-reviewed by K Eddens, J Alpert; comments to author 09.11.16; revised version received 28.12.16; accepted 11.02.17; published 24.04.17.

Please cite as:

Wang X, Zhao K, Street N

Analyzing and Predicting User Participations in Online Health Communities: A Social Support Perspective

J Med Internet Res 2017;19(4):e130

URL: <http://www.jmir.org/2017/4/e130/>

doi: [10.2196/jmir.6834](https://doi.org/10.2196/jmir.6834)

PMID: [28438725](https://pubmed.ncbi.nlm.nih.gov/28438725/)

©Xi Wang, Kang Zhao, Nick Street. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 24.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Use of Machine Learning Classifiers and Sensor Data to Detect Neurological Deficit in Stroke Patients

Eunjeong Park¹, PhD; Hyuk-Jae Chang², MD, PhD; Hyo Suk Nam³, MD, PhD

¹Cardiovascular Research Institute, Yonsei University College of Medicine, Seoul, Republic Of Korea

²Department of Cardiology, Yonsei University College of Medicine, Seoul, Republic Of Korea

³Department of Neurology, Yonsei University College of Medicine, Seoul, Republic Of Korea

Corresponding Author:

Hyo Suk Nam, MD, PhD

Department of Neurology

Yonsei University College of Medicine

50 Yonsei-ro, Seodaemun-gu

Seoul, 03722

Republic Of Korea

Phone: 82 2 2228 1617

Fax: 82 2 393 0705

Email: hsnam@yuhs.ac

Abstract

Background: The pronator drift test (PDT), a neurological examination, is widely used in clinics to measure motor weakness of stroke patients.

Objective: The aim of this study was to develop a PDT tool with machine learning classifiers to detect stroke symptoms based on quantification of proximal arm weakness using inertial sensors and signal processing.

Methods: We extracted features of drift and pronation from accelerometer signals of wearable devices on the inner wrists of 16 stroke patients and 10 healthy controls. Signal processing and feature selection approach were applied to discriminate PDT features used to classify stroke patients. A series of machine learning techniques, namely support vector machine (SVM), radial basis function network (RBFN), and random forest (RF), were implemented to discriminate stroke patients from controls with leave-one-out cross-validation.

Results: Signal processing by the PDT tool extracted a total of 12 PDT features from sensors. Feature selection abstracted the major attributes from the 12 PDT features to elucidate the dominant characteristics of proximal weakness of stroke patients using machine learning classification. Our proposed PDT classifiers had an area under the receiver operating characteristic curve (AUC) of .806 (SVM), .769 (RBFN), and .900 (RF) without feature selection, and feature selection improves the AUCs to .913 (SVM), .956 (RBFN), and .975 (RF), representing an average performance enhancement of 15.3%.

Conclusions: Sensors and machine learning methods can reliably detect stroke signs and quantify proximal arm weakness. Our proposed solution will facilitate pervasive monitoring of stroke patients.

(*J Med Internet Res* 2017;19(4):e120) doi:[10.2196/jmir.7092](https://doi.org/10.2196/jmir.7092)

KEYWORDS

medical informatics; machine learning; motor; neurological examination; stroke

Introduction

Stroke is one of the main causes of death and disability worldwide [1]. One-third of stroke patients experience recurrent strokes. Muscle weakness is the most frequent sign of stroke and is related to disability [2]. Pronator drift, an indication of arm weakness, is mainly caused by subtle upper motor neuron disorders and is measured using the pronator drift test (PDT)

[3]. PDT has higher sensitivity than other neurological examinations including forearm roll, segmental motor exam, the Barr test, the Mingazzinis movements, and tendon reflexes [4]. Most stroke patients are diagnosed with the help of trained neurologists who perform bedside neurological examination, including PDT. However, early detection of stroke is critical because the effectiveness of thrombolytic therapy is time-dependent, and earlier treatment results in better outcomes

[5]. In addition to the need for instant examination, objectivity and accuracy need to be improved, because the conventional PDT performed by an inexperienced observer can result in missed rapid jitter of arm movement. To improve subjective decision making in the context of the conventional PDT, we developed an objective tool (the iPronator) to measure drift and pronation, and reported its feasibility and usefulness in a previous study [6]. In this study, we propose a decision support solution that can distinguish between the PDT properties of stroke patients and healthy people using representative machine learning algorithms.

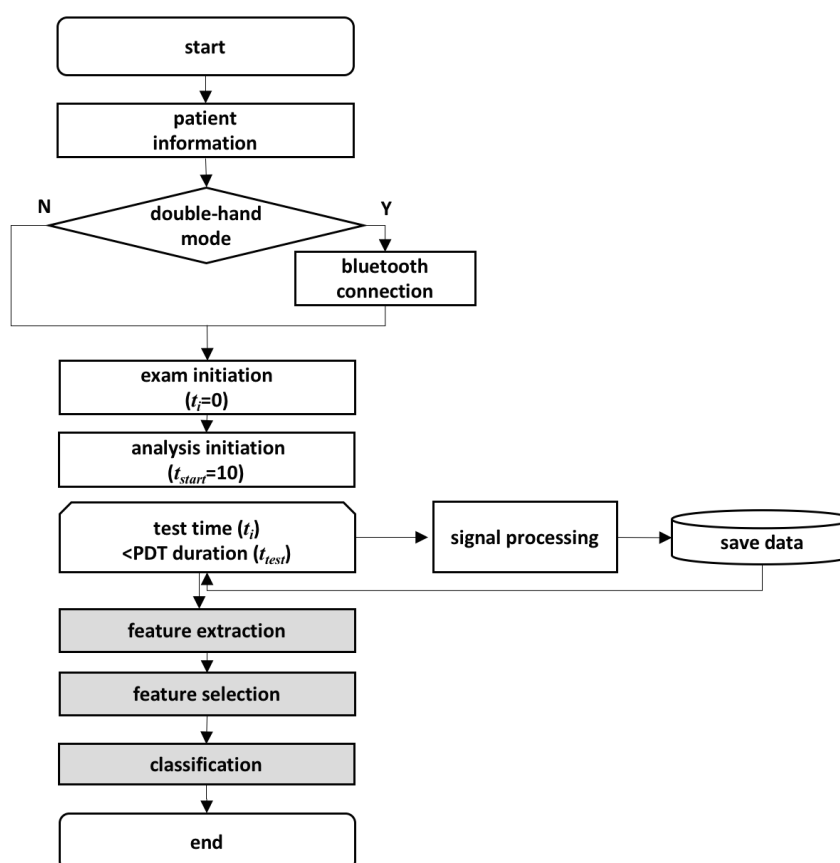
Methods

Study Design

We applied machine learning methods to detect arm weakness in stroke patients (Figure 1). First, accelerometer data from PDT

were collected from patients and healthy controls for a predefined period. We separated the start time for examination and analysis to exclude the effect of initial dip, which is commonly observed for upper extremity weakness [6]. In this work, the duration of PDT was set to 20 seconds, and the analysis began 10 seconds after the examination started. Next, our feature extraction task produced PDT features from the collected signals. Then, the feature selection task chose effective predictors among extracted features for the enhanced classification. Finally, after feature selection, machine learning algorithms modeled the classification for PDT. This study was approved by the Severance Hospital Institutional Review Board, and informed consent was obtained from all subjects.

Figure 1. Flowchart of pronator drift test (PDT) software.



Participants

A total of 26 subjects (10 men and 16 women) were recruited and assigned to the patient group or healthy control group. The ages of the participants ranged from 27 to 84 years, with an average of 58.2 (SD 17.8). During the study period, 16 consecutive stroke patients with mild upper arm weakness were enrolled. Exclusion criteria were patients who were unable to sit and had bilateral arm weakness or preexisting chronic arm weakness. A neurologist graded the muscle power of patients

and healthy controls using the Medical Research Council (MRC) scale, which is widely used to evaluate motor weakness (Table 1) [7]. Patients with MRC scores between 0 and 3 were also excluded because PDT was designed for subjects who are able to resist gravity and the weight of the measuring device on the arm. Healthy controls consisted of subjects who had no upper arm weakness and no history of neurological disease. All healthy controls were graded MRC 5. In total, 6 stroke patients were graded MRC 4, 7 were graded MRC 4+, and 3 were graded MRC 5.

Table 1. Muscle power grading using the Medical Research Council scale.

Grade	Description
0	No contraction
1	Flicker or trace of contraction
2	Active movement with gravity eliminated
3	Active movement against gravity
4	Active movement against gravity and moderate resistance
4+	Active movement against gravity and strong resistance
5	Normal power

Sensor Signal Processing

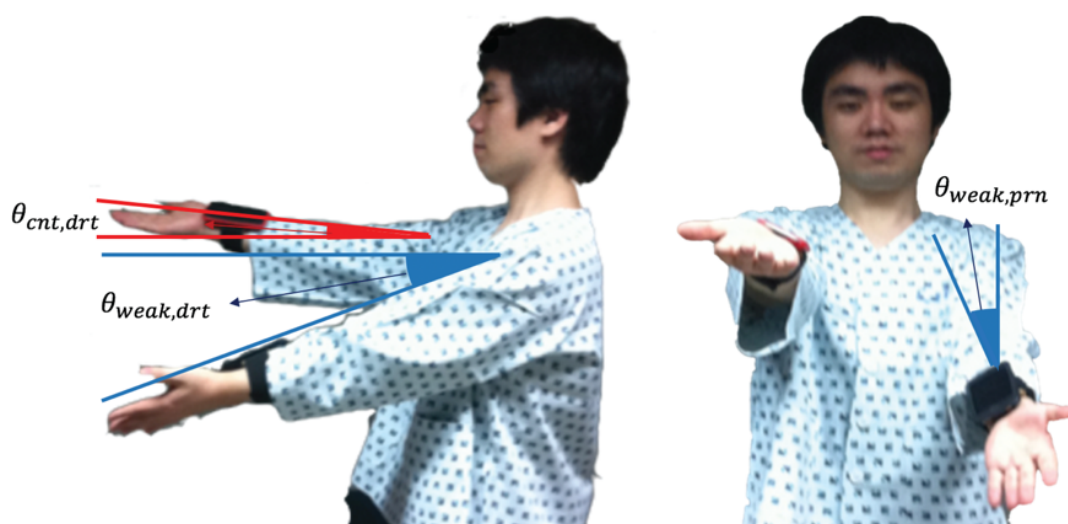
The integrated, low-power, three-axis accelerometer (LIS331DLH, ST-Microelectronics) in the mobile phone was used to measure pronation and drift. The accelerometer has a low-power mode and high accuracy of 1% on its lowest measurement range ($\pm 2g$) and approximately 0.1% on its highest measurement range ($\pm 8g$) [8]. Any device equipped with sensors, including mobile phones or wrist bands, can be used as a sensorized PDT tool.

Demographic information was collected according to predefined protocol. Patients were asked to conduct the PDT trial after registration. Two sensing devices were placed on each of a subject's wrists, as shown in Figure 2. When the mode was set to double-hand mode, the two devices were paired with a

Bluetooth connection. In the initial state of PDT, patients were asked to extend both arms anteriorly and hold them at shoulder height with palms facing up.

The time frame of the test was initialized, and then the PDT software initiated the measurement of arm movement to calculate the degree of pronation and drift. The procedure continued for the predefined test duration, calculating drift and pronation of the weak side. PDT simultaneously measures the movement of the counter-side by calculating the drift and pronation based on the fact that the counter-limb of the defective side also moves [6,9]. Measured values were subsequently saved for feature extraction in the analysis step. On the basis of collected data for the test duration, the properties of PDT were extracted and input into the classifiers (see Multimedia Appendix 1).

Figure 2. The pronator drift test: (a) the degree of drift in the weak arm and counter-arm of a patient was measured by the drift angle from the horizontal plane, and (b) the degree of pronation was assessed in front of the patient.



Decision Support by Pronator Drift Test Classifiers

We performed machine learning classification of PDT results from stroke patients versus those from healthy controls using MATLAB (Mathworks) [10], WEKA (University of Waikato) [11], and Medcalc (Medcalc Software) [12]. As a preprocessing step before machine learning, we implemented feature selection to identify which features were discriminant predictors to enhance the performance of the machine learning algorithms by eliminating redundant and irrelevant attributes [13]. We used

a wrapper approach for feature selection; this assesses subsets of extracted features according to their contribution to the classification performance [14,15]. Feature selection considers the employed classification model as an unseen part and assesses the subset of features according to their usefulness to a given classifier. Best-first search was used to traverse the space of candidate subsets and greedily find the optimal subset [15].

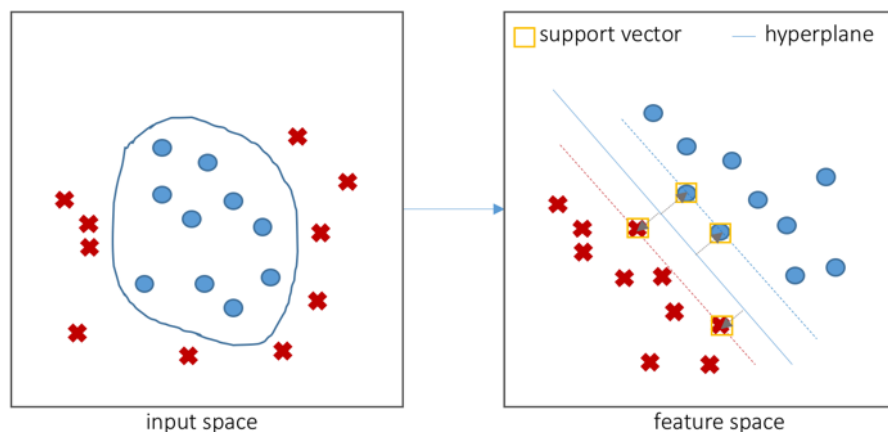
Next, a series of machine learning techniques [16], namely support vector machine (SVM), radial basis function network

(RBFN), and random forest (RF), were implemented. We selected such methods based on the findings in the research that compared 17 families of classifiers using 121 datasets, resulting in RF, SVM, and neural network–ranked top families [17]. Details of these machine learning algorithms are beyond the scope of this paper; thus, we only provide a brief description of each method.

Support Vector Machine

SVM is a machine learning algorithm developed by Cortes and Vapnik [18]. An SVM as a classifier trains a function that calculates a score for a new input to separate samples into two classes by building a hyperplane, which maintains a maximum margin between support vectors (Figure 3).

Figure 3. An example of a support vector machine with four support vectors in feature space.



If the output of the scoring function is negative, then the input is classified as belonging to the negative class; if the score is positive, the input is classified as belonging to the positive class. The scoring function is expressed as (eq.1; see [Multimedia Appendix 2](#)), where $x^{(i)}$ represents the i^{th} input vector, $y^{(i)}$ represents the class label of the i^{th} training data, and α_i is the coefficient associated with the training sample. The function K , which is called a kernel function, operates on the two vectors and reduces dimensions to simplify computation of the product of vectors. Among various kernel functions, we applied the polynomial kernel for the PDT classifier [19].

Radial Basis Function Network

RBFN is a neural network classifier that computes the Euclidean distance between a new input vector and the prototype vector corresponding to each neuron to measure the similarity between them [20] (Figure 4).

Each neuron's activation function is denoted as (eq. 2; see [Multimedia Appendix 2](#)), where μ_i and β_i are the prototype vector and the coefficient of the corresponding neuron i , respectively. The training process for an RBFN selects the

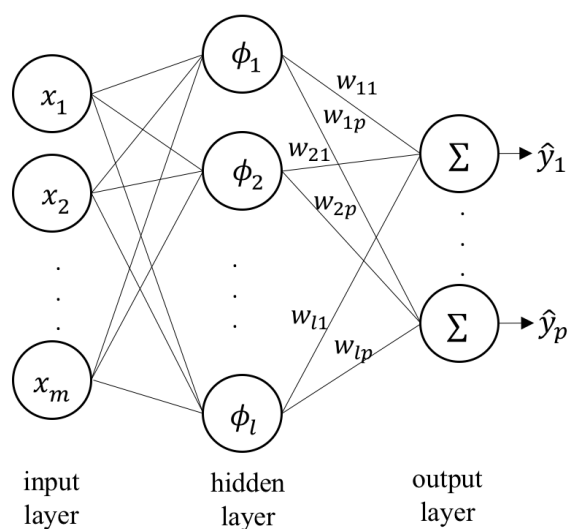
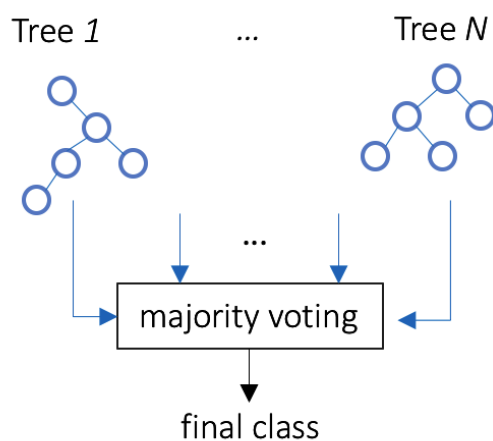
prototype vector, coefficient for each of the RBF neurons, and the matrix of output weights w_{ij} between the RBF neurons and the output node j . The decision for each class j is decided by (eq. 3).

Various approaches have been proposed to select prototypes from input vectors. We applied K means clustering as the base function to select prototypes [21].

Random Forest

RF is an ensemble predictor that uses a combination of multiple decision trees [22]. Prediction in the training stage is determined by voting from the forest in which an individual tree predicts the target class depending on the values of a random vector sampled independently (Figure 5).

We applied SVM, RBFN, and RF classifiers to the entire set of PDT features. Leave-one-out cross-validation (LOOCV) was applied, because we had a small number of training samples. The performance of classifiers was measured by calculating sensitivity, specificity, the F measure, and area under the receiver operating characteristic curve (AUC).

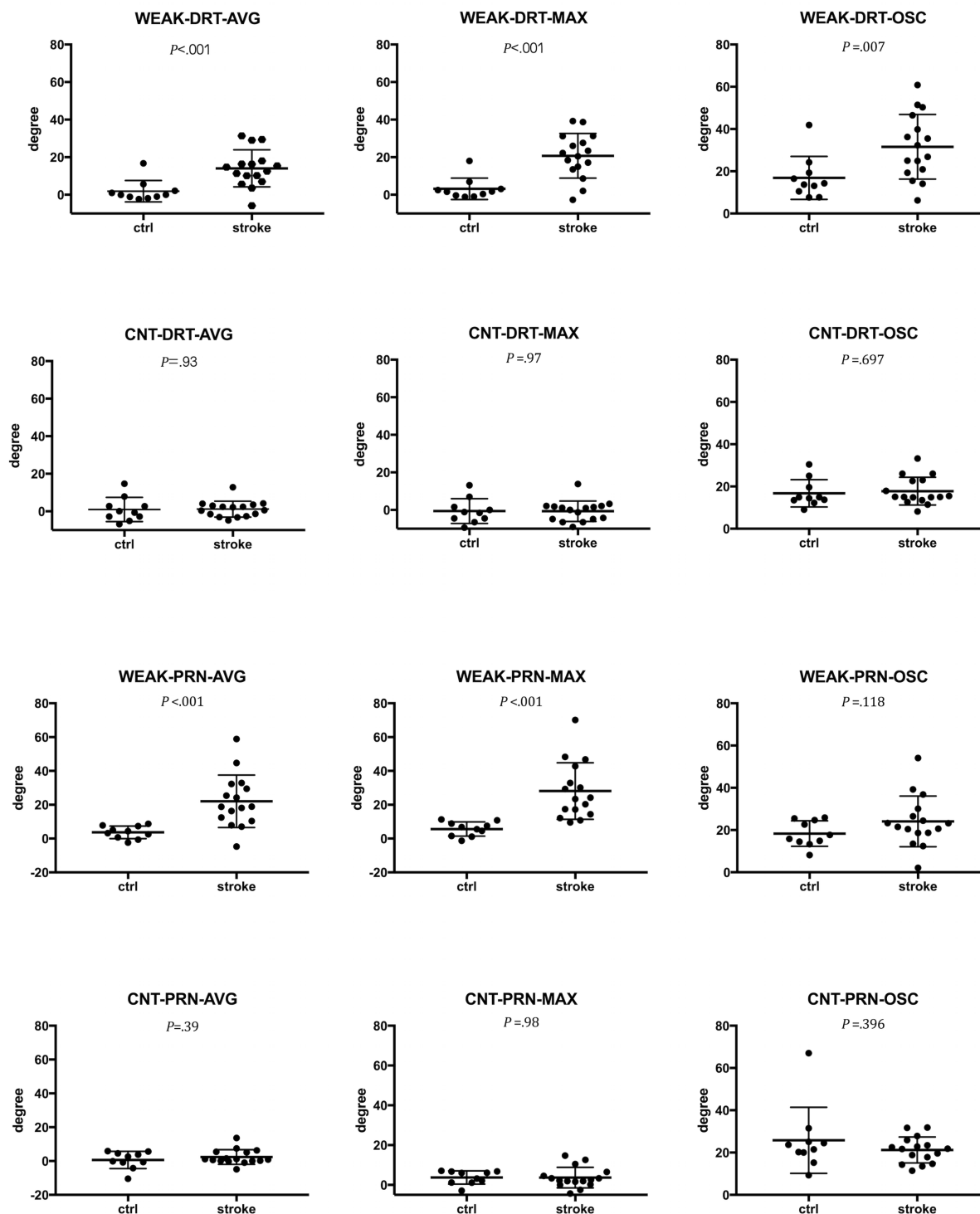
Figure 4. The architecture of an radial basis function network.**Figure 5.** A simplified random forest.

Results

Statistical Properties of Pronator Drift Test Features

We compared PDT features using the t test. Figure 6 shows the means and standard deviations of the PDT features. Among the

12 PDT features, WEAK-DRT-AVG, WEAK-DRT-MAX, WEAK-DRT-OSC, WEAK-PRN-AVG, and WEAK-PRN-MAX were significantly different between stroke patients and controls.

Figure 6. Differences of degree in PDT features between stroke patients and controls. Values are mean, standard deviation, and P value.

Selected Attributes

Among the extracted PDT features, SVM, RBFN, and RF classifiers selected discriminative features (Table 2). We applied the wrapping approach for feature selection. Feature selection procedure for SVM calculates the usefulness of features and extracts a discriminant feature set {WEAK-PRN-MAX,

WEAK-DRT-AVG} for SVM. Feature selection for RBFN reduced all features to three features of pronation on the weak side {WEAK-PRN-MAX, WEAK-PRN-AVG, WEAK-PRN-OSC}. Feature selection for RF resulted in identification of the maximum degree of pronation on the weak side and average drift of the counter-side as useful features for RF classification {WEAK-PRN-MAX, CNT-DRT-AVG}.

Table 2. Features selected in pronator drift test classification.

Selected feature	Machine learning classifier			No. of classifiers that selected the feature
	SVM	RBFN	RF	
WEAK-PRN-MAX	X	X	X	3
WEAK-PRN-AVG		X		1
WEAK-PRN-OSC		X		1
WEAK-DRT-AVG	X			1
CNT-DRT-AVG			X	1
No. of features	2	3	2	

Stroke Classifiers With Selected Pronator Drift Test Predictors

Using the selected features for the SVM classifier, we built an SVM PDT classifier with a polynomial kernel. PDT feature vector $SVM-PDT_{vec}$ and score function $f(SVM-PDT_{vec})$ for the SVM classifier were modeled as (eq. 4; see [Multimedia Appendix 2](#)).

The derived score function was used to assign training instances into positive class and negative class that contained positive and negative values of the score function, respectively. As

shown in [Figure 7](#) (a), stroke patients' PDT features were mapped on the surface of the score function above the cut-plane. Two control cases were misclassified as belonging to the stroke group. As shown in [Figure 7](#) (b), the score function for the control group produced values less than the cut-plane, and one stroke case was misclassified as a control case.

The RBFN classifier for stroke patients constructed four clusters to calculate radials in the RBFN without feature selection and two clusters for the RBFN including feature selection. The RF classifier combined decision trees as depicted in [Figure 8](#); two cases were misclassified ([Figure 9](#)).

Figure 7. Plot of support vector machine (SVM) score function and decision by the SVM classifier: (a) positive scores of the SVM classifier for input (above the plane); two control cases were misclassified as patients, and (b) negative scores of the SVM classifier for input (below the plane); one stroke patient case was misclassified as a control case.

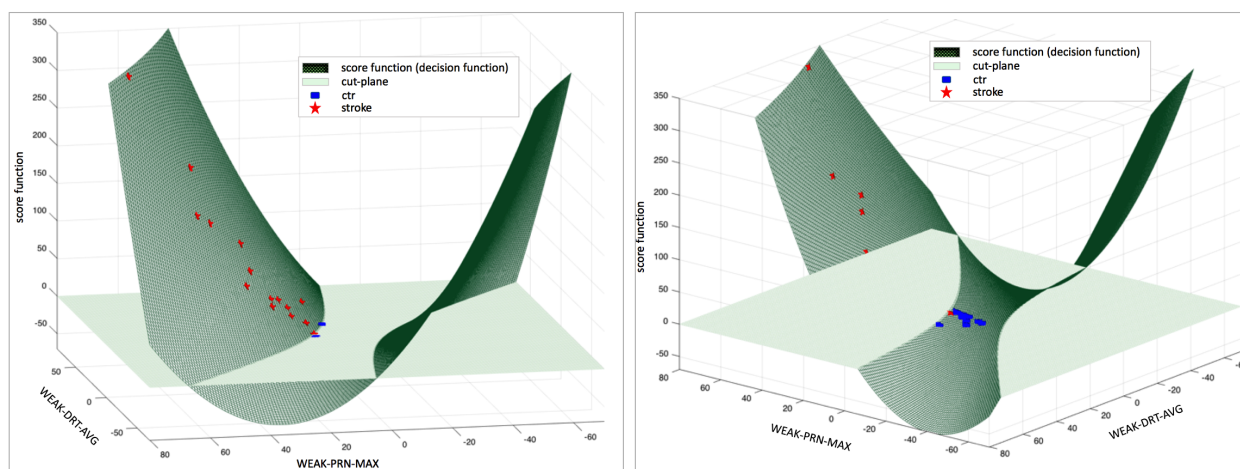
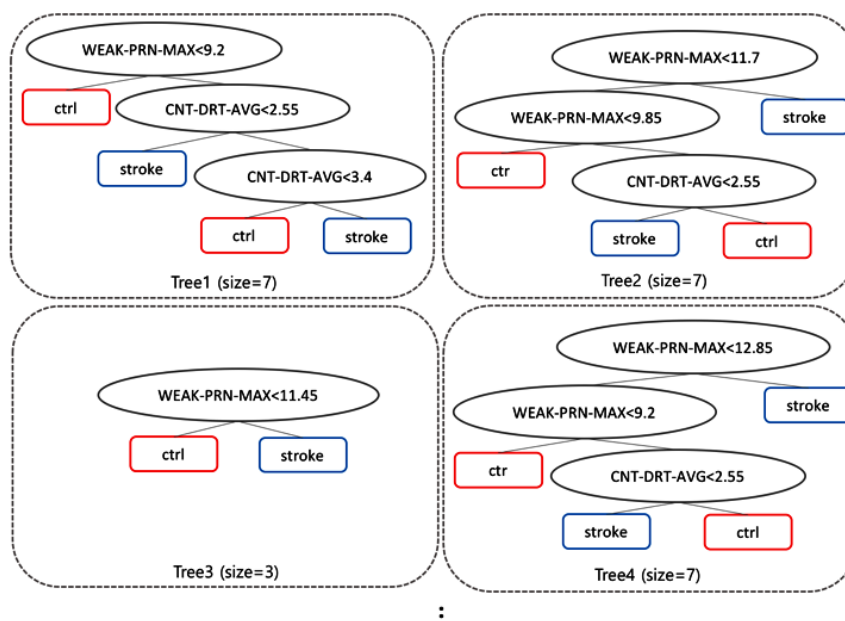
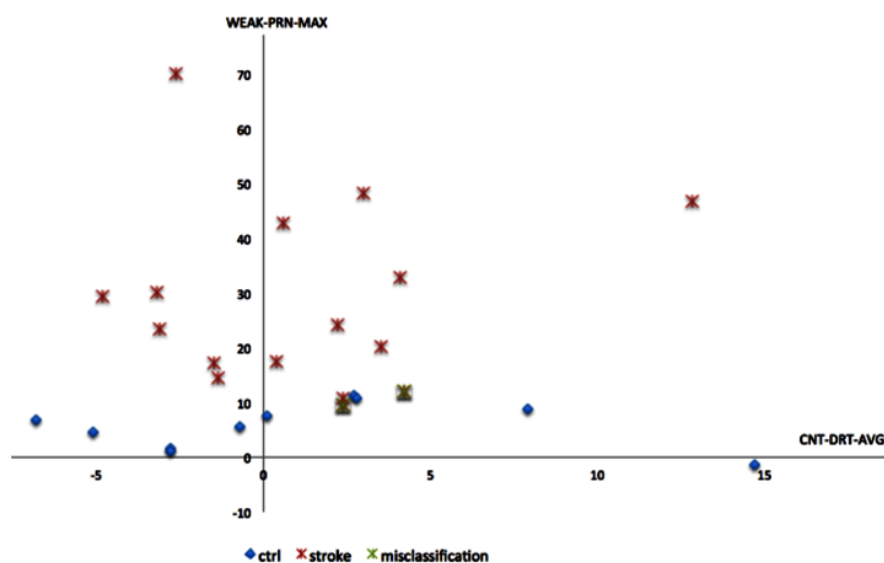


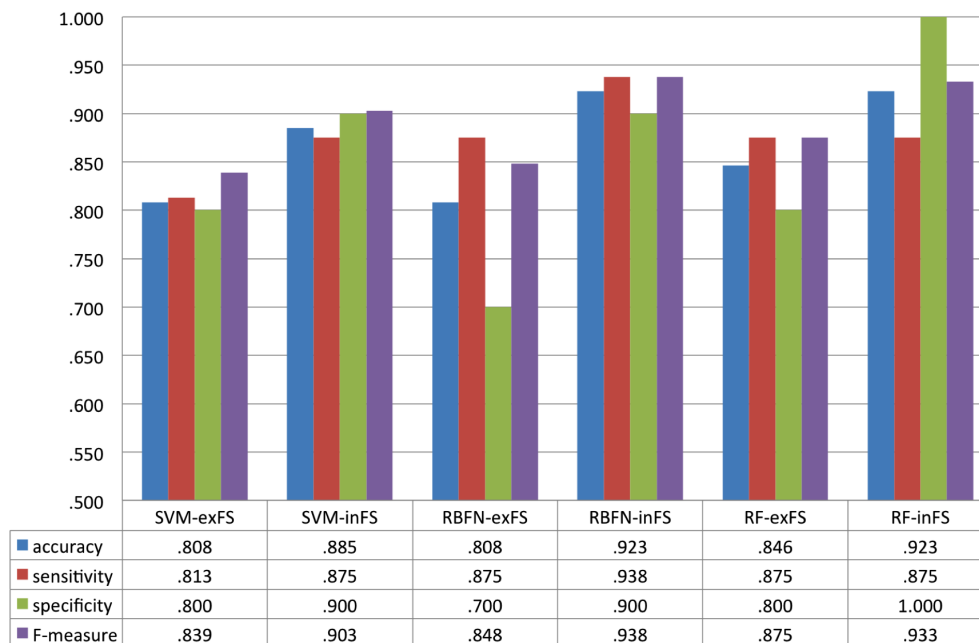
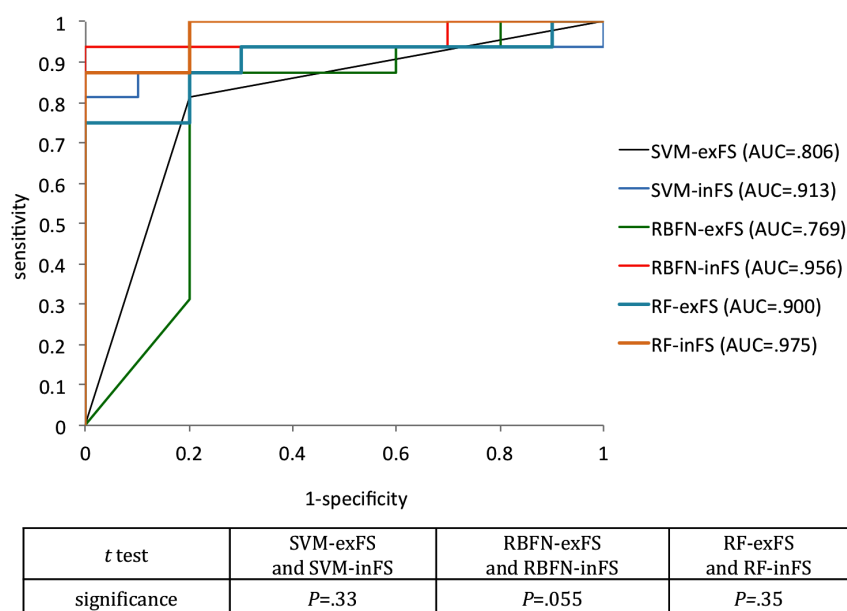
Figure 8. Random forest composed of decision trees as a pronator drift test classifier.**Figure 9.** Weakness detection using a random forest, including feature selection.

Performance of Stroke Classifiers With Feature Selection

Figure 10 shows the performance of classifiers in terms of accuracy, sensitivity, specificity, and F-measure. The accuracy of all classifiers was improved by feature selection; accuracies of the classifiers improved by 9.53% (.808 with SVM-exFS [excluding feature selection] vs .885 with SVM-inFS [including feature selection]), 14.23% (.808 with RBFN-exFS vs .923 with

RBFN-inFS), and 9.10% (.846 with RF-exFS vs .923 with RF-inFS), respectively.

The stroke classifiers had an accuracy of up to 92.3% for detecting stroke (RBFN-inFS, AUC = .956; RF-inFS, AUC = .975), and RF had the best AUC of .975 when feature selection was applied (Figure 11). To compare the means of the individual AUCs of methods with and without feature selection, *t* test was also performed.

Figure 10. Performance of stroke classifiers excluding/including feature selection.**Figure 11.** Receiver operating characteristic (ROC) curve of pronator drift test classifiers and t test of area under the receiver operating characteristic (AUC) (SVM-exFS and SVM-inFS: support vector machine excluding and including feature selection; RBFN-exFS and RBFN-inFS: radial basis function excluding and including feature selection; RF-exFS and RF-inFS: random forest excluding and including feature selection).

Discussion

Summary

We developed a novel method to monitor pronator drift using sensor-equipped devices. We investigated and demonstrated the feasibility of machine learning analysis of the information obtained via the sensors and found that the combination of these methods can detect the neurological deficit of subtle motor weakness. We demonstrated that machine learning-based classifiers correctly classified up to 92.3% of PDT cases.

Review of Previous Studies

Machine learning has recently been adopted in medicine and its usage includes various medical studies: health care utilization based on patients' social network data [23,24]; predicting mortality after surgery [25]; estimating the risk of treatment outcomes [26]; predicting deterioration using electronic medical records with physiological signals [27]; and activity monitoring [28]. This popularity is because of the advantage of easily incorporating new data to improve prediction performance [29] and to identify discriminant variables for prediction [30]. Machine learning has also improved assessment and outcome prediction in stroke studies. Decision tree [31], SVM, and neural

network [29] have been utilized to predict the outcome of acute ischemic stroke. SVM-integrated regression models have also been proposed to predict stroke [32].

In addition to machine learning analysis, sensor-based measurement improved the detection of abnormality and outcome prediction. Task-oriented, arm-hand training using sensor measurement was introduced in [33], and a machine learning method with pressure sensor-embedded smart shoes discriminated the alcohol-induced gait [34].

In this study, we utilized an off-the-shelf smart device embedding accelerometer for the measurement of arm weakness. The use of mobile phones or general activity trackers elaborated the high accessibility of users. Recent studies demonstrated the validity of using accelerometer in iPhone for the physical activity monitoring [28], the extraction of heart rate [35], and applications for Parkinson and Holmes tremor [36].

We previously developed a sensor-based mobile tool (the iPronator) and reported that the iPronator app was useful and feasible for detecting mild arm weakness and quantifying the degree of weakness. Moreover, the iPronator can also detect functional recovery after one week in patients with acute stroke [6]. In this study, we further evaluated whether machine learning could improve detection of the presence of mild arm weakness after stroke. Although information technology and mobile devices are increasingly used in the management of stroke [37], most researchers have focused on analyzing medical records, including laboratory results, to predict mortality and the outcome of care. However, as far as we know, no studies have reported using machine learning-based classifiers to detect weakness associated with stroke and PDT.

Predictors for Stroke Decision Support

PDT is known to be a sensitive neurological test of weakness. If a patient has pronator drift, positive test indicates the damage in motor pathway from the opposite side of the brain [38]. The pronator drift is determined by various conditions including motor deficit, sensory deficit, cerebellar drift, parietal lobe lesions, and conversion disorders: cerebellar disease causes outward and upward drift; patients with parietal lobe lesions exhibits loss of position sense, which causes updrift with the involved arm rising overhead; and functional upper limb paresis causes drift without pronation [39]. Due to such various causes of pronator drift, the result of PDT varies on the condition of patients: one study showed that patients with subtle difficulty in routine activity had positive PDT in 38 (76%) out of 50 patients [40], whereas another study showed positive PDT in only 43.8% of patients with cerebral lesions [41]. We infer that such variability originated the outperformance of machine learning methods in the classification of PDT, since machine learning is strong in sophisticated pattern recognition by delineating patterns from relations between less significant variables as well as key variables. As shown in Figure 6 and Table 2, the key variable WEAK-PRN-MAX, which showed a significant difference between patients and controls in statistical analysis, was the dominant feature selected by all classifiers and can therefore be considered a dominant predictor for detecting weakness. In addition, machine learning methods (RF and RBFN) utilized CNT-DRT-AVG and WEAK-PRN-AVG

resulting in high detection rate, although both WEAK-PRN-OSC ($P=.12$) and CNT-DRT-AVG ($P=.93$) were not significantly different between the patient and control groups statistically.

In searching for optimal subset of features for classifiers, the wrapper method resolves the problem of high-dimensional features space and feature redundancy to improve the intelligent decision [42]. In [43], feature selection of RF and SVM conducted the phenotyping through limiting the number of variables based on the importance in RF. Especially, feature selection was prominent in the classification with insufficient sample data by restricting the number of features in the classifier to $n/10$ for the best performance [44]. In this study, 12 PDT features extracted from sensor signal processing were narrowed down to two or three features to support decision for stroke. We investigated the effect of reduction of dominant features by comparing the performance of stroke classifiers including and excluding feature selection. As shown in Figure 10, the accuracy of RF-exFS classifier (accuracy = .846) obtained higher accuracy than SVM-exFS and RBFN-exFS (accuracy = .808). The result is induced from the RF's intrinsic property that RF contains the feature selection mechanism in the classification as it randomly selects different variables to construct each tree within its forest. This randomization is known to be effective in eliminating noises and reflecting multivariate interactions with other variables [22]. Therefore, the effect of feature selection appears stronger in RBFN and SVM than RF as shown in Figures 10 and 11. Feature selection improved accuracy by 14.23%, 9.53%, and 9.1% in RBFN, SVM, and RF, respectively. AUC was also improved along with FS by 15.3 % in average (SVM: .806-.913, RBFN: .769-.956, and RF: .900-.975). We conducted the t test between AUCs of classifiers with and without feature selection and the difference between RBFN-inFS and RBFN-exFS was most significant ($P=.06$).

Implications of This Study and Perspectives

Although many mobile devices using sensors have been developed and marketed to doctors and health care providers for years, adoption of machine learning in stroke patients is still in its infancy. In particular, patients or stroke witnesses do not have any tools to detect stroke or communicate with health care providers.

As we described, time is critical in acute stroke management, including thrombolytic treatments. Thrombolytic treatment should begin within 4.5 hours after the onset of a stroke. Moreover, earlier treatment results in better outcomes within the treatment window. Therefore, rapid evaluation of motor weakness is important. To reduce hospital delay and efficiently dispatch patients in emergent medical services, integration of machine learning methods with mobile devices with sensors might be useful.

In addition, evaluation by neurologists may be delayed in busy emergency room. To overcome these limitations and improve patients' care, a simple bedside tool and objectifying the results are important. The proposed solution can connect patients and health care providers in rapid communication and, ultimately, these approaches may improve the care of stroke patients at low cost.

As another application, the proposed tool might be helpful in monitoring of stroke recurrence in subacute-to-chronic period after stroke. Although we previously demonstrated that the objective of PDT was useful in detecting functional recovery in patients with acute stroke, further long-term follow-up studies can provide its usefulness in detecting stroke occurrence, because machine learning model can be improved with the big data, and personalized history of measurement can provide tailoring in stroke management.

Limitations and Future Works

In this study, a total of 26 sample data were analyzed by machine learning methods. The performance of machine learning algorithms is known to be affected by the quality and quantity of training data. We adopted LOOCV to complement the small

number of instances, and the large data accumulation in further study may diminish the requirement of LOOCV, which requires more computing time and resources.

We also plan to develop a new version of iPronator with small-sized, 3-axis accelerometer and 3-axis gyroscope, since the weight of smart devices may affect the result of PDT. In this study, we excluded the initial dip caused by the mobile phone's own weight.

The diverse causes of pronator drift can be another limitation for this tool in the detection of stroke, because there exist false positive signs in PDT caused by other lesions outside the motor pathway. The future development extends the current binary classification into multi-classification clustering various causes of PDT.

Acknowledgments

This study was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (HC15C1056), and grants from the R&D Program of Fire Fighting Safety and 119 Rescue Technology, funded by the Ministry of Public Safety and Security, Republic of Korea (MPSS-2015-70).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Features extracted from the pronator drift test.

[[PNG File, 276KB](#) - [jmir_v19i4e120_app1.png](#)]

Multimedia Appendix 2

Equations for pronator drift test classifiers.(eq.1): score function of support vector machine; (eq.2): activation function of neurons in radial basis function network; (eq.3): decision function of radial basis function network; (eq.4): feature vector and score function of support vector machine for pronator drift test.

[[PNG File, 403KB](#) - [jmir_v19i4e120_app2.png](#)]

References

1. GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016 Oct 08;388(10053):1459-1544 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(16\)31012-1](#)] [Medline: [27733281](#)]
2. Rathore SS, Hinn AR, Cooper LS, Tyroler HA, Rosamond WD. Characterization of incident stroke signs and symptoms: findings from the atherosclerosis risk in communities study. *Stroke* 2002 Nov;33(11):2718-2721 [[FREE Full text](#)] [Medline: [12411667](#)]
3. Darcy P, Moughty AM. Images in clinical medicine. Pronator drift. *N Engl J Med* 2013 Oct 17;369(16):e20. [doi: [10.1056/NEJMicm1213343](#)] [Medline: [24131195](#)]
4. Teitelbaum JS, Eliasziw M, Garner M. Tests of motor function in patients suspected of having mild unilateral cerebral lesions. *Can J Neurol Sci* 2002 Nov;29(4):337-344. [Medline: [12463489](#)]
5. Hacke W, Kaste M, Bluhmki E, Brozman M, Dávalos A, Guidetti D, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *N Engl J Med* 2008 Sep 25;359(13):1317-1329. [doi: [10.1056/NEJMoa0804656](#)] [Medline: [18815396](#)]
6. Shin S, Park E, Lee DH, Lee K, Heo JH, Nam HS. An objective pronator drift test application (iPronator) using handheld device. *PLoS One* 2012;7(7):e41544 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0041544](#)] [Medline: [22911811](#)]
7. Vanpee G, Hermans G, Segers J, Gosselink R. Assessment of limb muscle strength in critically ill patients: a systematic review. *Crit Care Med* 2014 Mar;42(3):701-711. [doi: [10.1097/CCM.0000000000000030](#)] [Medline: [24201180](#)]
8. Dunn P. Measurement and Data Analysis for Engineering and Science, Third Edition. Boca Raton: CRC Press; 2014.

9. Kobori S, Takizawa S, Sekiyama S, Takagi S. Ten centimeter elevation of the contralateral leg is enough to evaluate Hoover sign. *Intern Med* 2007;46(1):55-56 [[FREE Full text](#)] [Medline: [17202735](#)]
10. Moore H. *MATLAB for Engineers* (3rd Edition). Upper Saddle River: Prentice Hall; 2012.
11. Smith TC, Frank E. Introducing Machine Learning Concepts with WEKA. *Methods Mol Biol* 2016;1418:353-378. [doi: [10.1007/978-1-4939-3578-9_17](#)] [Medline: [27008023](#)]
12. Schoonjans F, Zalata A, Depuydt CE, Comhaire FH. MedCalc: a new computer program for medical statistics. *Comput Methods Programs Biomed* 1995 Dec;48(3):257-262. [Medline: [8925653](#)]
13. Saey Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* 2007 Oct 1;23(19):2507-2517 [[FREE Full text](#)] [doi: [10.1093/bioinformatics/btm344](#)] [Medline: [17720704](#)]
14. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell* 1997 Dec;97(1-2):273-324. [doi: [10.1016/S0004-3702\(97\)00043-X](#)]
15. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res* 2003;3:1157-1182. [doi: [10.1162/153244303322753616](#)]
16. Larrañaga P, Calvo B, Santana R, Bielza C, Galdiano J, Inza I, et al. Machine learning in bioinformatics. *Brief Bioinform* 2006 Mar;7(1):86-112 [[FREE Full text](#)] [Medline: [16761367](#)]
17. Fernández-Delgado M, Cernadas E, Barro S, Amorim D. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *J Mach Learn Res* 2014;15:3133-3181.
18. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995;20(3):273-297. [doi: [10.1023/A:1022627411411](#)]
19. Schoˆlkopf B, Smola A. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. Cambridge, Mass: MIT Press; 2002.
20. Schwenker F, Kestler HA, Palm G. Three learning phases for radial-basis-function networks. *Neural Netw* 2001 May;14(4-5):439-458. [doi: [10.1016/S0893-6080\(01\)00027-2](#)]
21. Wettschereck D, Dietterich T. Improving the performance of radial basis function networks by learning center locations. In: *Proceedings of the 4th International Conference on Neural Information Processing Systems*. 1991 Presented at: NIPS'91; December 2; Denver, Colorado p. 1133-1140.
22. Breiman L. Random forests. *Mach Learn* 2001;45(1):5-32. [doi: [10.1023/A:1010933404324](#)]
23. Agarwal V, Zhang L, Zhu J, Fang S, Cheng T, Hong C, et al. Impact of predicting health care utilization via web search behavior: a data-driven analysis. *J Med Internet Res* 2016 Sep 21;18(9):e251 [[FREE Full text](#)] [doi: [10.2196/jmir.6240](#)] [Medline: [27655225](#)]
24. Cole-Lewis H, Varghese A, Sanders A, Schwarz M, Pugatch J, Augustson E. Assessing electronic cigarette-related tweets for sentiment and content using supervised machine learning. *J Med Internet Res* 2015;17(8):e208 [[FREE Full text](#)] [doi: [10.2196/jmir.4392](#)] [Medline: [26307512](#)]
25. Allyn J, Allou N, Augustin P, Philip I, Martinet O, Belghiti M, et al. A comparison of a machine learning model with EuroSCORE II in predicting mortality after elective cardiac surgery: a decision curve analysis. *PLoS One* 2017;12(1):e0169772 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0169772](#)] [Medline: [28060903](#)]
26. Koutsouleris N, Kahn RS, Chekroud AM, Leucht S, Falkai P, Wobrock T, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry* 2016 Oct;3(10):935-946. [doi: [10.1016/S2215-0366\(16\)30171-7](#)] [Medline: [27569526](#)]
27. Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. *JMIR Med Inform* 2016 Sep 30;4(3):e28 [[FREE Full text](#)] [doi: [10.2196/medinform.5909](#)] [Medline: [27694098](#)]
28. Wu W, Dasgupta S, Ramirez EE, Peterson C, Norman GJ. Classification accuracies of physical activities using smartphone motion sensors. *J Med Internet Res* 2012;14(5):e130 [[FREE Full text](#)] [doi: [10.2196/jmir.2208](#)] [Medline: [23041431](#)]
29. Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. *PLoS One* 2014;9(2):e88225 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0088225](#)] [Medline: [24520356](#)]
30. Laux da Costa L, Delcroix M, Dalla Costa ER, Prestes IV, Milano M, Francis SS, et al. A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases. *Tuberculosis (Edinb)* 2015 Jul;95(4):421-425 [[FREE Full text](#)] [doi: [10.1016/j.tube.2015.04.008](#)] [Medline: [26025597](#)]
31. Alexopoulos E, Dounias G, Vemmos K. Medical diagnosis of stroke using inductive machine learning. *Machine Learning and Applications: Machine Learning in Medical Applications* 1999:20-23.
32. Khosla A, Cao Y, Lin C, Chiu H, Hu J, Lee H. An integrated machine learning approach to stroke prediction. 2010 Presented at: *Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining*; July 2010; Washington, DC, USA.
33. Lemmens RJ, Timmermans AA, Janssen-Potten YJ, Pulles SA, Geers RP, Bakx WG, et al. Accelerometry measuring the outcome of robot-supported upper limb training in chronic stroke: a randomized controlled trial. *PLoS One* 2014;9(5):e96414 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0096414](#)] [Medline: [24823925](#)]
34. Park E, Lee SI, Nam HS, Garst JH, Huang A, Champion A, et al. Unobtrusive and continuous monitoring of alcohol-impaired gait using smart shoes. *Methods Inf Med* 2017;56(1):74-82. [doi: [10.3414/ME15-02-0008](#)] [Medline: [27782289](#)]

35. Kwon S, Kim H, Park KS. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. *Conf Proc IEEE Eng Med Biol Soc* 2012;2012:2174-2177. [doi: [10.1109/EMBC.2012.6346392](https://doi.org/10.1109/EMBC.2012.6346392)] [Medline: [23366353](https://pubmed.ncbi.nlm.nih.gov/23366353/)]
36. Araújo R, Tábuas-Pereira M, Almendra L, Ribeiro J, Arenga M, Negrão L, et al. Tremor frequency assessment by iPhone® applications: correlation with EMG analysis. *J Parkinsons Dis* 2016 Oct 19;6(4):717-721. [doi: [10.3233/JPD-160936](https://doi.org/10.3233/JPD-160936)] [Medline: [27662333](https://pubmed.ncbi.nlm.nih.gov/27662333/)]
37. Nam HS, Park E, Heo JH. Facilitating stroke management using modern information technology. *J Stroke* 2013 Sep;15(3):135-143 [FREE Full text] [doi: [10.5853/jos.2013.15.3.135](https://doi.org/10.5853/jos.2013.15.3.135)] [Medline: [24396807](https://pubmed.ncbi.nlm.nih.gov/24396807/)]
38. Pullen RL. Neurologic assessment for pronator drift. *Nursing* 2004 Mar;34(3):22. [Medline: [15179998](https://pubmed.ncbi.nlm.nih.gov/15179998/)]
39. Campbell W, DeJong R. DeJong's the neurologic examination. Philadelphia: Lippincott Williams & Wilkins; 2005.
40. Weaver DF. A clinical examination technique for mild upper motor neuron paresis of the arm. *Neurology* 2000 Jan 25;54(2):531-532. [Medline: [10668739](https://pubmed.ncbi.nlm.nih.gov/10668739/)]
41. Maranhão ET, Maranhão-Filho P, Lima MA. The Digits Quinti Sign, Souques' Interosseous Phenomenon, and Pronator Drift Test as Subtle Motor Signs in Patients with Monohemispheric Brain Tumor. *Revista Brasileira de Cancerologia* 2008;54(3):227-230.
42. Bolón-Canedo V, Sánchez-Marño N, Alonso-Betanzos A. Recent advances and emerging challenges of feature selection in the context of big data. *Knowl Based Syst* 2015 Sep;86:33-45. [doi: [10.1016/j.knosys.2015.05.014](https://doi.org/10.1016/j.knosys.2015.05.014)]
43. Zhao J, Bodner G, Rewald B. Phenotyping: using machine learning for improved pairwise genotype classification based on root traits. *Front Plant Sci* 2016;7:1864 [FREE Full text] [doi: [10.3389/fpls.2016.01864](https://doi.org/10.3389/fpls.2016.01864)] [Medline: [27999587](https://pubmed.ncbi.nlm.nih.gov/27999587/)]
44. Raudys S, Jain A. Small sample size effects in statistical pattern recognition: recommendations for practitioners. *IEEE Trans Pattern Anal Machine Intell* 1991;13(3):252-264. [doi: [10.1109/34.75512](https://doi.org/10.1109/34.75512)]

Abbreviations

AUC: area under the receiver operating characteristic

FS: feature selection

MRC: Medical Research Council

PDT: pronator-drift test

RBFN: radial basis function network

RF: random forest

SVM: support vector machine

Edited by G Eysenbach; submitted 01.12.16; peer-reviewed by L Zhang, N Tran; comments to author 05.01.17; revised version received 02.02.17; accepted 05.03.17; published 18.04.17.

Please cite as:

Park E, Chang HJ, Nam HS

Use of Machine Learning Classifiers and Sensor Data to Detect Neurological Deficit in Stroke Patients

J Med Internet Res 2017;19(4):e120

URL: <http://www.jmir.org/2017/4/e120/>

doi: [10.2196/jmir.7092](https://doi.org/10.2196/jmir.7092)

PMID: [28420599](https://pubmed.ncbi.nlm.nih.gov/28420599/)

©Eunjeong Park, Hyuk-Jae Chang, Hyo Suk Nam. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 18.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Ethics and Privacy Implications of Using the Internet and Social Media to Recruit Participants for Health Research: A Privacy-by-Design Framework for Online Recruitment

Jacqueline Lorene Bender^{1,2}, PhD; Alaina B Cyr³, BSc; Luk Arbuckle⁴, BSc; Lorraine E Ferris², LLM, PhD

¹Electronic Living Laboratory for Interdisciplinary Cancer Survivorship Research (ELLICSR) Health, Wellness, and Cancer Survivorship Centre, Department of Supportive Care, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

²Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

³Cancer Education, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

⁴Electronic Health Information Lab, Children's Hospital of Eastern Ontario (CHEO) Research Institute, Ottawa, ON, Canada

Corresponding Author:

Jacqueline Lorene Bender, PhD

ELLICSR Health, Wellness, and Cancer Survivorship Centre

Department of Supportive Care

Princess Margaret Cancer Centre, University Health Network

B PMB 130 Toronto General Hospital

585 University Ave

Toronto, ON, M5G 2C4

Canada

Phone: 1 416 581 8606

Email: jackie.bender@uhnresearch.ca

Abstract

Background: The Internet and social media offer promising ways to improve the reach, efficiency, and effectiveness of recruitment efforts at a reasonable cost, but raise unique ethical dilemmas. We describe how we used social media to recruit cancer patients and family caregivers for a research study, the ethical issues we encountered, and the strategies we developed to address them.

Objective: Drawing on the principles of Privacy by Design (PbD), a globally recognized standard for privacy protection, we aimed to develop a PbD framework for online health research recruitment.

Methods: We proposed a focus group study on the dietary behaviors of cancer patients and their families, and the role of Web-based dietary self-management tools. Using an established blog on our hospital website, we proposed publishing a recruitment post and sharing the link on our Twitter and Facebook pages. The Research Ethics Board (REB) raised concern about the privacy risks associated with our recruitment strategy; by clicking on a recruitment post, an individual could inadvertently disclose personal health information to third-party companies engaged in tracking online behavior. The REB asked us to revise our social media recruitment strategy with the following questions in mind: (1) How will you inform users about the potential for privacy breaches and their implications? and (2) How will you protect users from privacy breaches or inadvertently sharing potentially identifying information about themselves?

Results: Ethical guidelines recommend a proportionate approach to ethics assessment, which advocates for risk mitigation strategies that are proportional to the magnitude and probability of risks. We revised our social media recruitment strategy to inform users about privacy risks and to protect their privacy, while at the same time meeting our recruitment objectives. We provide a critical reflection of the perceived privacy risks associated with our social media recruitment strategy and the appropriateness of the risk mitigation strategies that we employed by assessing their alignment with PbD and by discussing the following: (1) What are the potential risks and who is at risk? (2) Is cancer considered “sensitive” personal information? (3) What is the probability of online disclosure of a cancer diagnosis in everyday life? and (4) What are the public’s expectations for privacy online and their views about online tracking, profiling, and targeting? We conclude with a PbD framework for online health research recruitment.

Conclusions: Researchers, REBs, ethicists, students, and potential study participants are often unaware of the privacy risks of social media research recruitment and there is no official guidance. Our PbD framework for online health research recruitment is a resource for these wide audiences.

(*J Med Internet Res* 2017;19(4):e104) doi:[10.2196/jmir.7029](https://doi.org/10.2196/jmir.7029)

KEYWORDS

Internet; social media; ethics; privacy; recruitment; cancer

Introduction

Increasingly, health researchers are turning to the Internet to recruit people for research studies [1-4]. The wide penetration of the Internet and the increasing use of social media (eg, wikis, blogs, online communities, and social networking sites) create many new avenues for research recruitment. In particular, social networking sites, such as Facebook, Twitter, and Google+, offer several potential advantages. These have considerable reach, providing access to large heterogeneous populations as well as small, hard-to-reach subpopulations dealing with sensitive, stigmatizing, or rare health conditions. They offer powerful sharing features that researchers can leverage to engage the public in spreading the word about a research project and recruitment by “liking,” “favoriting,” “replying to,” or “retweeting.” They are flexible; recruitment notices can be turned on and off and content can be changed in real time, allowing researchers the ability to control and evaluate recruitment efforts [5]. They are economical, reducing the time and effort involved in recruitment at reduced cost relative to other approaches [3].

However, the use of the Internet and social media as a health research recruitment tool raises unique ethical issues in part because personal and sensitive information may be collected from individuals without their knowledge or consent before they enroll in a study. The simple act of clicking on a recruitment notice is providing data to online behavioral advertising companies, leaving a potentially identifiable trail [6]. Online behavioral advertising (OBA) is a set of practices that companies engage in to track consumers’ online activities over time to deliver advertisements targeted to their inferred interests [7]. The problem is that many individuals either are unaware of the privacy risks of online activity or consciously accept a trade-off to their privacy [8]. For example, a man with sleep apnea was shocked to be followed by ads for such devices when he visited websites unrelated to the condition [8]. This man’s experience prompted an investigation by the Office of the Privacy Commissioner of Canada (OPC), which revealed that Google’s online advertising service used sensitive personal information about individuals’ online activities to deliver targeted health-related ads, which violates Canadian privacy law [9].

Although regulators like the OPC are mandated to enforce privacy laws, privacy breaches are not uncommon, and there is little guidance for researchers seeking to use social media for research recruitment. There are basic ethical principles, such as Respect for Persons, Concern for Welfare, and Justice, codified in the UN Declaration of Human Rights [10], the Nuremberg Code [11], the Declaration of Helsinki [12], and the Belmont

Report [13]. There are general consensus statements, such as the Tri-Council Policy Statement (TCPS) [14] developed by Canada’s three federal funding agencies, that provide guidance on how to interpret and apply these basic ethical principles. For example, the TCPS explains that Respect for Persons can be achieved through “free, informed, and ongoing consent”; Concern for Welfare can be achieved by “minimizing risks and respecting and maintaining the welfare of participants,” which includes protecting their privacy; and Justice can be achieved by “treating all people fairly and equitably” [14]. In addition, the Ethics Working Committee of the Association of Internet Researchers (AoIR), an international professional association, has produced a set of guiding questions for researchers seeking to use the Internet for research [15]. However, these documents predate the Internet or social media, do not adequately address the unique ethical issues of social media as a recruitment tool, or do not offer practical solutions.

Many forms of Internet-based research could be considered ethically challenging because of the blurred public and private boundaries of online spaces [16], the dynamic and interactive nature of the media [17], and ease with which sensitive data can be accessed, shared, hacked, and/or replicated [18]. Online research recruitment introduces unique ethical issues because it may pose threats to the principles of Respect for Persons and Concern for Welfare in regard to privacy even before the consent to enroll in a study. Privacy is defined as an “individual’s right to be free from intrusion or interference by others” [14]. An important aspect of privacy is the right to control information about oneself. In the context of health research, this means that an individual should have the opportunity to exercise control over personal information by consenting to, or withholding consent for, the collection, use, and/or disclosure of information. Confidentiality, a related but distinct concept, refers to the obligation to “safeguard entrusted information from unauthorized access, use, disclosure, modification, loss, or theft” [14]. We [16], along with a handful of other researchers [19-23], explored the ethical and legal issues related to social media as a source of qualitative data, resulting in some recommendations. There is only one known study that explored the ethical aspects of social media as a recruitment tool. Curtis describes the ethical challenges of social networking and online recruitment for HIV research and concludes with a set of recommended best practices for HIV researchers [6].

Critical dialogue is needed to understand the pertinent ethical issues involved in online health research recruitment and the procedural solutions to protect the rights and safety of potential research participants. In this paper, we describe how we used the Internet and social media to recruit cancer patients and their family caregivers for a focus group study on dietary

self-management behaviors, the ethical concerns raised by our institutional Research Ethics Board (REB), and the privacy-enhancing strategies we developed to address them. We include a critical reflection of the perceived privacy risks associated with our social media recruitment strategy and the appropriateness of the risk mitigation strategies that we employed by assessing their alignment with the principles of Privacy by Design (PbD) [24], a globally recognized standard for the protection of privacy [25]. We conclude by offering a PbD framework for online health research recruitment. While primarily directed at researchers, this framework for achieving PbD in online health research recruitment is intended to support and inform a wide array of stakeholders responsible for making decisions about the ethics of online health research recruitment.

Methods

Overview

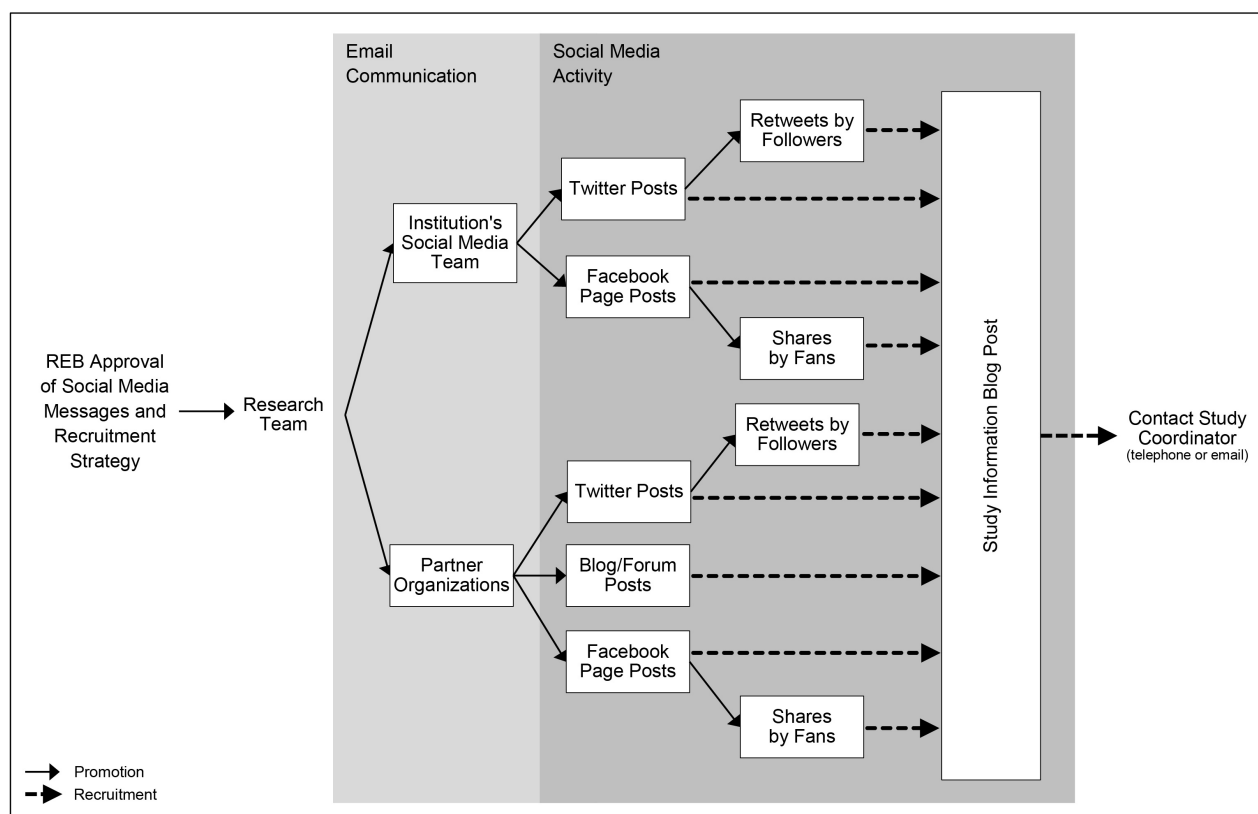
We (JLB and ABC) explored the nutrition and culinary knowledge, attitudes, and behaviors of cancer patients and their family caregivers, and their views on Web-based tools to enhance dietary self-management behaviors. Lack of nutritional

knowledge and culinary skills reduces the likelihood of practicing dietary self-management behaviors [26].

Initially, we relied on traditional recruitment methods, including posters placed at strategic locations (eg, elevators and clinics) in the hospital, in-person recruitment at our cooking and nutrition education classes, and targeted promotion of our study by email to our community partners. Despite this effort, these strategies did not help us reach our recruitment target and composition. Recruitment challenges are a persistent problem faced by researchers. A retrospective review of 404 clinical trials funded by two major funding agencies in the United Kingdom found that only 55% reached their recruitment target [27].

Encouraged by the evidence on the potential effectiveness of social media as a health research recruitment tool [1], we applied for institutional REB approval to use the Internet and social media to recruit study participants. Our social media recruitment strategy was multichannel (see Figure 1). We proposed to publish a recruitment notice on an established blog on our hospital website, share the link to the blog post on our Twitter and Facebook pages over 4 weeks, and ask our social media followers and community partners to share the link with their networks of connections.

Figure 1. Initial social media recruitment strategy.



Ethical Concerns Raised by the Research Ethics Board

Our institutional REB raised concerns about the privacy risks associated with our proposed use of the Internet and social media for research recruitment. Specifically, they were concerned that by clicking on our social media recruitment messages (eg, “Seeking cancer patients for a study of nutrition and cooking”),

individuals may unknowingly add personal and sensitive health information to their online profile, leaving an identifiable trail that may be used and disclosed by marketers.

The REB asked us to revise our social media recruitment strategy with the following questions in mind:

1. How will you inform users about the potential for privacy breaches and their implications?
2. How will you protect users from privacy breaches or inadvertently sharing potentially identifying information about themselves?

Results

Privacy-Enhanced Social Media Recruitment Strategy

Overview

Our revised social media recruitment strategy served to inform users about privacy risks and protect their privacy, while at the same time meeting our recruitment objectives. This *win-win* approach is a fundamental principle of PbD [24].

PbD was developed by the former Information and Privacy Commissioner of Ontario, Canada, Dr Ann Cavoukian in the

late 1990s. It is an overarching framework for embedding privacy and data protection into information technologies, organizational processes, networked architectures, and entire systems of oversight in a credible and effective way [24,28]. It is based on the following seven foundational principles (verbatim): (1) Proactive not Reactive, Preventative not Remedial; (2) Privacy as the Default Setting; (3) Privacy Embedded into Design; (4) Full Functionality—Positive-Sum, not Zero-Sum; (5) End-to-End Security—Full Lifecycle Protection; (6) Visibility and Transparency—Keep it Open; and (7) Respect for User Privacy—Keep it User Centric [24].

In this section, we describe our revised social media recruitment strategy and reflect on the extent to which the privacy-enhancing measures that we used aligned with PbD. The principles of PbD and their descriptions are summarized verbatim in Table 1, along with our assessment of the extent to which our recruitment measures aligned with them.

Table 1. Applying the principles of Privacy by Design [24] to our case study^a.

Principle	Short description	Alignment with Privacy by Design
1.Proactive not Reactive; Preventative not Remedial	PbD ^b seeks to anticipate and prevent privacy-invasive events before they happen. PbD does not wait for privacy risks to materialize nor offer remedies after the fact.	Privacy notices proactively informed users about the privacy risks of social media, but required individuals to take action to protect their privacy. On the other hand, marketing headlines proactively protected individuals' privacy by ensuring that those interested in the study were concealed within a broader population than just those targeted for recruitment. In contrast, editing or removing posts after publication represents a remedial, after-the-fact solution.
2.Privacy as the Default Setting	PbD seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected. No action is required on the part of the individual to protect their privacy. It is built into the system, by default.	We built privacy protection into the recruitment strategy using marketing headlines and a hospital blog with a disabled comment feature to recruit interested individuals. Those that chose to enroll in the study did so through the hospital's private data collection system without tracing back to social media.
3.Privacy Embedded into Design	PbD is embedded into the design and architecture of the system. It is not bolted on as an add-on, after the fact. Privacy is integral to the system, without diminishing functionality.	We embedded privacy into the design of the recruitment strategy using marketing headlines, without diminishing the functionality of social media. On the other hand, we lost functionality that could have enhanced the spread and exposure of our recruitment messages by opting to use a blog with a disabled comment feature and by proposing to edit and delete sensitive posts before publication.
4. Full Functionality —Positive-Sum, not Zero-Sum	PbD seeks to accommodate all legitimate interests and objectives in a positive-sum, win-win manner, not through a dated, zero-sum approach where unnecessary trade-offs are made.	Using marketing headlines is an example of a win-win, privacy-enhancing strategy. It increased the reach of the recruitment strategy (which one would expect to increase enrollment) without compromising privacy. Disabling the comment feature on the hospital blog, on the other hand, is not win-win because we traded function for privacy.
5. End-to-End Security —Full Lifecycle Protection	PbD explains that strong security measures are essential to PbD from start to finish. Embedding PbD into the system prior to the first element of information being collected ensures that all data are securely retained throughout the entire lifecycle of the data involved.	We used social media to garner interest in the research study, embedding privacy protection in the <i>consideration phase</i> well before enrollment, then used the hospital's private and secure data collection system and procedures to protect interested and consenting study participants' privacy and confidentiality from start to finish. However, we could have done a better job explaining our strategy to our community partners to ensure that they used it fully. We had no control over how the public responded to or shared our social media recruitment messages.
6. Visibility and Transparency —Keep it Open	PbD seeks to assure all stakeholders that whatever the business practice or technology involved, it is, in fact, operating according to the stated promises and objectives, subject to independent verification.	Our aim with privacy notices was two-fold: (1) to inform users about privacy risks and their implications; and (2) to be as open and transparent as possible. We also adhered to the procedural practices and requirements set by our governing bodies to protect the rights and safety of potential research participants. This included Research Ethics Board review of the research protocol and approval of all social media posts and privacy notices prior to publication.
7. Respect for User Privacy —Keep it User Centric	PbD requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options.	We were cautious in our use of marketing headlines so as not to risk deceiving people or wasting their time. We used privacy notices to offer users appropriate notice and attempted to design them effectively, but we did not use a user-centered design approach to develop them nor did we test their effectiveness. In addition, we do not know people's views on the marketing headline strategy. Some may have disliked the lack of directness in the notice to get them to the second site.

^aThe principles and their descriptions are described verbatim [24].

^bPbD: Privacy by Design.

A. Inform About Privacy Risks With Privacy Notices

Providing notice and choice about data practices is an essential element of data protection frameworks like PbD [24]. Providing participants with enough information to adequately assess risks and potential benefits associated with their participation in research is a basic requirement of ethical research practice [14]. Privacy notices are a common strategy to make a system's users aware of data practices involving personal information, which is supposed to enable users to make informed decisions [29]. If designed effectively, the notices can function to proactively

alert the user about potential privacy risks and prompt them to take action to protect their privacy. Privacy notices can take many different forms, ranging from a privacy policy on a website, cookie consent notices shown in a banner on a webpage, to consumer warnings or permission notices in pop-up dialog boxes.

We developed privacy notices for the hospital blog and Facebook page and regularly tweeted disclaimers about the privacy risks of Twitter. We also included privacy notices in our email requests to community partners to spread the word

about our research study. Privacy notices were written in plain language [30] and approved by a plain-language expert. Plain language is an evidence-based, patient-centered approach to writing health information. Plain language uses “familiar words,

not jargon; active voice; and a conversational study to convey information clearly” [30]. All privacy notices were reviewed and approved by the REB before posting (see Table 2 and Figure 2).

Table 2. Privacy notices and disclaimers.

Medium	Privacy notice/disclaimer
Email	“Please note that the security of email messages is not guaranteed. Messages may be forged, forwarded, kept indefinitely, or seen by others using the Internet. Do not use email to discuss information you think is sensitive. Do not use email in an emergency since email may be delayed.”
Facebook	“Please also note that the privacy and confidentiality of content (text or pictures) shared on social media platforms is not guaranteed. Content may be forged, forwarded, kept indefinitely, or seen by others using the Internet whether you share publicly to everyone or privately to specific people. Do not use social media to discuss information you think is sensitive. While you may share this information with a select group of people, someone in your networks may share it more widely without your consent.”
Twitter	“The security of social media is not guaranteed. Contact us about the study. Don’t post if concerned about privacy.” ^a

^aPlease note that this tweet focuses on security as a possible threat to privacy if data is leaked. Privacy is not limited to security issues.

Figure 2. Facebook recruitment post with privacy disclaimer.



B. Protect Privacy Using Privacy-Enhanced Social Media Messages

We built privacy protection into our social media recruitment strategy using an Internet marketing approach known as *marketing headlines*. Internet marketing headlines aim to attract traffic by providing just enough information to make the reader curious, but not enough to satisfy their curiosity without clicking through. Marketing headlines are often associated with the less savory marketing practice *clickbait* that aims to trick people into following links online for the purpose of generating ad revenue [31]. In our case, we provided just enough information in our social media recruitment messages to attract the desired population, but not too much information that may cause them to inadvertently disclose personal health information through social media. Our goal was to garner public interest in our research while also attracting study participants.

For example, we originally proposed the following tweet to recruit participants for our study: “Seeking cancer patients for

a study of nutrition and cooking @ELLICSRKitchen [URL].” Upon request by our REB, we removed the term “cancer patient” from all social media posts. The following is an example of a privacy-enhanced tweet: “Does #nutrition matter to you? Tell us what you think about #cooking and #cancer @ELLICSRKitchen [URL].”

This small change accomplished two goals: (1) it broadened the reach of our recruitment strategy by attracting a larger population of social media users; and (2) it protected patients’ privacy by default. Used in this way, marketing headlines is a win-win because we attract more interest in our work while pooling cancer patients we wish to recruit into a broader population of people interested in the subject of our research.

We asked our community partners to use our privacy-enhanced social media messages. All social media messages were reviewed and approved by a plain-language expert and the REB, and were published without modification.

C. Protect Privacy by Disabling Comment Feature or Moderating Comments

All social media messages included a link that directed interested individuals to the study recruitment notice on our hospital blog. At the time of publishing the recruitment notice on our hospital website, comments were not enabled on the blog platform due to hospital policy. Had commenting been enabled, we proposed to moderate any comments before they were made visible on the blog and remove references to potentially identifying or personal health information. While this strategy would have offered privacy protection, it does not represent a win-win because the blog software functionality was diminished to accomplish the privacy objectives. Allowing readers to freely post and share comments on the hospital blog could have generated online discussion about the study, which could have attracted more study participants, and represents a way to engage the public in spreading the word about a research project.

Limitations

First, we used PbD to assess the appropriateness of our revised social media recruitment strategy *after the fact*. We encourage others to proactively use the PbD framework from the outset of the study design. Second, although our privacy notices were designed based on plain-language principles [30] and displayed prominently near the relevant contact information, it is possible that they were not seen or read. The evidence suggests that most privacy notices are not effective at informing consumers [32]. Based on a comprehensive review of research, Schaub et al offer best practices for improving the effectiveness of privacy notices [29]. These include the following: starting with a thorough understanding of a system's information flows and data practices; tailoring notices to different audiences; providing concise, relevant, and actionable information; layering and contextualizing notices (eg, just-in-time notices without too much repetition to avoid habituation); and employing user-centered design to evaluate user attention, comprehension, and recall. Third, our social media recruitment strategy engaged other actors—our community partners and social media followers—to spread our social media recruitment messages. Although we provided our community partners and social media followers with privacy-enhanced social media messages to promote our study, we had no control over whether or how they adopted them or how the public responded to them, which could have resulted in advertent disclosures of personal information. Fourth, we do not know about the public's views on marketing headlines as a research recruitment strategy. Some people may have disliked the lack of directness in the notice. Furthermore, we do not know the impact of this recruitment strategy on people who clicked but were ineligible to participate, and we do not know if eligible people found the recruitment strategy acceptable or if it negatively influenced their willingness to participate. Lastly, interested participants were required to contact the study coordinator by phone or email. A secure email form on the hospital webpage would have offered more privacy protection. Using a form controlled by our hospital server would ensure that the message was delivered to the intended recipient, with fewer chances of it being hacked from server to server, and that the message content is not scanned for keywords to trigger targeted ads, as is this case with Google email client [33].

Discussion

Were the Privacy Measures Appropriate?

Overview

Guidelines for the ethical conduct of human subject research state that risk mitigation strategies should be proportional to the magnitude and probability of risks involved [14]. Known as the proportionate approach, the level of risk posed by the research is used to determine “the level of review (eg, delegated or full board review), the approach to the actual review of the research itself, and the risk mitigation strategies required to protect the rights and safety of research subjects” [14]. This means that “the most intensive scrutiny, time, resources, and correspondingly, protection should be applied to the most ethically challenging research.” Similarly, PbD suggests that the strength of the privacy measure should match the sensitivity of the data [24].

We reflect on the perceived privacy risks associated with our social media recruitment strategy and the appropriateness of the risk mitigation strategies that we employed by discussing the following: (1) What are the potential risks and who is at risk? (2) Is cancer considered sensitive personal information? (3) What is the probability of online disclosure of a cancer diagnosis in everyday life? and (4) What are the public's expectations for privacy online and their views about online tracking, profiling, and targeting?

A. What Are the Potential Risks and Who Is at Risk?

The primary risk associated with our recruitment strategy was the potential harm that a person may experience from the disclosure, collection, and use of personal and sensitive information—in this case a diagnosis of cancer—triggered by clicking on our social media recruitment messages. Potential harms associated with disclosure of health information like a cancer diagnosis could include stigmatization, discrimination, or damage to reputation, and may negatively affect relationships, job opportunities, and insurance options. However, we cannot assume that a person clicking on the recruitment message would experience these harms. What we do know is that they will likely receive advertising messages about cancer and/or eating well. It is possible that seeing such messages could be personally troubling for them, but we do not know if this is the case.

It is worth mentioning that there are documented cases of health data located in big data repositories or biobanks being repurposed by third parties for legal and security purposes. These unintended secondary uses of health data have included forensic investigations, civil lawsuits, border security, and identification of victims in mass casualty events [34]. For a thorough discussion of documented and hypothetical secondary uses of online health data collections, see O'Doherty et al [34].

In terms of who is at risk, it cannot be assumed that the person clicking on the recruitment message was revealing information about himself or herself at all. Spouses, children, siblings, other family members, and friends play a vital role in searching for health information. Research conducted by the Pew Research Center indicates that half of online health information research

is on behalf of someone else [35]. Furthermore, as our recruitment blog post explained, we were seeking cancer patients and their family caregivers to participate in our research. Hence, if a caregiver clicked on our social media recruitment message, he or she would not have been revealing information about himself/herself, which was presumed to be the case in the ethics review. However, these individuals could have still received advertising messages about cancer and/or eating well, which they may or may not have found troubling.

B. Is Cancer “Sensitive” Personal Information, Requiring More Privacy Protection?

The Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) defines personal information as “information about an identifiable individual” [36]. The OPC takes the position that information involved in online tracking and targeting constitutes personal information [7]. Principle 3 of PIPEDA states that “the knowledge and consent of the individual are required for the collection, use, or disclosure of personal information, except where inappropriate” [36]. Furthermore, the privacy act goes on to explain that “organizations must obtain an individual’s consent for all disclosures of their personal information to any third party unless one of PIPEDA’s exceptions to consent can be applied.” The magnitude or seriousness of harms associated with the disclosure of personal information depends, in part, on whether the information is considered “sensitive.” According to PIPEDA, some information is almost always considered to be sensitive (eg, medical records and income records); however, any information can be sensitive depending on the context. The US National Committee on Vital and Health Statistics (NCVHS) has done some further work defining sensitive health information, which they explain carries unusually high risks in the event of disclosure [37]. Based on public consultations and expert deliberation, categories of health information considered sensitive by the NCVHS include those related to domestic violence, genetics, mental health, reproductive care, and substance abuse [37]—not cancer. That being said, as explained in PIPEDA, sensitivity is subjective and depends on the individual’s circumstances, and the context in which the information is shared [36]. The Google health ads case is evidence of this as sleep apnea was considered sensitive personal information [9]. The main point is that health information is considered sensitive personal information, but within health information, there are gradients of sensitivity and cancer may be considered less sensitive personal health information.

C. What Is the Probability of the Risks and Harms Occurring in Everyday Life?

It is highly probable that cancer patients who clicked on our social media recruitment messages already disclosed their cancer diagnosis online, thereby exposing themselves to related harms. First, the majority of cancer patients report using the Internet as a source of health information. For example, 86% of a sample of 202 thyroid cancer patients [38] surveyed from the same hospital where this study was conducted, and 68% of a sample of 824 Canadian prostate cancer patients [39], reported using the Internet to search for information related to their cancer. As people spend more time online, they leave a digital trail. Second,

given the scope and scale of information collected by third-party advertisers and the sophisticated means of collecting and analyzing disparate pieces of data [7], it is reasonable to assume that Internet search queries about cancer could be linked to an individual. Typical information collected in Internet log files includes the following: IP address, pages visited, length of time spent on pages, advertisements viewed, articles read, purchases made, search terms or other information entered on a site, user preferences such as language, operating system, and geographical location [8]. Additional data may be gathered from social networking sites where individuals volunteer significant amounts of personal information. Third, we used our departmental Facebook and Twitter pages to promote our research study to our social media followers. The people who follow us on social media have likely already “disclosed” to third-party trackers that, at the very least, they are interested in cancer by choosing to follow a social media account affiliated with a cancer center. Therefore, using our marketing headline approach, we would not subject cancer patients to disclose more than an interest in cancer, which they likely have already provided online.

D. What Are the Public’s Privacy Expectations and Views on Online Behavioral Advertising?

A total of 90% of Canadians are concerned about the privacy impact of new technologies and 98% want strong privacy laws [8]. People between the ages of 45 and 65 years are more likely to express high levels of concern about the privacy impact of new technologies than those 25 and under [34]. However, teenage social media users seem to care more about online privacy when it comes to their personal health information. Motivated by a need for self-protection as a chronically ill patient and self-definition as a regular teenager, a qualitative study (N=20) revealed that Canadian teenagers (12-18 years old) with a chronic illness were selective about sharing personal health information on social media [40]. In general, teenagers are less concerned about the collection of personal information by governments and companies, but very concerned about their social privacy, or having control over the content of their interactions with others [40]. When it comes to OBA, 50% of Canadians surveyed in 2009 were “somewhat uncomfortable” with tracking-based advertising [8]. However, a 2011 report by KPMG consulting firm revealed that 46% of Canadians were “somewhat willing” to have their online usage tracked by advertisers if there are benefits [41]. Benefits of OBA for the consumer include free online content, more relevant advertising, and enhanced browsing experience [41]. A population-based telephone survey of Americans suggests that consumers would be more willing to accept OBA if there was more transparency, consumer choice, and data retention limits [42]. Complicating a clear understanding of the public’s views and expectations with regard to online privacy is the well-known privacy paradox [43]. Most people would say that they care about their online privacy but do not act on that concern, revealing increasing amounts of personal information that can be used and disclosed by governments and marketers [43].

Disparate Norms Within and Across Disciplines and Research Ethics Boards

A further challenge for researchers seeking to use the Internet and social media for research recruitment is the disparate norms about what is and what is not ethical across research communities. Researchers are guided by different disciplinary methodological approaches, norms, and conventions, and regulations for ethical online research vary across disciplines. What is considered ethically acceptable in one discipline may not be in another [44]. The same holds true for different REBs. Moreover, the same REB may reach different conclusions about the same technological approach across studies. Nebeker and colleagues show that visual imaging and location-tracking devices (eg, Global Positioning System) are reviewed inconsistently in one institution [45]. While research plans incorporated consistent descriptions of each device and associated potential risks, REB letters revealed inconsistent perceptions of potential study risks associated with the collocation of location data should a data breach occur [45]. Inconsistent perceptions about the potential risks involved in research that uses new technologies like social media make the REB protocol development and review process challenging for researchers. However, researchers are not the only ones grappling with the unique ethical issues of online research. REBs may be unfamiliar with these new technologies, prompting confusion about what actions are necessary and appropriate to effectively evaluate and mitigate potential risks. Furthermore, there may also be some differences in where different REBs draw the line between participant autonomy versus participant protection.

Privacy by Design for Online Health Research Recruitment

We have shown that PbD is a useful framework for designing, evaluating, and achieving privacy in online health research recruitment. Applying the principles of PbD helped to identify the privacy strengths, weaknesses, and gaps in our recruitment strategy. Based on alignment with PbD principles, use of marketing headlines was the strongest privacy measure used whereas privacy notices were the weakest. Contrary to the principles of PbD, we made trade-offs in favor of privacy protection, such as agreeing to disable the comment functionality on the hospital blog, which traded function for privacy. PbD also alerted us to areas in need of improvement, such as the privacy gaps created by engaging others in implementing our recruitment strategy. To fully embed privacy into the design of a recruitment strategy, all parties involved in implementing it should endorse the PbD approach.

By applying PbD, we also identified areas in need of further research. While PbD is becoming the standard for privacy protection in many jurisdictions around the world [28], there is little practical guidance on how to apply the seven foundation principles [46]. For example, transparency and empowering

user-centered options are key principles of PbD, but the framework provides little practical guidance on how to effectively design privacy notices using these principles. Schaub et al's compilation of best practices for privacy notices is an excellent complementary resource in this regard [29]. As a first step, we need a better understanding of the public's views on the privacy risks of online health research recruitment and Web-based research, including the probability and magnitude of harm as well as what privacy protection would be appropriate or may create potential barriers to access. In parallel, further research is needed to understand how to effectively design strong privacy defaults, appropriate notice mechanisms, and empowering options, and to examine the impact of these privacy measures on the public's online behaviors, including participation in health research studies. In this study, we did not consider informing users about the various strategies to protect their online privacy, but we think this is important. Future research should consider designing and evaluating educational efforts to teach patients and their families about these strategies. These strategies include the following: clearing your Internet browser history (eg, cookies); installing Internet browser extensions that block ads, or that reveal and block the websites that track your browser history; or using InPrivate Browsing to stop the computer from tracking your website history.

Another privacy tool, Privacy Impact Assessment (PIA), deserves mention. PIAs aim to "identify the potential privacy risks of new or redesigned programs and to eliminate or reduce those risks to an acceptable level" [46]. They are generally used to ensure that an organization is complying with legislative and regulatory requirements. PIAs may be useful tools to consider for the assessment of the privacy risks of an online health research recruitment strategy. However, typical PIAs are not grounded in the PbD framework and they do not provide overarching principles to guide the design and implementation of privacy protection. Jeselon and Fineberg recommend using the PbD framework to augment PIAs to achieve a more holistic approach to privacy protection and offer practical guidance on how to apply PbD to PIAs [47].

Based on our experiences with this case study, we offer a PbD framework for online health research recruitment. We drew on the principles of PbD [24] and examples of its application [28,47], as well as recommendations from the AoIR [15] and the Secretary's Advisory Committee on Human Research Protections [18]. In this framework, we offer a set of privacy questions and considerations to guide the ethical design and conduct of studies that use the Internet and social media as a health research recruitment tool. We describe the principles, guiding questions, and application considerations of this framework in Table 3. The PbD principles are verbatim. We have drawn on recommendations from cited sources to aid the reader in their application.

Table 3. Privacy by Design framework for online health research recruitment: Proposed considerations for researchers and institutional Research Ethics Boards.

Privacy-by-Design principles	Considerations
A. Consider the nature of the study, the target population, and the sensitivity of the data	
Justification	Why is it necessary to use the Internet and social media to recruit participants for your research project?
Context	Where does the study recruitment take place? What are the terms of use and privacy policies of the recruitment sites or applications? What are users' privacy expectations regarding the recruitment sites or applications?
Sensitivity	What is the subject of study? Is the data considered personal information? Is the data considered "sensitive" personal information? What are the privacy expectations commonly associated with these types of data?
Vulnerability	Who are the recruitment targets? What additional privacy measures may be required to protect the privacy of vulnerable individuals?
B. Apply Privacy by Design [24]	
Proactive not Reactive; Preventative not Remedial	<p>What are the potential privacy risks and related harms associated with the recruitment strategy? Do certain data, people, or groups require more privacy protection?</p> <p>PbD^a Application: Anticipate and prevent privacy-invasive events before they happen—before individuals are even exposed to the recruitment strategy—as opposed to offering remedies for resolving privacy breaches once they have occurred [24]. Adopt and implement strong privacy practices early and continuously, and use systematic methods to recognize and correct weakening links, privacy measures, or data protection practices before privacy risks occur [28].</p>
Privacy as the Default Setting	<p>If an individual does nothing, is their privacy still intact when they are exposed to the recruitment strategy or do they have to take action (eg, opt out or add a privacy measure) to protect their privacy?</p> <p>PbD Application: Aim to deliver the maximum degree of privacy by ensuring that personal data are automatically protected without the individual having to do anything to protect their privacy [47]. Keep the collection of personal information to a minimum, justify additional data collection on a data-by-data basis, and use default settings of technologies that offer the most privacy protection [28].</p>
Privacy Embedded into Design	<p>Is your privacy-enhancing measure built into the design of your recruitment strategy or has it been bolted on as an add-on, after the fact?</p> <p>PbD Application: Make privacy a core component of your recruitment strategy from the outset of the study, so that it is an essential component of the study design [47]. Embed privacy into the recruitment technologies, operations, and information architectures in a holistic, integrative way [28].</p>
Full Functionality—Positive-Sum not Zero-Sum	<p>Does your recruitment strategy offer privacy protection without sacrificing your recruitment goals and objectives?</p> <p>PbD Application: Consider all legitimate interests and objectives of the recruitment strategy and aim to accommodate them in optimal ways to ensure the individual's privacy is protected without any unnecessary trade-offs between privacy and functionality, security, or your recruitment goals [47]. Select privacy-enhancing measures that help to achieve your recruitment goals, maintaining full functionality and full security while protecting privacy [28].</p>
End-to-End Security—Full Lifecycle Protection	<p>Are there any weak links or gaps in the implementation or oversight of your recruitment strategy?</p> <p>PbD Application: Consider how information, particularly personal information, will flow and be accessed, and by whom, throughout the entire lifecycle of the study. Embed privacy-enhancing measures and data security measures into the recruitment strategy before the first element of data is collected by you as a researcher or by third parties, and extend that security in a comprehensive and systematic manner throughout the entire lifecycle of the data involved [28].</p>
Visibility and Transparency—Keep it Open	<p>Are all people and organizations involved in recruiting participants (directly or indirectly) operating according to stated promises and objectives, and is information about their privacy policies and practices readily available to the public?</p> <p>PbD Application: Ensure that all recruitment actors are operating according to their stated privacy practices (eg, policies and procedures related to the collection, use, and storage of personal information) and that these are made visible and transparent to enable users to make an informed choice about whether to participate in the study or not [47]. When sharing study information with collaborators or third parties, ensure that they use equivalent data protection measures through contractual processes or others means [28].</p>

Privacy-by-Design principles	Considerations
Respect for User Privacy—Keep it User Centric	<p>Are your privacy measures user centric? Have they been designed with the user in mind? Are they simple to use and written in easy-to-understand plain language? Have they been tested and approved by users?</p> <p>PbD Application: Respect for User Privacy is at the heart of PbD. Use user-centered and empowering, user-friendly, privacy-enhancing recruitment technologies, policies, and procedures so that individuals can exercise their privacy rights and make informed privacy decisions [28]. As explained by Dr Cavoukian, “The most privacy-enhancing solutions and results are usually those that are consciously designed around the interests, needs, and expectations of individuals and users, who typically have the greatest vested interest in the management of their personal data by others” [28].</p>

^aPbD: Privacy by Design.

Conclusions

Researchers, REBs, ethicists, students, and potential study participants are often unaware of the privacy risks of Internet and social media health research recruitment and there is no official guidance. From this case study, some may conclude that the REB's perceptions of the potential risks involved in our research study and our revised privacy-enhanced recruitment strategy did not match the magnitude and probability of the risks involved. On the other hand, others may argue that given that hospitals occupy an important trust relationship with

patients and the public, hospital REBs should apply the precautionary principle as their use of social media may provide a false sense of security. We have shown that PbD is a useful framework for designing, evaluating, and achieving privacy in Web-based research recruitment. We offer our PbD framework for online health research recruitment for researchers and REBs to guide the ethical design, review, and conduct of studies that use the Internet and social media as a health research recruitment tool. Future research should focus on designing effective privacy notices and measures and evaluating their impact.

Conflicts of Interest

None declared.

References

1. Alshaikh F, Ramzan F, Rawaf S, Majeed A. Social network sites as a mode to collect health data: A systematic review. *J Med Internet Res* 2014;16(7):e171 [FREE Full text] [doi: [10.2196/jmir.3050](https://doi.org/10.2196/jmir.3050)] [Medline: [25048247](https://pubmed.ncbi.nlm.nih.gov/25048247/)]
2. Reaves AC, Bianchi DW. The role of social networking sites in medical genetics research. *Am J Med Genet A* 2013 May;161A(5):951-957. [doi: [10.1002/ajmg.a.35903](https://doi.org/10.1002/ajmg.a.35903)] [Medline: [23554131](https://pubmed.ncbi.nlm.nih.gov/23554131/)]
3. 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](https://doi.org/10.1111/jnu.12032)] [Medline: [23676115](https://pubmed.ncbi.nlm.nih.gov/23676115/)]
4. Topolovec-Vranic J, Natarajan K. The use of social media in recruitment for medical research studies: A scoping review. *J Med Internet Res* 2016 Nov 7;18(11):e286 [FREE Full text] [doi: [10.2196/jmir.5698](https://doi.org/10.2196/jmir.5698)] [Medline: [27821383](https://pubmed.ncbi.nlm.nih.gov/27821383/)]
5. Frandsen M, Walters J, Ferguson SG. Exploring the viability of using online social media advertising as a recruitment method for smoking cessation clinical trials. *Nicotine Tob Res* 2014 Feb;16(2):247-251. [doi: [10.1093/ntr/ntt157](https://doi.org/10.1093/ntr/ntt157)] [Medline: [24127266](https://pubmed.ncbi.nlm.nih.gov/24127266/)]
6. Curtis BL. Social networking and online recruiting for HIV research: Ethical challenges. *J Empir Res Hum Res Ethics* 2014 Feb;9(1):58-70 [FREE Full text] [doi: [10.1525/jer.2014.9.1.58](https://doi.org/10.1525/jer.2014.9.1.58)] [Medline: [24572084](https://pubmed.ncbi.nlm.nih.gov/24572084/)]
7. Office of the Privacy Commissioner of Canada. 2015 Dec. Policy position on online behavioral advertising URL: https://www.priv.gc.ca/en/privacy-topics/advertising-and-marketing/behaviouraltargeted-advertising/bg_ba_1206/ [accessed 2017-03-10] [WebCite Cache ID 6ormML8Ou]
8. Office of the Privacy Commissioner of Canada. 2011 May. Report on the 2010 Office of the Privacy Commissioner of Canada's consultations on online tracking, profiling and targeting, and cloud computing URL: https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/report_201105/ [accessed 2017-03-10] [WebCite Cache ID 6ormEdYo3]
9. Office of the Privacy Commissioner of Canada. 2015 Jan 15. News Release: Google ads sparked by Web surfing on health sites violate privacy right, investigation finds URL: https://www.priv.gc.ca/en/opc-news/news-and-announcements/2014/nr-c_140115/ [accessed 2016-11-21] [WebCite Cache ID 6mBy6DZ42]
10. The Universal Declaration of Human Rights. New York, NY: The United Nations; 1948. URL: http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf [accessed 2017-03-30] [WebCite Cache ID 6pLuJcbrX]
11. Trials of War Criminals Before the Nuremberg Military Tribunals Under Control Council Law, No. 10. Washington, DC: US Government Printing Office; 1949. The Nuremberg Code URL: <https://history.nih.gov/research/downloads/nuremberg.pdf> [accessed 2016-11-21] [WebCite Cache ID 6mC1LwkWw]
12. World Medical Association. Declaration of Helsinki: Ethical Principles for Research Involving Human Subjects. *JAMA*; 2013 Nov 27. URL: <http://jamanetwork.com/journals/jama/fullarticle/1760318> [accessed 2017-03-30] [WebCite Cache ID 6pLw1RFqP]

13. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Office for Human Research Protections. 1979 Apr 18. The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research URL: <http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/> [accessed 2016-11-21] [WebCite Cache ID 6mBzBGsjB]
14. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Ottawa, CA: Government of Canada; 2014. URL: http://www.pre.ethics.gc.ca/pdf/eng/tcps2-2014/TCPS_2_FINAL_Web.pdf [accessed 2016-11-21] [WebCite Cache ID 6pLxVh5K7]
15. Markham A, Buchanan E. Ethical Decision-Making and Internet Research: Recommendations From the AoIR Ethics Working Committee (Version 2.0). Chicago, IL: Association of Internet Researchers; 2012. URL: <https://aoir.org/reports/ethics2.pdf> [accessed 2017-03-30] [WebCite Cache ID 6pLxgDwsU]
16. Bender J, Norman C, Jadad A. Negotiating consent in the Facebook era: Insights from research on online health communities with youth. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW): Association of Computing Machinery (ACM); 2010 Presented at: ACM Conference on Computer Supported Cooperative Work (CSCW); February 6-10, 2010; Savannah, GA.
17. O'Grady L, Witteman H, Bender JL, Urowitz S, Wiljer D, Jadad AR. Measuring the impact of a moving target: Towards a dynamic framework for evaluating collaborative adaptive interactive technologies. *J Med Internet Res* 2009;11(2):e20 [FREE Full text] [doi: [10.2196/jmir.1058](https://doi.org/10.2196/jmir.1058)] [Medline: [19632973](https://pubmed.ncbi.nlm.nih.gov/19632973/)]
18. Office for Human Research Protections. Considerations and Recommendations Concerning Internet Research and Human Subject Research Regulations, With Revisions. Washington, DC: US Department of Health and Human Services; 2013 Mar. URL: https://www.hhs.gov/ohrp/sites/default/files/ohrp/sachrp/mtgings/2013%20March%20Mtg/internet_research.pdf [accessed 2017-03-30] [WebCite Cache ID 6orl7v111]
19. Moreno MA, Fost NC, Christakis DA. Research ethics in the MySpace era. *Pediatrics* 2008 Jan;121(1):157-161. [doi: [10.1542/peds.2007-3015](https://doi.org/10.1542/peds.2007-3015)] [Medline: [18166570](https://pubmed.ncbi.nlm.nih.gov/18166570/)]
20. Flicker S, Haans D, Skinner H. Ethical dilemmas in research on Internet communities. *Qual Health Res* 2004 Jan;14(1):124-134. [doi: [10.1177/1049732303259842](https://doi.org/10.1177/1049732303259842)] [Medline: [14725180](https://pubmed.ncbi.nlm.nih.gov/14725180/)]
21. Eysenbach G, Till JE. Ethical issues in qualitative research on Internet communities. *BMJ* 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: [11701577](https://pubmed.ncbi.nlm.nih.gov/11701577/)]
22. Bond CS, Ahmed OH, Hind M, Thomas B, Hewitt-Taylor J. The conceptual and practical ethical dilemmas of using health discussion board posts as research data. *J Med Internet Res* 2013;15(6):e112 [FREE Full text] [doi: [10.2196/jmir.2435](https://doi.org/10.2196/jmir.2435)] [Medline: [23748147](https://pubmed.ncbi.nlm.nih.gov/23748147/)]
23. Zimmer M. "But the data is already public": On the ethics of research in Facebook. *Ethics Inf Technol* 2010;12(4):313-325 [FREE Full text] [doi: [10.1007/s10676-010-9227-5](https://doi.org/10.1007/s10676-010-9227-5)]
24. Cavoukian A. Privacy by Design: The 7 Foundational Principles. Toronto, ON: Information and Privacy Commissioner of Ontario; 2011 Jan. URL: <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf> [accessed 2016-11-21] [WebCite Cache ID 6mC0MzUgQ]
25. Unknown. Resolution on privacy by design. In: Proceedings of the 32nd International Conference on Data Protection and Privacy Commissioners. 2010 Presented at: 32nd International Conference on Data Protection and Privacy Commissioners; October 27-29, 2010; Jerusalem, Israel URL: https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Cooperation/Conference_int/10-10-27_Jerusalem_Resolutionon_PrivacybyDesign_EN.pdf
26. Garcia AL, Vargas E, Lam PS, Shennan DB, Smith F, Parrett A. Evaluation of a cooking skills programme in parents of young children--A longitudinal study. *Public Health Nutr* 2014 May;17(5):1013-1021. [doi: [10.1017/S1368980013000165](https://doi.org/10.1017/S1368980013000165)] [Medline: [23402548](https://pubmed.ncbi.nlm.nih.gov/23402548/)]
27. Sully B, Julious S, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: A review of trials funded by two UK funding agencies. *Trials* 2013;14:166. [doi: [10.1186/1745-6215-14-166](https://doi.org/10.1186/1745-6215-14-166)]
28. Cavoukian A. Privacy by Design: Leadership, methods, and results. In: Gutwirth S, Leenes R, de Hert P, Pouillet Y, editors. *European Data Protection: Coming of Age*. Dordrecht, the Netherlands: Springer Netherlands; 2013:175-202.
29. Schaub F, Balebako R, Durity A, Cranor L. A design space for effective privacy notices. In: Proceedings of the Symposium on Usable Privacy and Security (SOUPS). 2015 Presented at: Symposium on Usable Privacy and Security (SOUPS); June 22-25, 2015; Ottawa, ON URL: https://www.ftc.gov/system/files/documents/public_comments/2015/10/00038-97832.pdf
30. Wizowski L, Harper T, Hutchings T. Writing Health Information for Patients and Families: A Guide to Developing Patient Education Materials That Promote Health Literacy. 4th edition. Hamilton, ON: Hamilton Health Sciences; 2014.
31. English Oxford Living Dictionaries. Definition of clickbait URL: <https://en.oxforddictionaries.com/definition/clickbait> [accessed 2017-03-30] [WebCite Cache ID 6pLzbudsq]
32. Cranor L. Necessary but not sufficient: Standardizing mechanisms for privacy notice and choice. *J Telecommun High Technol Law* 2012;10(2):273-307 [FREE Full text]
33. Schofield J. The Guardian. 2013 Aug 15. Is Gmail secure enough for my private emails? URL: <https://www.theguardian.com/technology/askjack/2013/aug/15/gmail-google-email-privacy> [accessed 2016-11-21] [WebCite Cache ID 6mC1XIQ6G]

34. O'Doherty KC, Christofides E, Yen J, Bentzen H, Burke W, Hallowell N, et al. If you build it, they will come: Unintended future uses of organised health data collections. *BMC Med Ethics* 2016 Sep 06;17(1):54 [FREE Full text] [doi: [10.1186/s12910-016-0137-x](https://doi.org/10.1186/s12910-016-0137-x)] [Medline: [27600117](#)]
35. Fox S, Duggan M. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013 Dec 16. Health fact sheet URL: <http://www.pewinternet.org/fact-sheets/health-fact-sheet/> [accessed 2016-11-21] [WebCite Cache ID [6mC27KEaF](#)]
36. Personal Information Protection and Electronic Documents Act. Ottawa, ON: Minister of Justice; 2015 Jun 23. URL: <http://laws-lois.justice.gc.ca/PDF/P-8.6.pdf> [accessed 2016-11-21] [WebCite Cache ID [6pMOK7XWQ](#)]
37. Recommendations Regarding Sensitive Health Information. Hyattsville, MD: National Committee on Vital and Health Statistics; 2010. URL: <http://www.ncvhs.hhs.gov/wp-content/uploads/2014/05/101110lt.pdf> [accessed 2016-11-21] [WebCite Cache ID [6mC2XDJee](#)]
38. Bender JL, Wiljer D, Sawka AM, Tsang R, Alkazaz N, Brierley JD. Thyroid cancer survivors' perceptions of survivorship care follow-up options: A cross-sectional, mixed-methods survey. *Support Care Cancer* 2016 May;24(5):2007-2015. [doi: [10.1007/s00520-015-2981-5](https://doi.org/10.1007/s00520-015-2981-5)] [Medline: [26530226](#)]
39. Bender JL, Feldman-Stewart D, Tong C, Pai H, Au JWY, Brundage MD, et al. Prostate cancer patients information technology use, preferences and needs: A cross-sectional population based survey. In: Proceedings of the Society of Behavioral Medicine 38th Annual Meeting and Scientific Sessions. 2015 Presented at: Society of Behavioral Medicine 38th Annual Meeting and Scientific Sessions; March 29-April 1, 2017; San Diego, CA.
40. van der Velden M, El Emam K. "Not all my friends need to know": A qualitative study of teenage patients, privacy, and social media. *J Am Med Inform Assoc* 2013 Jan 01;20(1):16-24 [FREE Full text] [doi: [10.1136/amiajnl-2012-000949](https://doi.org/10.1136/amiajnl-2012-000949)] [Medline: [22771531](#)]
41. The Converged Lifestyle: Consumers and Convergence 5.: KPMG China; 2012. URL: <https://assets.kpmg.com/content/dam/kpmg/pdf/2012/01/Converged-lifestyle-O-201112.pdf> [accessed 2016-11-21] [WebCite Cache ID [6ormu76VI](#)]
42. Turow JM, King J, Hoofnagle C, Bleakley A, Hennessey M. Social Science Research Network (SSRN). 2009 Sep 29. Americans reject tailored advertising and the three activities that enable it URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214 [accessed 2016-11-21] [WebCite Cache ID [6mC30dfoi](#)]
43. Radin TJ. The privacy paradox: E-commerce and personal information on the Internet. *Bus Prof Ethics J* 2001;20(3/4):145-170.
44. Bos N, Karahalios K, Musgrove-Chávez M, Poole E, Thomas J, Yardi S. Research ethics in the Facebook era: Privacy, anonymity, and oversight. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI). 2009 Presented at: Conference on Human Factors in Computing Systems (CHI); April 4-9, 2009; Boston, MA.
45. Nebeker C, Linares-Orozco R, Crist K. A multi-case study of research using mobile imaging, sensing and tracking technologies to objectively measure behavior. *J Res Adm* 2015;46(1):118-137.
46. Office of the Privacy Commissioner of Canada. 2016 Nov 01. Privacy impact assessments URL: <https://www.priv.gc.ca/en/privacy-topics/privacy-impact-assessments/> [accessed 2017-03-10] [WebCite Cache ID [6orqSj8Mo](#)]
47. Jeselon P, Fineberg A. A Foundational Framework for a Privacy by Design Privacy Impact Assessment. Canada; 2011 Nov 01. URL: <http://10stappen.nl/files/2012/02/8686D0C8638A9A7E6F40F0BE708046BE.pdf> [accessed 2017-03-10] [WebCite Cache ID [6orp5Upls](#)]

Abbreviations

AoIR: Association of Internet Researchers

CHEO: Children's Hospital of Eastern Ontario

ELLICSR: Electronic Living Laboratory for Interdisciplinary Cancer Survivorship Research

NCVHS: National Committee on Vital and Health Statistics

OBA: online behavioral advertising

OPC: Office of the Privacy Commissioner of Canada

PbD: Privacy by Design

PIA: Privacy Impact Assessment

PIPEDA: Personal Information Protection and Electronic Documents Act

REB: Research Ethics Board

TCPS: Tri-Council Policy Statement

Edited by G Eysenbach; submitted 22.11.16; peer-reviewed by A Cavoukian, C Nebeker, K Devon, M Morrison; comments to author 04.01.17; revised version received 10.03.17; accepted 16.03.17; published 06.04.17.

Please cite as:

Bender JL, Cyr AB, Arbuckle L, Ferris LE

Ethics and Privacy Implications of Using the Internet and Social Media to Recruit Participants for Health Research: A Privacy-by-Design Framework for Online Recruitment

J Med Internet Res 2017;19(4):e104

URL: <http://www.jmir.org/2017/4/e104/>

doi: [10.2196/jmir.7029](https://doi.org/10.2196/jmir.7029)

PMID: [28385682](https://pubmed.ncbi.nlm.nih.gov/28385682/)

©Jacqueline Lorene Bender, Alaina B Cyr, Luk Arbuckle, Lorraine E Ferris. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 06.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

A Content Analysis of Health and Safety Communications Among Internet-Based Sex Work Advertisements: Important Information for Public Health

Julie Kille^{1*}, RN, MSN; Vicky Bungay^{2*}, RN, PhD; John Oliffe^{3*}, RN, MEd, PhD; Chris Atchison⁴, MA

¹Providence Health Care, Vancouver, BC, Canada

²School of Nursing, University of British Columbia, Vancouver, BC, Canada

³University of British Columbia, Vancouver, BC, Canada

⁴University of Victoria, Victoria, BC, Canada

* these authors contributed equally

Corresponding Author:

Vicky Bungay, RN, PhD

School of Nursing

University of British Columbia

111-2176 Health Sciences Mall

Vancouver, BC, V6T1Z3

Canada

Phone: 1 6048227933

Fax: 1 6048227466

Email: vicky.bungay@ubc.ca

Abstract

Background: The capacity to advertise via the Internet continues to contribute to the shifting dynamics in adult commercial sex work. eHealth interventions have shown promise to promote Internet-based sex workers' health and safety internationally, yet minimal attention has been paid in Canada to developing such interventions. Understanding the information communicated in Internet-based sex work advertisements is a critical step in knowledge development to inform such interventions.

Objective: The purpose of this content analysis was to increase our understanding of the health and safety information within the Internet advertisements among women, men, and transgender sex workers and to describe how this information may be utilized to inform eHealth service development for this population.

Methods: A total of 75 Internet-based sex worker advertisements (45 women, 24 men, and 6 transgender persons) were purposefully selected from 226 advertisements collected as part of a larger study in Western Canada. Content analysis was employed to guide data extraction about demographic characteristics, sexual services provided, service restrictions, health practices and concerns, safety and security, and business practices. Frequencies for each variable were calculated and further classified by gender. Thematic analysis was then undertaken to situate the communications within the social and commercialized contexts of the sex industry.

Results: Four communications themes were identified: (1) demographic characteristics; (2) sexual services; (3) health; and (4) safety and security. White was the most common ethnicity (46/75, 61%) of advertisements. It was found that 20-29 years of age accounted for 32 of the 51 advertisements that provided age. Escort, the only legal business title, was the most common role title used (48/75, 64%). In total, 85% (64/75) of advertisements detailed lists of sexual services provided and 41% (31/75) of advertisements noted never offering uncovered services (ie, no condom). Gender and the type of Web-based platform mattered for information communicated. It was found that 35 of the 45 women's advertisements were situated in personal websites and hosted details about nonsexual aspects of an appointment. Men and transworkers used Internet classified advertisement platforms with predetermined categories. Communications about sexually transmitted infections (STIs) occurred in only 16% (12/75) of advertisements with men accounting for 7. Women's advertisements accounted for 26 of the 37 advertisements noting safety restrictions. Zero men or transpersons restricted alcohol or drug use. In total, 75% (56/75) of advertisements offered out-call services and the average minimal hourly rate ranged from Can \$140/h to Can \$200/h.

Conclusions: The study findings contribute to understandings about the diverse platforms used in commercial sex advertisements, and how sex workers frame information for potential clients. This information affords health care providers and policy makers insights to how they might assist with promoting the health of Internet-based sex workers and their clients.

(*J Med Internet Res* 2017;19(4):e111) doi:[10.2196/jmir.6746](https://doi.org/10.2196/jmir.6746)

KEYWORDS

eHealth; communication; confidentiality; cross sectional studies; gender; health behavior; Internet; sex industry; sexual health

Introduction

Background

The capacity to advertise via the Internet continues to contribute to the shifting dynamics in adult commercial sex work [1,2] creating unprecedented opportunities for sex workers to determine how they will conduct business and allowing for greater control over their work [3]. Health-related benefits associated with Internet-based sex work, defined as the use of Internet by adults to facilitate consensual exchange of sexual services for money, include increased income, autonomy, and safety when compared with street-level sex work or working for a third party [1,4-6]. Internet-based sex workers can also, however, experience depression, isolation, violence, sexually transmitted infections (STIs), and substantial barriers in accessing and receiving health services [2,7-13].

Internationally, eHealth interventions have shown promise to promote health and safety among Internet-based sex workers [14,15]. Minimal attention has been paid in Canada, however, to developing interventions tailored to this population. The dearth of information about Internet-based sex work and the health practices and experiences of those involved may be contributing to oversights in eHealth intervention development [8,9,16].

Currently, the Internet is the primary advertisement medium used within the commercial sex industry. Internet advertisements for independent (ie, self-employed) Internet-based sex workers are in the hundreds of thousands each day on an international scale [1,17-23]. The analysis of information communicated within Internet-based sex work advertisements offers great promise to help improve our understanding of this population and the health and safety norms and expectations within the sex industry [8,10,21]. Indeed, an analysis of Internet advertisement content is a critical step in contributing to the knowledge needed to develop targeted and effective eHealth interventions.

To date, a small number of international studies have contributed to an evolving knowledge base about the content of Internet sex work advertisements. Although variation in the level of detail in advertisements occurs, preliminary evidence indicates that information about sexual services, pricing, locations, physical and interpersonal characteristics, and contact instructions are commonly included [8,13,19,21-28]. Gender variations in content have been reported suggesting that descriptions of personal spaces, personalities, and noncommercial services such as dinner and companionship are more common to women's advertisements [19]. By contrast, personal biographical statements and information about who they provide services to (eg, men, women, and couples) appear more consistently in

men's advertisements [22]. Photographs are frequently included although many are partial images where the person's face is not fully recognizable; a situation that may reflect concerns about confidentiality and privacy [19,27].

Much of the extant research on the content of Internet advertisements emphasizes men selling sexual services to other men with some conflicting results about their health practices [12,25,26]. In one US study, for instance, it was reported that 19% of Internet advertisements by men cited exclusive safer sex practices [24]. Conversely 90% of advertisements in a separate US project cited always engaging in safe sex and 16% included a statement of a preference for drug use during sexual activities [8]. Research concerning women Internet-based sex work advertisements and health and safety is minimal. Women's advertisements are frequently investigated from male clients' perspectives emphasizing clients' discussions and reviews of sex workers embedded in Internet social networking forums [28] or analysis framed by queries about what male clients find attractive [27]. Moreover, research concerned with the content of Internet advertisements has focused on marketing strategies, sexual service information, and privacy concerns about sex workers' identities. Less attention has been paid to the types of health, safety, and transactional information that are communicated within these advertisements.

Purpose of the Study

In this paper, we contribute to the growing knowledge about Internet-based sex work by detailing the results of a content analysis of Web-based advertisements within a Canadian context. The study purpose was to increase understandings of the health and safety information within Internet-based sex work advertisements and to offer suggestions for how this information may be utilized to inform eHealth program development. Because people of all genders engage in sex work [29] and gender matters for health [30], the content of Internet advertisements were examined within and between subgroups of women, men, and transgender or transpersons (defined as individuals who were assigned a sex at birth but express and experience their gender differently; however, they do not necessarily undergo sex reassignment or corrective surgery) [16] engaged in sex work.

Methods

Overview

This study was conducted using a content analysis approach [31,32] following university ethics approval. Content analysis has been used extensively by gender scholars to compare the information noted in Internet- and media-based communications of men and women [33]. Included were a series of iterative and

systematic processes that involved creating an empirically derived sample of Internet advertisements, determining the unit of analysis (eg, words, phrases, photos), developing a categorical scheme to code the data, and conducting data analysis [32].

Sample

The sample consisted of 75 verified Internet advertisements for sexual services provided by women (n=45), men (n=24), and transgender people (n=6) advertising specifically in Vancouver, British Columbia, Canada. The empirically driven sample [32] was selected from a database of 226 Web-based Canadian advertisements of women (n=173), men (35), and transgender (n=20) workers collected during ethnographic mapping activities of a larger study examining the relationships between health, safety, and the organizational features of an indoor sex industry. Ethnographic mapping involved a series of key informant interviews to detail prominent classified websites and search strategies for personal business websites situated within the local context, visiting the websites and classifieds 2 times per week over a 4-week period to collate number of posts and strategize for duplication removal, and narrowing to 3 Web-based classified platforms that were used most consistently;

the details of which are described elsewhere [20]. Purposeful sampling was used to capture the diversity in Web-based sex advertisement formats including advertisements from 35 personal websites and 40 from 3 Web-based classified advertising platforms. All advertisements were publicly accessible and membership or sign-in was not a requirement.

Data Collection

Due to the potential temporary nature of Internet advertisements, screenshots were taken and downloaded to a password-protected folder (see [Multimedia Appendix 1](#)). Drawing from previous research on the content of information communicated within Internet-based sex work advertisements [8,12,19,21,22,25-27,34] and the study objectives, a coding scheme that included relevant, advertisement-specific variables and related operational definitions ([Table 1](#)) was used [21,33]. Each variable was also coded as a categorical “yes” or “no” to document variable presence or absence within the advertisement. Data extraction and coding were undertaken by the first two authors collaboratively to aid reliability of the coding process. Data extraction of the variables and associated measures was completed by using a Microsoft Excel file template.

Table 1. Variables and operational definitions.

Variable	Operational definition
Sexual services provided	Sexual service provision communications ^a refer to the types of sexual acts that the person offers within their advertisement.
Sexual service restrictions	Restriction communications refer to the sexual services that the person advertising reports being unwilling to engage in during the provision of sexual services.
Health	Health communications were operationalized into two categories: (1) communications that specifically speak to health-related practices that could be considered protective or nonprotective for health (eg, nonsmoker, drug use, condom use) and (2) presence or absence of specific health issues (eg, HIV ^b , sexually transmitted infections).
Safety and security	Safety and security communications refer to factors associated with positively or negatively affecting people's physical, emotional, and financial well-being including their privacy. Specific categories for extraction included (1) location where sexual services provided, (2) instructions for clients on how to prepare for the encounter, (3) security measures taken before encounter including screening processes, (4) security measures during the encounter, and (5) facial photo details.
Demographic characteristics	Demographic communications refer to salient information about the “structure” of the population and included age, ethnicity, gender, and business title (eg, escort, companion).
Business practices	Business practices refer to communications that detail the dollar amounts associated with the cost of sexual services, contact and booking processes, and related restrictions.

^aCommunication was defined as the process of sharing information and conveying meaning through a shared system of semiotic rules and symbols [35]. Hypertexts or the array of words, phrases, and images found within advertisements were the focus of data collection [32].

^bHIV: human immunodeficiency virus.

Data Analysis

Variable frequencies were calculated using descriptive statistics and further classified according to gender. The social and commercialized contexts in which the information was embedded were also considered [32], including norms of language, the legal status of sex work in Canada, and the diversity of sexual services and Web-based advertising formats within the marketplace [9]. Thematic analysis of the hypertexts and related frequencies were undertaken to synthesize and categorize types of information communicated and the gender similarities and differences within the social and commercial contexts of the Web-based advertisements. Four communication

themes were identified: demographics, sexual services, health, and safety and security.

Results

Overview

The 75 advertisements varied in structure and format in how the information was communicated and organized. All the 75 advertisements marketed sex services to men and 2 men's advertisements additionally advertised services for women clients. In total, 78% (35/45) of women's advertisements were within their personal business websites, whereas men's (22/24; 83%) and transpersons' advertisements (5/6; 83%) appeared

within Internet classified advertising platforms. Unlike tailored personal websites, the advertisements within the classified platforms adhered to an advertising template with information organized into predetermined categories (eg, age, weight, services offered, classification). Women's advertisements within their personal websites frequently described their workspaces in detail with emphasis on the provision of a relaxing atmosphere. Classified platforms did not afford the opportunity to detail a space, only whether or not one was provided. Health, safety, demographic characteristics, and sexual services were regularly communicated in the 75 advertisements regardless of platform used.

Demographic Communications

Sex industry language norms were evident in the work titles and physical characteristics communicated throughout the

advertisements with variations that mapped onto gender. Although "escort" was the most common work title (Table 2) across groups, it was the only term used in the men's and transpersons' advertisements. By contrast, almost half of the women's advertisements included other titles that evoked specific heterosexual erotic identities, such as "courtesan," "playmate," or "dominatrix." The use of escort for a work classification may also be understood within the constraints of the Canadian laws that criminalize communications associated with selling and purchasing sexual services. The term "escort," for example, is a legally recognized business title within the City of Vancouver municipal bylaws [36], whereas other titles are not. Additionally, the laws may contribute to restrictions on some of the Web-based classified platforms regarding language permitted within their site.

Table 2. Demographic characteristic communications.

Characteristics	Women (n=45)	Men (n=24)	Transgender persons (n=6)	Total n=75 n (%)
Work title				
Escort	18	24	6	48 (64)
Companion	5			5 (7)
Multiple roles	4			4 (5)
Other	10			10 (13)
Not specified	8			8 (10)
Age (years)				
20-29	14	16	2	32 (43)
30-39	9	4	1	14 (18)
40-49	1	3	1	5 (7)
Not specified	21	1	2	24 (32)
Ethnicity				
White	29	13	4	46 (61)
Black	2	1		3 (4)
Latino or Hispanic	1	3		4 (5)
Asian	9	4	1	14 (19)
Multiethnic	4	3	1	8 (10)
Weight				
Advertisements specifying weight	25	24	4	53 (71)
Average of specified weights (lbs)	118	167	135	

In total, 25 of the 75 advertisements did not specify the sex workers' ethnicity. In these cases, photos accompanying advertisements provided an indication from which ethnicity was estimated. White was the most common ethnicity stated or observed (Table 2). Men's advertisements were the only ones to consistently provide age, weight, and ethnicity (Table 2), which may reflect the template of the classifieds' advertising platforms. Two transpersons' advertisements communicated surgical or hormone replacement therapy details such as "undergone hormone replacement therapy" or "nonoperative transsexual" highlighting the significance of language to

communicate about their embodied specificities. Men's communications detailed their muscular build or fitness, for instance, "physically fit, muscular" or "lean or swimmer build" and included physical characteristics, the most common being weight, body hair, body type, and penis size. This differed from the women and transpersons' advertisements wherein physical descriptions such as bust size were present in some cases but did not appear uniformly.

Sexual Service Communications

The language used to communicate about sexual services among the 64 advertisements listing sexual services reflected sex industry norms. Variation between genders and homogeneity within was observed (Table 3). The term “full service,” communicating to clients a willingness to engage in sexual activities that included vaginal intercourse was found within women’s advertisements. Women also listed the “girlfriend

experience (GFE) or boyfriend experience (BFE),” which is defined as the provision of dating activities and sex that would be expected in a noncommercial adult relationship [37], more often than men. Men’s and transgender peoples’ advertisements used the terms “top,” “bottom,” or “versatile,” indicating their sexual positions for anal intercourse. Men also used the term “anal” explicitly, whereas women used the term “greek” to indicate anal services.

Table 3. Sexual services communications.

Characteristics	Women (n=45)	Men (n=24)	Transgender persons (n=6)	Total n=75 n (%)
Advertisements with services listed	35	23	6	64 (85)
Type of services				
Full service	20			20 (27)
Massage	18	18		36 (48)
GFE ^a or BFE ^b	17	2	1	20 (27)
Companionship	12	12		24 (32)
Duos	12	1		13 (17)
BDSM	12	12	1	25 (33)
Greek or anal	6	16	3	25 (33)
Uncovered services	11	6	1	18 (24)
Top (position)		6		
Bottom (position)		6		
Versatile (top or bottom)		11	2	
Advertisements with service restrictions	26	9	2	37 (49)

^aGFE: girlfriend experience.

^bBFE: boyfriend experience.

The language of men’s advertisements was further associated with terms that are used to describe specific sexual practices between men, for instance, “rimming, fisting, and watersports,” and included detailed lists of sexual acts that they would engage in. Diverse BDSM services were listed across genders. BDSM is a term used to denote varied consensual sexual practices that include bondage, domination or submission, pain or sensation play, power exchange, leathersex, role playing, and fetishes [38]. The women’s advertisements offering BDSM services additionally communicated information about health and safety. Clients were asked to provide information about any medical issues and a safe word to signal when they wanted to stop the activity before services. Women’s advertisements more often identified and provided hyperlinks to the websites of other men and women sex workers with whom they worked as part of their “duo” services (Table 3).

Communications about condom use included the terms “covered” (ie, condom) and “uncovered” (ie, no condom required) with service variations noted among the 64 advertisements listing services. Uncovered services were usually specific to oral-genital sex (Table 3). Women used industry acronyms such as BBBJ (bareback blowjob), CIM (cum in mouth), DATY (dining at the Y/cunnilingus), and DATO (dining at the O/analagus) to detail the service. Uncovered analagus

or “rimming,” was offered in 6 of the 23 men’s advertisements listing services. In contrast, 41% (31/75) of advertisements stated that they *never* provided uncovered services, often describing their services as “always safe” (see [Multimedia Appendix 2](#)). Women communicated more restrictions overall, sometimes posting an “etiquette” or “frequently asked questions” section that outlined restrictions.

Health Communications

There were noteworthy gender-based variations in the Web-based information about personal health practices. Transworkers’ advertisements rarely communicated any personal health information, whereas men and women consistently provided health-related details. Communications about cigarette smoking and substance use were the most common types of personal health practice communications (Table 4). It was found that 7% (3/45) of women’s advertisements also contained communications requesting clients refrain from smoking before or during appointments, and one of the women’s advertisements indicated that she would not visit smoking rooms in hotels.

Although none of the men or transgender persons explicitly restricted clients’ drug or alcohol use, 22% (10/45) of women’s advertisements stated a refusal to see a client who was, or appeared to be, under the influence of drugs or alcohol. It was

found that 25% (6/24) of men's advertisements stating that they *did* use drugs noted being okay with "party and play" (PNP), a euphemism for acknowledgment of the availability for drugs, frequently crystal methamphetamine, to be taken in conjunction with having sex [39]. None of the advertisements stated specific illicit drugs with the exception of a 2 women's advertisements that specified being "420 friendly" indicating marijuana was acceptable.

Communications about STIs were infrequent (12/75, 16%) and higher among men's advertisements. "Clean" and

"infection-free" were the terms used. One of the men's advertisements communicated that he was "HIV-positive" but "undetectable," a status referring to HIV viral load as measured through serum testing [40]. His advertisement further noted that although bareback was preferred, he would provide covered services. Client hygiene was commented upon in women's advertisements only. Expectations of "good hygiene" were noted and they also included information about the facility to shower at their in-call before services (see [Multimedia Appendix 2](#)).

Table 4. Health communications.

Characteristics	Women (n=45)	Men (n=24)	Transgender persons (n=6)	Total n=75 n (%)
Health communications	29	24	1	64 (85)
Sex workers' personal health behaviors communications				
Nonsmoker	16	9		25 (33)
Nondrug using	13	3		16 (21)
"Social" alcohol use	7	4		11 (15)
Uses drugs	2	6		8 (11)
Other (physical activities such as yoga, gym)	3	10	1	14 (19)
Communicable infection communications	5	8		12 (16)

Safety and Security Communications

Communications associated with personal and financial safety and security occurred in all advertisements and were regularly embedded in information for potential clients to learn about the operational features of connecting with and making an appointment with a worker ([Table 5](#)). For example, in addition to detailing the location where appointments would be held, specific contact instructions were provided. Multiple options for making contact were consistently available across genders ([Table 5](#)); however, the patterns of use varied with men and transpersons' advertisements more frequently using the messaging features of the hosting classified advertisement platform. Women utilized a live (ie, phone) booking process with potential clients. It was found that 9% (7/75) advertisements noted not accepting blocked numbers for phone calls, of which 5 were in women's advertisements.

The advertisements also illustrated that Internet-based sex work occurs within in-call spaces such as apartments provided by the sex worker, and out-call locations where a sex worker travels to the client. Although there were limited restrictions communicated concerning out-call locations, gender variance was observed with only women communicating that they would not provide out-call services to private residences, or in hotels rated below a 4-5 star accommodation. It was found that 9% (4/45) of the women required references from another service provider and one woman required clients to check-in with hotel

security before the exchange. None of the transworkers shared any communications about willingness to travel; although just under one-third and one-quarter of men and women's advertisements, respectively, noted this as an option.

Attempts to maintain privacy and security were also reflected in the photos provided within the advertisements. One-third of advertisements contained identifiable facial photos, including all of transpersons' advertisements, and one-half provided a viewable photo of some kind ([Table 5](#)). The partial face photos revealed features such as eyes or lips, or in some cases a side profile with shadowing that made it more difficult to identify the person in the photo. In total, 2 of the 45 women's advertisements explicitly noted that clients were not to take photos or videos during an appointment and one woman's advertisement stated that she would agree to photos and videos only if she was confident that they would never be posted on the Internet, but she did not state how that would be negotiated.

The communications about pricing provided concise information about fees. The price of service varied with women charging on average approximately 25% less than men per exchange ([Table 5](#)). Admittedly, the highest rate in the sample, Can \$700/h, was listed on a woman's site; however, the advertisement was an outlier in terms of fee. The men and transworkers' sites listed consistent prices per hour; and both men and women charged significantly more for travel and did not include the price of airfare, hotels, and meals, which was specified separately.

Table 5. Safety and security communications.

Characteristics	Women (n=45)	Men (n=24)	Transgender persons (n=6)	Total n=75 n (%)
Safety and security restrictions for client behaviors identified	26	9	2	37 (49)
Service locations				
In-call	35	12	3	50 (67)
Out-call	34	18	4	56 (75)
Willing to travel out of town	9	7		16 (21)
Viewable facial photo	22	11	5	38 (51)
Full	13	9	5	27 (36)
Partial	9	2		11 (15)
Booking contact process				
Phone	35	15	5	55 (73)
Website	15	19	3	37 (49)
Text	8	1	1	10 (13)
Email	18	5	9	32 (43)
Rates (Canadian dollar)				
Minimum	140	180	200	
Maximum	700	450	250	
Average	313	247	233	

Discussion

Principal Findings

The study findings build on nascent empirical understandings of the content of Web-based advertisements [19,22,23,27,41]. This research also contributes particular topics for consideration in public health and eHealth programming among the growing population of sex workers who use the Internet as their primary means to communicate, at least initially, occupational health and safety details to potential clients. In these advertisements, sex workers communicated significant information that reflected their personal health practices both within and outside commercial sex exchanges (eg, drug and alcohol use, condom use, physical fitness). These health communications provide important insights to the nature and norms of Internet-based sex work and the practices, health behaviors, and overall wellbeing of those advertising in this context.

The findings also illustrate that gender matters when considering health and safety practices and Internet-based communications. Women list more safety communications overall and do so by communicating restrictions, particularly about place and substance use. Women also use personalized screening processes involving live phone communications to determine the suitability of the client [5]. Men and transpersons communicate minimal safety details. These findings add nuance to our understandings about different strategies between men and women sex workers to mitigate risks of violence [13,42,43] by demonstrating the role of Internet-based communications within the realm of the commercialized exchange. Gender and technology also intersect to influence the content of communications including how

information is presented, language used, and the actual types of details included. Women tend to use personal business websites to advertise providing substantial details about etiquette and atmosphere. Men and transpersons use Web-based classified platforms with predetermined categories and emphasize physical characteristics. Understanding the gender nuance of advertising platforms is essential to developing targeted eHealth messaging that takes into consideration the norms of communication within the realm of the diverse platforms used [1,18].

The findings also provide some new insights into population characteristics within Canadian advertisements situated in this locale as predominantly white, young workers who charged over Can \$200/h, on average, for their services in both in-call and out-call settings. These characteristics are similar to those described in other North American Internet-based sex work research [27,44] and in direct contrast to research situated in street-based marketplaces that demonstrate higher numbers of Indigenous, Black, and Hispanic people aged over 30 years and diverse financial arrangements where fees charged are amount per service with much lower income potential [25,45]. As expected, the snapshot of the Internet-based sex industry captured in these 75 advertisements also indicates that men, women, and trans-sex workers provide services almost exclusively to male clients.

The health communications regarding condom use raised several important issues warranting attention. Condom use was regularly communicated for anal or vaginal intercourse, yet less than half of the advertisements communicated that a condom was required for oral-genital sex. Although there is growing evidence that many sex workers may experience STI infection rates similar to or less than the general population [12,46,47], the lack of

condom requirement for oral-genital sex may still pose a significant health threat for them [41,48,49]. For example, STIs such as syphilis, gonorrhea, and chlamydia can be transmitted via oral-genital contact and can be asymptomatic [50]. Women aged under 30 years and men having sex with men are disproportionately affected by some STIs [51]; an important fact given the demographic characteristics of the study sample. Moreover, there is some evidence that sex workers, like many in the general population, may not fully comprehend the possibilities of transmission of STIs through oral-genital body fluid exchange [52,53].

Developing health promotion strategies for condom use, however, requires nuanced understandings of evolving commercial sex industry norms. Men and women Internet-based sex workers are engaged in offering emotionally erotic exchanges for sale and purchase that are designed to foster intimacy and replicate aspects of noncommercial relationships [54,55]. Additionally, the demand for unprotected sexual services is steadily on the rise and is directly contributing shifting norms and economic influences shaping safer sex practices [47,53]. It has been demonstrated, for instance, that women's willingness to engage in uncovered oral sex is associated with increased earnings offered by clients for those services irrespective of their health-related knowledge [53]. How the issues of knowledge, intimacy, companionship, economy, and shifting norms influence people's practices, their health and safety highlights the need for contextually appropriate approaches to sexual health promotion and STI prevention. Contextually appropriate approaches that take into consideration the multiple cofactors that can contribute to STI transmission and the diverse needs that people may have with regards to building prevention capacity have been shown effective in diverse populations [56]. It is, therefore, critical that public health consider how to target multiple interrelated issues such as money, intimacy, and familiarity in the campaigns aimed at promoting sexual health and people's safer sex behaviors.

Finally, Internet-based sex workers are often considered a hard-to-reach group particularly because their businesses operate in limitless Web-based spaces afforded by Internet-based advertising services [20]. Isolation and barriers in access to health promotions services have been noted among this population [7-13]. eHealth programs delivered via the Internet are often successful because of their potential to reach people who experience isolation [57]. To be effective, however, it is essential that such health programming be developed to reach the diversity of sex workers and to do so, programming must involve the active participation of those engaged in sex work in the design and implementation [11,58,59]. As the World

Health Organization (WHO) strategy for the health promotion of sex workers noted, "health services should be made available, accessible, and acceptable to sex workers" (p.xix). Moreover, health promotion programs that build upon the existing strengths and the knowledge base of sex workers have a proven track record as the most successful to promote health and safety [11]. Additionally, health promotion strategies must include and extend beyond STI prevention. The need for greater knowledge and skills in the areas of safe and effective Internet advertising strategies, attracting and screening appropriate clients, addressing legal concerns, and financial planning—an important aspect of transitioning out of the industry have been identified [44,60]. Given the evidence of increased autonomy and independence associated with Internet-based sex work [4-6], there is perhaps an opportunity for health programming to capitalize on the benefits of autonomy and control in sex workers' choice to communicate explicit Web-based restrictions, and to promote and expand self-health in the sex industry as a highly desirable and marketable characteristic.

Limitations

The research is limited by the small sample and the limited information gleaned from the transworkers' advertisements. And, although content analysis provides some information to help understand health and safety practices, additional research with a larger sample that includes the perspectives and experiences of Internet-based sex workers is warranted.

Conclusions

This descriptive study is one of the first to analyze the content of Internet-based sex work advertisements in Canada, and is one of a few to include men, women, and transgender people in a single study. Content analysis is a feasible approach to understanding Internet-based communications and must be nuanced within the social and commercialized context of the sex industry to help inform health service programming for Internet-based sex workers [44]. Strategies for eHealth promotion must be informed by more relational approaches to sex and intimacy that are inclusive of sex workers and their clients and considers gender-specific aspects of these relationships. As the Internet continues to grow, greater numbers of sex workers will likely advertise services via the Internet. Important insights can be gleaned by describing Internet advertisement content about industry practices including the emergence of self-management akin to small business ownership. With this in mind, we might thoughtfully continue to describe Internet-based sex workers' advertisements as a means to developing more effective eHealth programming to better support the health of sex workers and their clients.

Acknowledgments

This paper is based on the research of Julie Ann Kille for her Master of Science degree in Nursing, 2015, at the University of British Columbia, Vancouver, Canada and the SPACES (sex, power, agency, consent, environment, safety) study. Funding for this project was provided by the Canadian Institutes for Health Research (grant number MOP-11947). Vicky Bungay is partially supported by a Canada Research Chair in Gender, Equity and Community Engagement and Michael Smith Foundation for Health Scholar Award (grant number 5212).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample screenshots from a personal business website and a Web-based classified advertising site.

[PDF File (Adobe PDF File), 225KB - [jmir_v19i4e111_app1.pdf](#)]

Multimedia Appendix 2

Additional advertising details are provided in table format concerning communications about restrictions, infections, and rules for client health behaviors.

[PDF File (Adobe PDF File), 38KB - [jmir_v19i4e111_app2.pdf](#)]

References

1. Ashford C. Male sex work and the Internet effect: time to re-evaluate the criminal law? *J Crim Law* 2009 Jun;73(3):258-280. [doi: [10.1350/jcla.2009.73.3.573](#)]
2. Murphy AK, Venkatesh SA. Vice careers: the changing contours of sex work in New York City. *Qual Sociol* 2006 Jun 27;29(2):129-154. [doi: [10.1007/s11133-006-9012-2](#)]
3. Pitcher J. Sex work and modes of self-employment in the informal economy: diverse business practices and constraints to effective working. *Soc Policy Soc* 2015 Jan;14(1):113-123 [FREE Full text] [doi: [10.1017/S1474746414000426](#)] [Medline: [25506264](#)]
4. Argento E, Taylor M, Jollimore J, Taylor C, Jennex J, Krusi A, et al. The loss of Boystown and transition to online sex work: strategies and barriers to increase safety among men sex workers and clients of men. *Am J Mens Health* 2016 Jun 28:1-12 Epub ahead of print. [doi: [10.1177/1557988316655785](#)] [Medline: [27352925](#)]
5. Atchison C, Benoit C, Burnett P, Jansson M. NSW. 2015. The influence of time to negotiate on control in sex worker-client interactions URL: http://www.nswp.org/sites/nswp.org/files/R4SW%202015_issue14_PDFV.pdf [accessed 2016-10-04] [WebCite Cache ID 612cU7yIN]
6. Minichiello V, Scott J, Callander D. New pleasures and old dangers: reinventing male sex work. *J Sex Res* 2013;50(3-4):263-275. [doi: [10.1080/00224499.2012.760189](#)] [Medline: [23480072](#)]
7. Bar-Johnson MD, Weiss P. A comparison of male sex workers in Prague: Internet escorts versus men who work in specialized bars and clubs. *J Sex Res* 2015;52(3):338-346. [doi: [10.1080/00224499.2013.848256](#)] [Medline: [24423089](#)]
8. Blackwell CW, Dziegielewska SF. Risk for a price: sexual activity solicitations in online male sex worker profiles. *J Soc Serv Res* 2013 Mar;39(2):159-170. [doi: [10.1080/01488376.2012.744617](#)]
9. Minichiello V, Scott J, Callander D. A new public health context to understand male sex work. *BMC Public Health* 2015;15:282 [FREE Full text] [doi: [10.1186/s12889-015-1498-7](#)] [Medline: [25879716](#)]
10. Nguyen M, Venne T, Rodrigues I, Jacques J. Why and according to what consultation profiles do female sex workers consult health care professionals? A study conducted in Laval, Québec. *Health Care Women Int* 2008 Feb;29(2):165-182. [doi: [10.1080/07399330701738226](#)] [Medline: [18350422](#)]
11. World Health Organization, UNFPA, UNAIDS, NSWP, World Bank, UNDP. Implementing comprehensive HIV/STI programmes with sex workers: practical approaches from collaborative interventions. Geneva, Switzerland: World Health Organization; 2013.
12. Uy J, Parsons J, Bimbi D, Koken J, Halkitis P. Gay and bisexual male escorts who advertise on the internet: understanding reasons for and effects of involvement in commercial sex. *Int J Mens Health* 2004 Jan 1;3(1):11-26. [doi: [10.3149/jmh.0301.11](#)]
13. Manning E, Bungay V. 'Business before pleasure': the golden rule of sex work, payment schedules and gendered experiences of violence. *Cult Health Sex* 2017 Mar;19(3):338-351. [doi: [10.1080/13691058.2016.1219767](#)] [Medline: [27600513](#)]
14. Hong Y, Li X, Fang X, Lin X, Zhang C. Internet use among female sex workers in China: implications for HIV/STI prevention. *AIDS Behav* 2011 Feb;15(2):273-282. [doi: [10.1007/s10461-010-9846-1](#)] [Medline: [21082341](#)]
15. Xiao Z, Li X, Lin D, Tam CC. Mass media and HIV/AIDS prevention among female sex workers in Beijing, China. *J Health Commun* 2015;20(9):1095-1106. [doi: [10.1080/10810730.2015.1018575](#)] [Medline: [25950448](#)]
16. The SPACES Team. Open.library.ubc. 2016. Recommendations from the off-street sex industry in Vancouver URL: <https://open.library.ubc.ca/cIRcle/collections/facultyresearchandpublications/52383/items/1.0340040> [accessed 2017-01-01] [WebCite Cache ID 6o5AnJgUG]
17. Cunningham S, Kendall TD. Risk behaviours among internet-facilitated sex workers: evidence from two new datasets. *Sex Transm Infect* 2010 Dec;86(Suppl 3):iii100-iii105. [doi: [10.1136/sti.2010.044875](#)] [Medline: [20852309](#)]
18. McLean A. 'You can do it from your sofa': the increasing popularity of the internet as a working site among male sex workers in Melbourne. *J Sociol* 2015 Dec 08;51(4):887-902. [doi: [10.1177/1440783313480416](#)]

19. Agresti BT. search.proquest. Washington: The George Washington University, ProQuest Dissertations Publishing; 2009. E-prostitution: A Content Analysis of Internet Escort Websites: Master's Thesis URL: <http://search.proquest.com/docview/304882040> [accessed 2017-03-24] [WebCite Cache ID 6pDUPuPCc]
20. Bungay V, Oliffe J, Atchison C. Addressing underrepresentation in sex work research: reflections on designing a purposeful sampling strategy. *Qual Health Res* 2016 Jun;26(7):966-978. [doi: [10.1177/1049732315613042](https://doi.org/10.1177/1049732315613042)] [Medline: [26589337](https://pubmed.ncbi.nlm.nih.gov/26589337/)]
21. Koken J, Bimbi D, Parsons J. Malefemale escorts: a comparative analysis. In: *Sex for sale*. 2nd ed. New York, NY: Routledge; 2010:205-232.
22. Lee-Gonyea JA, Castle T, Gonyea NE. Laid to order: male escorts advertising on the Internet. *Deviant Behavior* 2009 Apr 03;30(4):321-348. [doi: [10.1080/01639620802168858](https://doi.org/10.1080/01639620802168858)]
23. Pruitt MV. Online boys: male-for-male Internet escorts. *Socio Focus* 2005 Aug;38(3):189-203. [doi: [10.1080/00380237.2005.10571265](https://doi.org/10.1080/00380237.2005.10571265)]
24. Logan TD. Personal characteristics, sexual behaviors, and male sex work: a quantitative approach. *Am Sociol Rev* 2010 Oct 08;75(5):679-704. [doi: [10.1177/0003122410379581](https://doi.org/10.1177/0003122410379581)]
25. Mimiaga MJ, Reisner SL, Tinsley JP, Mayer KH, Safren SA. Street workers and internet escorts: contextual and psychosocial factors surrounding HIV risk behavior among men who engage in sex work with other men. *J Urban Health* 2009 Jan;86(1):54-66 [FREE Full text] [doi: [10.1007/s11524-008-9316-5](https://doi.org/10.1007/s11524-008-9316-5)] [Medline: [18780186](https://pubmed.ncbi.nlm.nih.gov/18780186/)]
26. T. Parsons, David Bimbi, Perry N. H J. Sexual compulsivity among gay/bisexual male escorts who advertise on the internet. *Sexual Addict Compulsivity* 2001 Apr;8(2):101-112. [doi: [10.1080/10720160127562](https://doi.org/10.1080/10720160127562)]
27. Castle T, Lee J. Ordering sex in cyberspace: a content analysis of escort websites. *Int J Cult Stud* 2008 Mar 01;11(1):107-121. [doi: [10.1177/1367877907086395](https://doi.org/10.1177/1367877907086395)]
28. Langanke H, Ross MW. Web-based forums for clients of female sex workers: development of a German internet approach to HIV/STD-related sexual safety. *Int J STD AIDS* 2009 Jan;20(1):4-8. [doi: [10.1258/ijsa.2008.008202](https://doi.org/10.1258/ijsa.2008.008202)] [Medline: [19103884](https://pubmed.ncbi.nlm.nih.gov/19103884/)]
29. Vanwesenbeeck I. Prostitution push and pull: male and female perspectives. *J Sex Res* 2013 Jan;50(1):11-16. [doi: [10.1080/00224499.2012.696285](https://doi.org/10.1080/00224499.2012.696285)] [Medline: [23030049](https://pubmed.ncbi.nlm.nih.gov/23030049/)]
30. Johnson JL, Repta R. Sex and gender: beyond the binaries. In: *Designing and conducting gender, sex and health research*. Thousand Oaks, CA: SAGE Publishing; 2012:17-36.
31. Polit D, Hungler B. *Nursing research: principles and methods*. 4th edition. Philadelphia: Lippincott; 1991.
32. Krippendorff K. *Content Analysis: An Introduction to Its Methodology*. 2nd edition. Thousand Oaks, Calif: Sage; 2004.
33. Neuendorf KA. Content analysis: a methodological primer for gender research. *Sex Roles* 2010 Nov 20;64(3-4):276-289. [doi: [10.1007/s11199-010-9893-0](https://doi.org/10.1007/s11199-010-9893-0)]
34. Parsons JT, Koken JA, Bimbi DS. The use of the Internet by gay and bisexual male escorts: sex workers as sex educators. *AIDS Care* 2004 Nov;16(8):1021-1035. [doi: [10.1080/09540120412331292405](https://doi.org/10.1080/09540120412331292405)] [Medline: [15511734](https://pubmed.ncbi.nlm.nih.gov/15511734/)]
35. Knapp M, Daly J. *The SAGE Handbook of Interpersonal Communication*. Thousand Oaks, California: Sage Publications, Inc; 2011.
36. City of Vancouver. 2016. License By-law No 4450 URL: <http://former.vancouver.ca/bylaws/4450c.pdf> [WebCite Cache ID 6l0j9CgfE]
37. Milrod C, Monto MA. The hobbyist and the girlfriend experience: behaviors and preferences of male customers of internet sexual service providers. *Deviant Behavior* 2012 Nov;33(10):792-810. [doi: [10.1080/01639625.2012.707502](https://doi.org/10.1080/01639625.2012.707502)]
38. Weiss M. *Techniques of Pleasure: BDSM and the Circuit of Sexuality*. Durham, NC: Duke University Press; 2012.
39. Frederick BJ, Perrone D. "Party N Play" on the Internet: subcultural formation, craigslist, and escaping from stigma. *Deviant Behavior* 2014 Jul 29;35(11):859-884. [doi: [10.1080/01639625.2014.897116](https://doi.org/10.1080/01639625.2014.897116)]
40. Van Den Boom W, Stolte IG, Witlox R, Sandfort T, Prins M, Davidovich U. Undetectable viral load and the decision to engage in unprotected anal intercourse among HIV-positive MSM. *AIDS Behav* 2013 Jul;17(6):2136-2142. [doi: [10.1007/s10461-013-0453-9](https://doi.org/10.1007/s10461-013-0453-9)] [Medline: [23525838](https://pubmed.ncbi.nlm.nih.gov/23525838/)]
41. Kolar K, Atchison C, Bungay V. Sexual safety practices of massage parlor-based sex workers and their clients. *AIDS Care* 2014;26(9):1100-1104. [doi: [10.1080/09540121.2014.894611](https://doi.org/10.1080/09540121.2014.894611)] [Medline: [24617632](https://pubmed.ncbi.nlm.nih.gov/24617632/)]
42. Smith MD, Seal DW. Sexual behavior, mental health, substance use, and HIV risk among agency-based male escorts in a small U.S. city. *Int J Sex Health* 2008 Mar 1;19(4):27-39 [FREE Full text] [doi: [10.1300/J514v19n04_04](https://doi.org/10.1300/J514v19n04_04)] [Medline: [19779600](https://pubmed.ncbi.nlm.nih.gov/19779600/)]
43. Bungay V, Halpin M, Halpin PF, Johnston C, Patrick DM. Violence in the massage parlor industry: experiences of Canadian-born and immigrant women. *Health Care Women Int* 2012;33(3):262-284. [doi: [10.1080/07399332.2011.603868](https://doi.org/10.1080/07399332.2011.603868)] [Medline: [22325026](https://pubmed.ncbi.nlm.nih.gov/22325026/)]
44. Grov C, Rodríguez-Díaz CE, Ditmore MH, Restar A, Parsons JT. What kinds of workshops do Internet-based male escorts want? Implications for prevention and health promotion. *Sex Res Soc Policy* 2014 Mar 7;11(2):176-185. [doi: [10.1007/s13178-014-0151-z](https://doi.org/10.1007/s13178-014-0151-z)]
45. Benoit C, Shumka L. *Understandingsexwork*. 2015. Sex work in Canada URL: http://www.understandingsexwork.com/sites/default/files/uploads/2015%2005%2007%20Benoit%20&%20Shumka%20Sex%20Work%20in%20Canada_2.pdf [WebCite Cache ID 6l2cmu1C]

46. Donovan B, Harcourt C, Egger S, Fairley CK. Improving the health of sex workers in NSW: maintaining success. *N S W Public Health Bull* 2010;21(3-4):74-77 [FREE Full text] [doi: [10.1071/NB10013](https://doi.org/10.1071/NB10013)] [Medline: [20513305](https://pubmed.ncbi.nlm.nih.gov/20513305/)]
47. El-Hayek C, van GC, Bowring A, Feigin A, Stoové M. Burnet.edu. 2011. Review of clinical and health promotion services for sex workers in Western Australia: prepared for the Western Australian Department of Health URL: http://www.burnet.edu.au/system/asset/file/1125/2011_WA_review_of_sex_worker_project_FINAL.pdf [WebCite Cache ID 6l0jDn2WR]
48. Bimbi DS, Parsons JT. Barebacking among internet based male sex workers. *J Gay Lesb Ment Health* 2005 Oct 21;9(3-4):85-105. [doi: [10.1300/J236v09n03_06](https://doi.org/10.1300/J236v09n03_06)]
49. Bungay V, Kolar K, Thindal S, Remple VP, Johnston CL, Ogilvie G. Community-based HIV and STI prevention in women working in indoor sex markets. *Health Promot Pract* 2013 Mar;14(2):247-255. [doi: [10.1177/1524839912447189](https://doi.org/10.1177/1524839912447189)] [Medline: [22885289](https://pubmed.ncbi.nlm.nih.gov/22885289/)]
50. Holmes K, Sparling P, Stamm W, Piot P, Wasserheit J, Corey L, et al. Sexually Transmitted Diseases. 4th edition. New York, NY: McGraw Hill; 2008:0071417486.
51. Public Health Agency of Canada. Catie.ca. Report on sexually transmitted infections in Canada 2014 URL: http://www.catie.ca/sites/default/files/64-02-14-1200-STI-Report-2011_EN-FINAL.pdf [accessed 2016-10-05] [WebCite Cache ID 6l2cKNuLz]
52. Minichiello V, Mariño R, Browne J. Knowledge, risk perceptions and condom usage in male sex workers from three Australian cities. *AIDS Care* 2001 Jun;13(3):387-402. [doi: [10.1080/09540120120044035](https://doi.org/10.1080/09540120120044035)] [Medline: [11397340](https://pubmed.ncbi.nlm.nih.gov/11397340/)]
53. Handlovsky I, Bungay V, Kolar K. Condom use as situated in a risk context: women's experiences in the massage parlour industry in Vancouver, Canada. *Cult Health Sex* 2012 Oct;14(9):1007-1020. [doi: [10.1080/13691058.2012.712720](https://doi.org/10.1080/13691058.2012.712720)] [Medline: [22900640](https://pubmed.ncbi.nlm.nih.gov/22900640/)]
54. Bernstein E. Sex Work for the Middle Classes. *Sexualities* 2007 Oct 01;10(4):473-488. [doi: [10.1177/1363460707080984](https://doi.org/10.1177/1363460707080984)]
55. Walby K. Touching Encounters: Sex, Work, and Male-for-Male Internet Escorting. Chicago, IL: University of Chicago Press; 2012.
56. Stall R, Herrick A, Guadamuz T, Friedman M. Updating HIV prevention with gay men: current challenges and opportunities to advance health among gay men. In: *HIV prevention : a comprehensive approach*. Cambridge: Elsevier; 2009:267-280.
57. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. *Annu Rev Public Health* 2009;30:273-292. [doi: [10.1146/annurev.publhealth.031308.100235](https://doi.org/10.1146/annurev.publhealth.031308.100235)] [Medline: [19296777](https://pubmed.ncbi.nlm.nih.gov/19296777/)]
58. Bowen R, Bungay V. Taint: an examination of the lived experiences of stigma and its lingering effects for eight sex industry experts. *Cult Health Sex* 2016;18(2):186-199. [doi: [10.1080/13691058.2015.1072875](https://doi.org/10.1080/13691058.2015.1072875)] [Medline: [26671209](https://pubmed.ncbi.nlm.nih.gov/26671209/)]
59. Parsons JT, Koken JA, Bimbi DS. Looking beyond HIV: eliciting individual and community needs of male internet escorts. *J Homosex* 2007;53(1-2):219-240. [doi: [10.1300/J082v53n01_10](https://doi.org/10.1300/J082v53n01_10)] [Medline: [18019076](https://pubmed.ncbi.nlm.nih.gov/18019076/)]
60. Bowen RR. Squaring up: experiences of transition from off-street sex work to square work and duality--concurrent involvement in both--in Vancouver, BC. *Can Rev Sociol* 2015 Nov;52(4):429-449. [doi: [10.1111/cars.12085](https://doi.org/10.1111/cars.12085)] [Medline: [26577882](https://pubmed.ncbi.nlm.nih.gov/26577882/)]

Abbreviations

BBBJ: bareback blowjob
BFE: boyfriend experience
CIM: cum in mouth
DATO: dining at the O/analingus
DATY: dining at the Y/cunnilingus
GFE: girlfriend experience
PNP: party and play
STI: sexually transmitted infection
WHO: World Health Organization

Edited by G Eysenbach; submitted 05.10.16; peer-reviewed by V Minichiello, Y Hong; comments to author 09.01.17; revised version received 14.02.17; accepted 04.03.17; published 13.04.17.

Please cite as:

Kille J, Bungay V, Oliffe J, Atchison C
A Content Analysis of Health and Safety Communications Among Internet-Based Sex Work Advertisements: Important Information for Public Health
J Med Internet Res 2017;19(4):e111
 URL: <http://www.jmir.org/2017/4/e111/>
 doi:[10.2196/jmir.6746](https://doi.org/10.2196/jmir.6746)
 PMID:[28408364](https://pubmed.ncbi.nlm.nih.gov/28408364/)

©Julie Kille, Vicky Bungay, John Oliffe, Chris Atchison. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

eHealth Search Patterns: A Comparison of Private and Public Health Care Markets Using Online Panel Data

Janina Anne Schneider^{1*}, BA(Hons), MSc, PhD; Christopher Patrick Holland^{1,2*}, BSc (Hons), PhD

¹Alliance Manchester Business School, University of Manchester, Manchester, United Kingdom

²Information Systems Group, University of Münster, Münster, Germany

* all authors contributed equally

Corresponding Author:

Christopher Patrick Holland, BSc (Hons), PhD

Alliance Manchester Business School

University of Manchester

Booth Street East

Manchester, M13 9PL

United Kingdom

Phone: 44 161 820 8344 ext 56460

Fax: 44 161 275 6596

Email: chris.holland@manchester.ac.uk

Abstract

Background: Patient and consumer access to eHealth information is of crucial importance because of its role in patient-centered medicine and to improve knowledge about general aspects of health and medical topics.

Objectives: The objectives were to analyze and compare eHealth search patterns in a private (United States) and a public (United Kingdom) health care market.

Methods: A new taxonomy of eHealth websites is proposed to organize the largest eHealth websites. An online measurement framework is developed that provides a precise and detailed measurement system. Online panel data are used to accurately track and analyze detailed search behavior across 100 of the largest eHealth websites in the US and UK health care markets.

Results: The health, medical, and lifestyle categories account for approximately 90% of online activity, and e-pharmacies, social media, and professional categories account for the remaining 10% of online activity. Overall search penetration of eHealth websites is significantly higher in the private (United States) than the public market (United Kingdom). Almost twice the number of eHealth users in the private market have adopted online search in the health and lifestyle categories and also spend more time per website than those in the public market. The use of medical websites for specific conditions is almost identical in both markets. The allocation of search effort across categories is similar in both the markets. For all categories, the vast majority of eHealth users only access one website within each category. Those that conduct a search of two or more websites display very narrow search patterns. All users spend relatively little time on eHealth, that is, 3-7 minutes per website.

Conclusions: The proposed online measurement framework exploits online panel data to provide a powerful and objective method of analyzing and exploring eHealth behavior. The private health care system does appear to have an influence on eHealth search behavior in terms of search penetration and time spent per website in the health and lifestyle categories. Two explanations are offered: (1) the personal incentive of medical costs in the private market incentivizes users to conduct online search; and (2) health care information is more easily accessible through health care professionals in the United Kingdom compared with the United States. However, the use of medical websites is almost identical, suggesting that patients interested in a specific condition have a motivation to search and evaluate health information, irrespective of the health care market. The relatively low level of search in terms of the number of websites accessed and the average time per website raise important questions about the actual level of patient informedness in both the markets. Areas for future research are outlined.

(*J Med Internet Res* 2017;19(4):e117) doi:[10.2196/jmir.6739](https://doi.org/10.2196/jmir.6739)

KEYWORDS

health information management; medical informatics; information science

Introduction

Background

The Internet hosts over 100,000 health and medical-related websites, termed as eHealth websites in this study. In the United States, 95 million adults, about 80% of all adult Internet users, have searched for eHealth-related information [1]. Retrieval of eHealth information is one of the most common reasons for accessing the Internet [1] and it is estimated that 4.5% of total Internet searches worldwide are related to eHealth [2,3]. Almost 7 million eHealth searches are conducted on *Google.com* each day [1,4,5]. It has been shown that most online health searches are guided by finding concrete answers to specific questions, for example, to obtain information about treatments, symptoms, diseases, and conditions [6]. In a detailed study of online users of NHS Direct, using survey and detailed interviews, four main motivations to search for health information were found: (1) to obtain reassurance, (2) to seek a second opinion, (3) to improve understanding, and (4) the ease of online search compared with using traditional sources [7]. The scale of health and medical-related online search and the importance of the outcomes in terms of informing and influencing patient behavior mean that it is important to gain a detailed understanding of the online search process [7,8]. *Prima facie*, techniques and approaches used in consumer markets [9] have significant potential in contributing to our understanding of health and medical online search behavior. That is, research concepts and frameworks that have been used in other sectors to measure and evaluate online search behavior can be adapted and used in a health care context.

Although the Internet has greatly reduced consumer search costs, the scale and complexity of the available information requires sophisticated search methods by patients [8,10]. In this context, Google in particular has been described as the “gatekeeper to Web information,” due to its dominance as a search engine, directing searchers to a selected set of websites [11-13]. However, many users still feel that they need help in searching for information [13] and in evaluating the trustworthiness and veracity of online health and medical information [14,15]. These search issues are exacerbated in an eHealth search context because it often concerns an unfamiliar and unknown problem, where the absence of well-defined keywords initiates a series of cognitive learning and reformulating processes [16].

The use of the Internet by patients has the potential to significantly increase patient informedness and is clearly an important part of a health care philosophy of patient-centered medicine, where the patient is placed at the center of decisions and treatments. Such informed patients actively engage in their health management, understand their conditions and diseases, and discuss the medical decision making with their health care professionals. Medical information that was previously exclusive to members of the medical profession is now much more widely available [17,18], a trend that has encountered some professional resistance. There are concerns over the accuracy and validity of the information, the patients’ ability to understand the information, and also to be able to discern the quality of

competing information sources. Given the scale, complexity, and extent of eHealth search, it is therefore important to gain a better understanding of the overall picture of eHealth search by examining a large sample of eHealth websites in terms of their purpose to develop a taxonomic structure and also by evaluating the detailed search patterns by patients or consumers on individual websites.

The following 4 main gaps have been identified in the literature:

First, very little information about the landscape of eHealth websites exists, that is, a high-level view of the most important websites measured by scale in terms of their categorization and usage. This is important because it would inform medical professionals and patients about the structure of website information defined in a taxonomy and would provide valuable insights into the scale and allocation of search effort by patients.

Second, there is very little research into the detailed measurement of eHealth search behavior on a large scale, raising the question of how eHealth search behavior can be analyzed and measured in a systematic and objective manner, using techniques and methods from other online markets [9,10].

Third, how do consumers actually navigate the eHealth landscape? That is, their search patterns have not been established.

Fourth, the differences between online behavior in private (United States) and public (United Kingdom) health care markets have not been investigated. This is an important issue for health policy because it starts to inform the research agenda of how private and public health care systems with different characteristics, for example, restricted versus universal access to health professionals and regulatory differences for advertising, affect online search behavior. These gaps are described in more detail below.

The eHealth Landscape

Previous research has investigated the influence of search engines on the structure of the eHealth landscape [12,13] and surveys reveal the approximate scale of usage in terms of the number of eHealth users and frequency of use [15], but there is very little research on mapping out the structure of the eHealth landscape in terms of different types of eHealth website and their scale. Although there have been some efforts to develop classifications [4,19], there was no attempt to evaluate the models and populate them with a large-scale realistic dataset and there is no agreed classification or taxonomy of health-related websites.

Measurement Framework

Previous research into eHealth behavior can be grouped into two categories: intensive research methods such as focus groups, experiments, and observational studies [3,10,20] and more extensive methods such as surveys [1,5]. Intensive research methods generate a high level of detail and allow the researcher to explore the motivation and logic of particular actions such as search strategies. They are based on relatively small samples of users, typically around $n=20$, which raises the problem of generalizability. This is exacerbated if the sample is of a specific nature; such as patients having a particular condition or being

members of a professional group [3,6]. In addition, such samples are often from a single country [19,21]. Surveys provide potentially much larger samples and better generalizability, but there are also problems, notably, the difficulty of balancing the length and complexity of the survey with the practicalities of users completing it in a comprehensive and accurate manner [22]. There is also a general issue of the accuracy of self-reported behavior in surveys. To overcome these issues in an eHealth context where the purpose of the research is to try and capture highly accurate eHealth search behavior from a very large sample, online panel data are used. Göritz et al [23] explained the use of online panel data in a social science context in an authoritative, general discussion. The methodology, ideas, and concepts used in this paper are adapted from marketing where these techniques have been more widely used [9,24,25]. This enables the accurate measurement of variables such as the number of unique visitors who use a set of websites, the time spent per website, and the breadth of the search process measured by the number of different websites accessed [26,27].

Patient Online Search Patterns

In general, there is very little systematic research and understanding of how consumers actually search for eHealth information in terms of their specific search patterns based on a very large sample of users. Previous research into eHealth search behavior tended to focus on very small groups of users, specific websites, or on the characteristics of the searchers themselves [4,15,28,29]. The bigger picture is therefore largely missing, except from surveys which capture the volume of activity based on self-reported online activities but not the related search behavior across multiple websites, for example, the number of websites visited within a related set of websites or detailed information such as the amount of time spent per website.

Differences Between the Private and Public Health Care Markets

One might reasonably expect privately funded eHealth users to be more engaged in their eHealth activities because of the financial motivation. An additional factor, direct to consumer advertising of drugs is also expected to stimulate patient interest, which would be reflected in more intensive online search patterns, as compared with those of users in a public market. The approach taken in this study is to analyze the US and UK markets as exemplars of private and public health markets. These countries are similar in that both have (1) high levels of Internet penetration, (2) sophisticated use of online technology across all market sectors, and (3) advanced health care systems. The main points of difference are in the nature of the funding, where the United States is predominantly a private system compared with the United Kingdom, which is mainly public [30], and the United States allows direct to consumer advertising of drugs, whereas it is prohibited in the United Kingdom and also in Europe. The two markets are significant from an economic perspective because the United States is the largest pharmaceutical market, and the United Kingdom is part of the second largest pharmaceutical market worldwide, the European Union. Previous studies of online search in health markets found that language strongly affects the behavior [5], warranting the

United Kingdom as the best choice within the European Union. In addition, the vast majority of health content on the Internet, regardless of country, is in English [31].

Research Questions

1. What is the structure of the eHealth landscape described as a taxonomy of eHealth websites?
2. How can eHealth search patterns be analyzed using an objective measurement framework?
3. What are the eHealth search similarities and differences between public and private health care markets?

Methods

Taxonomy

The taxonomy was developed using a combination of conceptual development based on categories from the literature and was then adapted and extended through empirical experimentation with a sample of the largest 100 eHealth websites taken from the comScore commercial database [25,32]. The size of the websites was measured by the number of unique visitors per month, which ranged from 15,000 to 10 million in the United Kingdom and 45,000 to 70 million in the United States.

Online Panel Data and Measurement Framework

The accurate measurement of online behavior requires objective data with international scope, very large scale, detailed granularity, and tracking ability across multiple websites. Online panel clickstream data were used because it fulfills all these data criteria [24-27]. "Clickstream" data recorded electronic data of Internet usage which was collected automatically from a panel of online users [33-35]. Online panel data are more reliable than interviews and questionnaires because it does not rely on self-reporting of Internet usage, and the automatic data collection can be implemented on a very large scale, which gives statistical reliability. It has been widely used in market research [9,24,33-36] and can also be used to make direct comparisons between international markets. In this research, the largest commercial international panel from comScore has been used, which tracks 1.5-trillion digital interactions monthly from more than 2-million registered users in 172 countries who access 3-million websites [25]. The data are gathered through a program installed on the computers of registered users which records the URLs of all pages viewed, how long a window has been active, and the actual pages viewed [26,27]. All data collected are encrypted to ensure complete privacy protection of users, and no personally identifiable information is released. A list of the websites used in the study is provided ([Multimedia Appendix 1](#)).

comScore tracks all types of websites, and the database of websites that are tracked is in effect defined by the users and is entirely dependent on the users' online behavior. The company then categorizes the database, which means that all health- and medical-related websites of significant size are tracked [27]. The more important limitation of the research sample is therefore in the selection of the websites by the researchers. The research sample was an iterative process in which the starting point was to use comScore's own categorization of health and medical websites and then to refine it using the expertise of the

researchers and their colleagues. A small number of additional websites were added, including two important US Government websites. There are two possible sampling errors in the dataset used in this study: the inclusion of irrelevant websites and the exclusion of important websites.

The first error is mitigated by checking and categorizing each website according to the proposed taxonomy. The second type of error cannot be removed altogether although it is mitigated by the large sample size and the sampling process. The large size of the sample means that it is likely that the data are representative of usage and that the results are therefore robust. The statistical significance of differences is also estimated and because of the very large user samples of 1 million in the US and approximately 80,000 in the UK, the measurement errors are very small. Refer to previous studies [1,23,24,36] for examples of how comScore has been used by other researchers and Government bodies. A detailed methodological discussion of the general use of online panels in research that covers panel recruitment, composition, and validity of the data is given in the following core texts [37,38].

A multilevel framework was used to capture different aspects of search. Each measurement is defined in the following conceptual terms.

Search Penetration

The simplest and arguably the most important measurement is search penetration: this is defined as the overall level of search within a defined market compared with the country's Internet population, that is the percentage of the Internet population that accesses eHealth content. Search penetration is reported for each category of eHealth website. It is important because it describes the overall level of market adoption of health-related search and represents the level of interest in health topics.

Allocation of Search Effort Across Categories

The composition of the search is defined as the distribution of users across categories of eHealth websites. It is important because it describes the allocation of search effort across categories and the relative size of each category.

Single Preferred Website and Searcher Split

This measure captures the distinction between those users who look exclusively at one, preferred website and those who search across two or more websites within a category. Those looking at a single website are likely to be accessing information from

their preferred website, whereas those looking at more than one website are actively conducting search based on a consideration set of eHealth websites.

Consideration Set

A consideration set is the number of websites a user visits to conduct his search. Those looking at two or more websites are most likely to be actively searching within a set of related websites, and the consideration set concept is used to measure the range of websites that are accessed [9,24].

Time Per Visitor Per Website

This measure indicates the average length of time spent per website. It is a measure of the intensity of the search process and is a good proxy for search effort or engagement per individual website [34].

Private and Public Index

For each of the variables, search penetration, consideration set, and time per visitor per website, a private and public index is defined as the ratio of the US:UK result. This gives a comparative measure of the development or use of the Internet in a private market compared with a public one.

Results

Taxonomy of eHealth Websites

The proposed taxonomy of websites is shown in Table 1.

The proposed taxonomy is based on the synthesis of individual categories identified in the literature, in particular, health [39], medical [15], lifestyle [40], e-pharmacy [41], social media [6], and professional [42]. This new eHealth taxonomy is broader in scope and more comprehensive than similar frameworks [4,19]. Although the classification scheme proposed by Di Giacomo et al [19] has some similarities to our model, it did not identify the health, lifestyle, and social media categories. The proposed taxonomy has therefore been derived in a mainly deductive manner and then applied empirically to a large sample of websites. The categorization of the empirical data indicates that the taxonomy appears to be comprehensive and useful, that is, the researchers could place all the researched websites naturally into an individual category. Our approach is therefore similar to the methods for developing taxonomies that combine both deductive and empirical approaches [32].

Table 1. Health website categories, their general description, and subcategories.

Category	General description	Subcategories	Number of websites
Health	Broad information about health encompassing a wide variety of information centered on everyday health issues.	General health information Health news Women's health Children's health Men's health Other health information	20
Medical	Websites providing specific information about diseases, conditions, treatments and symptoms.	General medical information Drug information Services, physician and hospital information Specific disease information Mental health and psychology	26
Lifestyle	Information websites centered on well-being and healthy living.	General healthy lifestyle information Natural health and alternative medicine	20
E-pharmacy	Websites providing information and products for medication.	Online pharmacies	20
Social media	Platforms dedicated to communication among patients, peers, and professionals to interact and discuss health and medical issues.	Social media platforms Health forums	4
Professional	Websites providing information specifically targeted to health care professionals.	Information for health professionals Pharmaceutical companies	10

Search Penetration

Online search penetration is an important measure because it is a proxy for the level of interest and awareness in a subject. Online penetration is an adaptation of the standard concept of market penetration from marketing and is an indication of the maturity of the market measured by the product lifecycle [43]. The number of unique visitors and search penetration in each category is shown in Table 2. To take out the effect of market size, the results are calculated as a search penetration percentage of the total online population in each country. The total Internet population is 251 million (US) and 48 million (UK) [44]. The search penetration index is the private (US):public (UK) ratio of the search penetration of eHealth websites in each eHealth category.

Health is the largest category in both markets and the search penetration index is a factor of 2, which is an enormous difference. Lifestyle shows a similar level of difference between the two markets. These two categories are general in nature and demonstrate a much higher level of interest in the private than the public health care market for general health topics. Medical search penetration is identical, and this may reflect the fact that very specific medical search originates from an interaction with a health care professional rather than from general interest in a topic. The level of search into e-pharmacy websites is relatively low in both the countries compared with other categories. The US e-pharmacy penetration is significantly higher than the UK, which is perhaps a reflection of direct to consumer advertising of prescription drugs. The exception to the general pattern is

professional websites, where a much higher proportion of UK users than US users access these websites.

Allocation of Search Effort Across Categories

The search distribution shows the relative size of each category relative to the other categories within each health care market and the results are shown in Table 3.

The distribution of overall search effort is quite similar and shows that health, medical and lifestyle account for around 90% of Internet activity in both markets. The exception is professional, which accounts for a higher proportion of overall UK activity compared to the US.

Single Preferred Website and Searcher Ratio

In general across all categories, most online users only access one website within each category (Table 4). The highest level of search is in health even in which only 22% of all users look at two or more websites. These results are consistent with previously reported levels of search in online consumer markets [34,35,45]. The category with the lowest level of searchers is the e-pharmacy category, which suggests that almost all patients use a single, preferred e-pharmacy. Considering the nature and impact of eHealth content, it is also likely that consumers are risk averse and therefore contain their online activity to a single, trusted website.

Consideration Set

The consideration set is relatively narrow in all cases, particularly when considering the wide choice of websites within

each category (Table 5) and the two markets are almost identical. Narrow search patterns are the norm in other consumer markets [9,24,34,45]. In the health market, it is reasonable to assume

that the complexity of the information naturally limits search patterns because of the effort taken to read and comprehend the information.

Table 2. Search penetration: unique visitors (millions) and search penetration index.

Category	Unique visitors (millions) and search penetration (%)		Search penetration index
	United Kingdom (public HCM ^a)	United States (private HCM)	
Health	19.8 (41%)	207.7 (83%)	2.0
Medical	16.8 (35%)	90.8 (36%)	1.0
Lifestyle	5.4 (11%)	60.8 (24%)	2.1
E-pharmacy	0.7 (1%)	7.1 (3%)	2.0
Social media	0.9 (2%)	14.6 (6%)	3.1
Professional	4.9 (10%)	9.6 (4%)	0.4

^aHCM: health care market.

Table 3. Distribution of search effort.

Category	Public HCM ^a	Private HCM
Health	41%	53%
Medical	35%	23%
Lifestyle	11%	16%
E-pharmacy	1%	2%
Social media	2%	4%
Professional	10%	2%

^aHCM: health care market.

Table 4. Single preferred website and search.

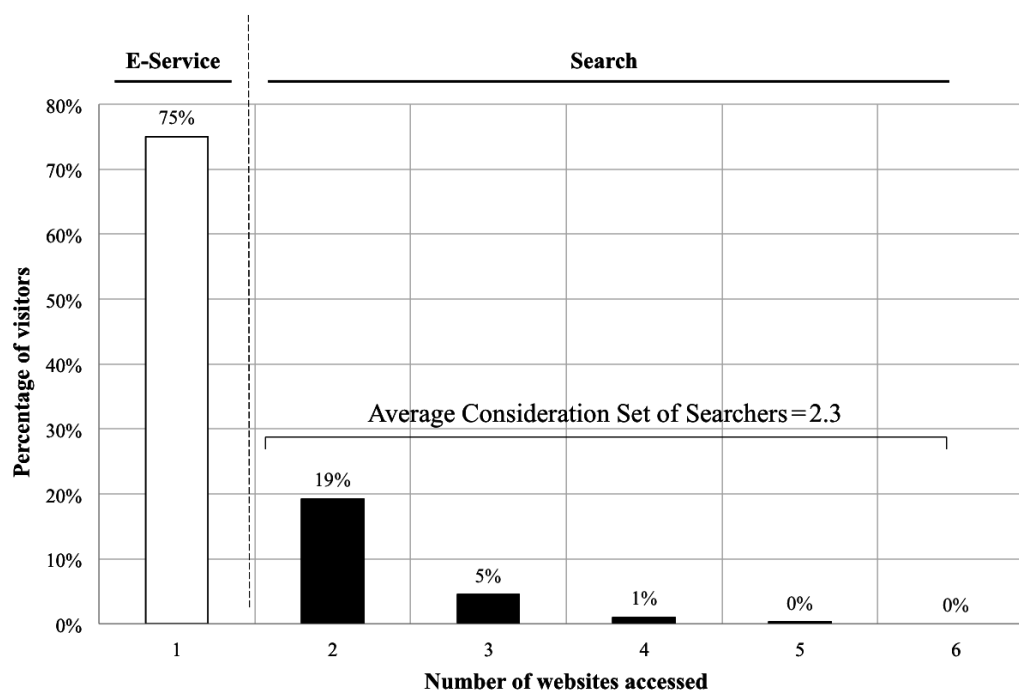
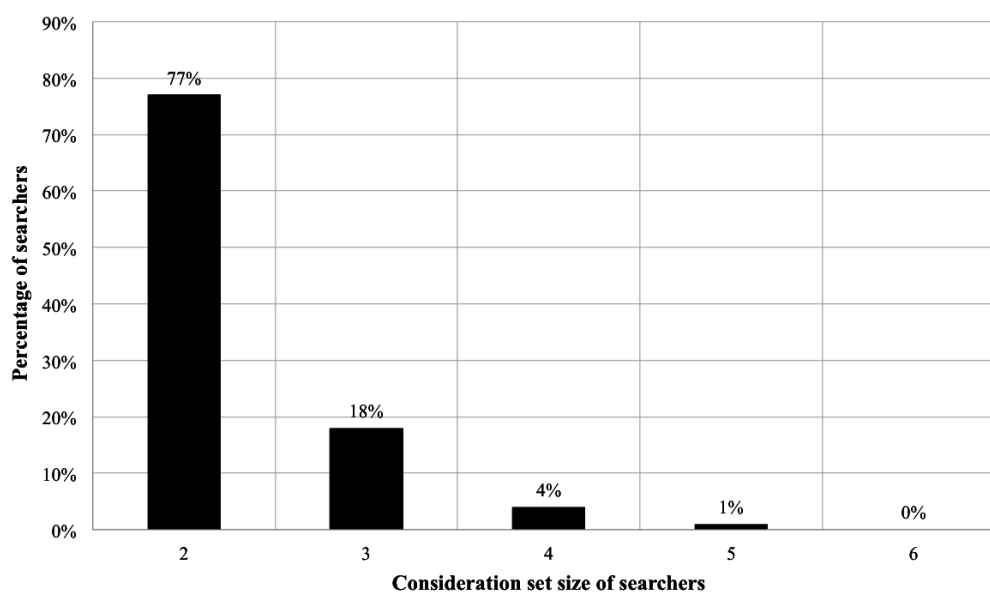
Category	Single website		Search	
	Public HCM ^a	Private HCM	Public HCM	Private HCM
Health	78%	75%	22%	25%
Medical	87%	84%	13%	16%
Lifestyle	88%	85%	12%	15%
E-pharmacy	92%	98%	8%	2%
Social media	91%	93%	9%	7%
Professional	95%	93%	5%	7%

^aHCM: health care market.

Table 5. Results of consideration set.

Category	Public HCM ^a	Private HCM	Consideration set index
Health	2.4	2.3	1.0
Medical	2.2	2.3	1.0
Lifestyle	2.3	2.3	1.0
E-Pharmacy	2.1	2.0	1.0
Social media	2.0	2.1	1.1
Professional	2.1	2.2	1.0

^aHCM: health care market.

Figure 1. The distribution of all visitors in the private (US) health category.**Figure 2.** The distribution of searchers in the private (US) health category, average consideration set size=2.3.

Note that this is an average consideration set. To understand its composition better, the distribution of all visitors is given in [Figure 1](#) and for searchers only is given in [Figure 2](#) for the private (United States) health category.

The distribution of all visitors ([Figure 1](#)) shows that the split between single website and searchers is 75:25%. The distribution of the searchers only ([Figure 2](#)) shows that 95% of searchers look at just two or three websites, with only 5% conducting

what could be termed an extensive search pattern of four or more websites.

Time Per Visitor Per Website

The average time spent per website per unique visitor is a measure of the depth of the search process. It is therefore a measure of engagement or attention between the user and the website content. The results are shown in [Table 6](#).

Patients in the private health care market are more engaged across all categories of eHealth content, and this is possibly due to the financial risk and costs of private health care. A more engaged search process could therefore be interpreted as a risk reduction strategy. Another explanation is that users in the United Kingdom may have better and easier access to health care professionals than in the United States, which gives all users, regardless of income, an alternative to online search.

Sample Size and Statistical Significance

A sample size table was calculated based on the number of visitors in each category, the size of the online panel, and the overall digital population in each country. The estimate of unique visitors is equal to the number of panel members who visit a set of websites (ie, the sample size for each category in the taxonomy) as a proportion of the online panel size multiplied by the digital population. The results for each category are shown in [Table 7](#).

The sample sizes range from 1100 to 828,000. In the United Kingdom, the sample sizes for health, medical, lifestyle, and professional are of the order of 10,000, and 1000 for e-pharmacy and social media. In the United States, the smallest sample is 28,200. With samples of this magnitude, differences between categories are almost certainly real. Search penetration in [Table 2](#) and average time per visitor in [Table 6](#) represent arguably the two most important measures of search effort, that is, the level of usage measured by penetration of the population, which is equivalent to the concept of adoption in marketing, and then average time spent per website per user, which is a measure of search effort or engagement that is independent of market size. The differences between the United Kingdom and the United States for the variables search penetration and average time per website across each of the categories are all significantly different, $P < .001$.

The very low P values can be attributed to the scale of the differences between the samples and the very large sample sizes. It is therefore essential and useful to also consider the effect sizes to give a complete picture of the implications of the results and relate them to the differences in the behavior of digital populations that are independent of the sample size [46]. The effect sizes were measured using odds ratio (OR) for search penetration and Cohen d for time per website [46]. The results are shown in [Table 8](#).

Table 6. Average time per website per unique visitor in minutes.

Category	Public HCM ^a	Private HCM	Time index
Health	3.13	4.36	1.4
Medical	3.14	3.42	1.1
Lifestyle	3.91	4.82	1.2
E-pharmacy	3.20	5.56	1.7
Social media	2.84	3.16	1.1
Professional	3.08	7.31	2.4

^aHCM: health care market.

Table 7. The sample size estimate of each category for the United Kingdom and the United States (represented in thousands).

Category	Public HCM ^a	Private HCM
Health	32.9	827.6
Medical	27.9	361.8
Lifestyle	9.0	242.1
E-pharmacy	1.1	28.2
Social media	1.5	58.3
Professional	8.1	38.4

^aHCM: health care market.

Table 8. Effect sizes for search penetration and time per website per visitor.

Variable	Search penetration OR (95% CI)	Time per website Cohen <i>d</i>
Health	2.01 (1.99-2.04)	0.37
Medical	1.04 (1.02-1.05)	0.02
Lifestyle	2.15 (2.10-2.20)	0.24
E-pharmacy	2.05 (1.93-2.18)	0.82
Social media	3.11 (2.95-3.27)	0.15
Professional	0.38 (0.37-0.39)	0.98

The OR was used to measure the effect size for search penetration because it is a binary outcome, that is, adoption or non-adoption of Internet search within a particular category of eHealth. The effect size interpretation used is small (1.5), medium (2.0), and large (3.0) [46]. For search penetration, the results show a medium effect size for health, lifestyle, and e-pharmacy, a large effect size for social media and a very weak effect size for medical. The effect is in the opposite direction for professional, and in terms of its magnitude is medium. For Cohen *d*, the thresholds of small (0.2), medium (0.5), and large (0.8) are adopted [46]. For time per website, large effect sizes exist in e-pharmacy and professional and small effect sizes in health and lifestyle. The time effect size is very weak for medical and social media.

Health and lifestyle categories are both large categories within the eHealth taxonomy (see Table 3). The effect sizes for search and time in these categories therefore represent important search differences between the US and UK health care markets. The very small effect sizes in the medical category confirm the strong similarity between the UK and US health care markets for medical search.

In e-pharmacy, the results may indicate a higher level of maturity for online ordering of prescriptions in the United States compared with the United Kingdom. The social media category represents just 2% and 4% of overall search effort in the United Kingdom and United States, respectively (see Table 3), although its use could grow as eHealth users become more confident and Web 2.0 technology improves. The effect size for search penetration in the professional category is opposite to the other categories and this requires further research.

Discussion

Principal Findings

A context is needed to measure online behavior in any market. In the case of eHealth, the most suitable framework was to develop a taxonomy to allow grouping of similar websites into health, medical, lifestyle, e-pharmacy, social media, and professional. A measurement framework was also developed to measure search behavior in a holistic manner, which took into account five key aspects of the search process: (1) search penetration, (2) allocation of search effort across categories, (3) distinguish between single website use and search across two or more websites, (4) Consideration set within each category, and (5) time per visitor per website. The application of the

measurement framework in the context of the taxonomy enabled us to compare a private (United States) and a public (United Kingdom) health care market in a much more sophisticated and nuanced manner than simple measures such as the number of online users.

The search penetration indices for health and lifestyle for the private market are twice as high as the public one. On the basis of Internet adoption rates, the private market is therefore much more developed than the public one. For both health and lifestyle, consumers of the private market also spent more time per website. The evaluation of market penetration and time together indicates that users in the private market are generally more interested and also more engaged in eHealth content. Note that these differences cannot be attributed to general differences in Internet sophistication, which are very similar by other measures as reported by authoritative sources [44,47,48]. The exception to search penetration is the professional category where UK Internet use is more extensive in terms of search penetration. This implies that UK users may be more sophisticated in searching this highly specialized content, where the specialized and detailed nature of the information also appears to limit search because 95% of UK and 93% of US users only access one website in this category. These conjectures require further research, which could focus on this category only and perhaps explore the use of a larger number of websites coupled with qualitative data about online search behavior.

The search behavior in the medical category is almost identical, which begs the question why online users in two very different health environments should behave so similarly. One explanation is that medical search is very specific, that is, it relates to a condition or treatment. In this context, it is much more likely that the search is by patients, or their carer, responding and reacting to a specific need for further information and advice on a medical topic. The stimulus for the search process is therefore most likely to be as a result of a medical consultation with a professional, or a patient informing themselves of the likely diagnosis or treatment of an ailment. In this particular medical context, the main factor initiating the search process is likely to be a medical event, which is equally likely in both countries.

The allocation of search across categories has the same rank order, where health, medical, and lifestyle websites attract approximately 90% of online activity in both markets. Social media websites are very small in comparison with the professional websites in health, medical, and lifestyle and may

indicate that eHealth users rely more on professional content rather than peer advice.

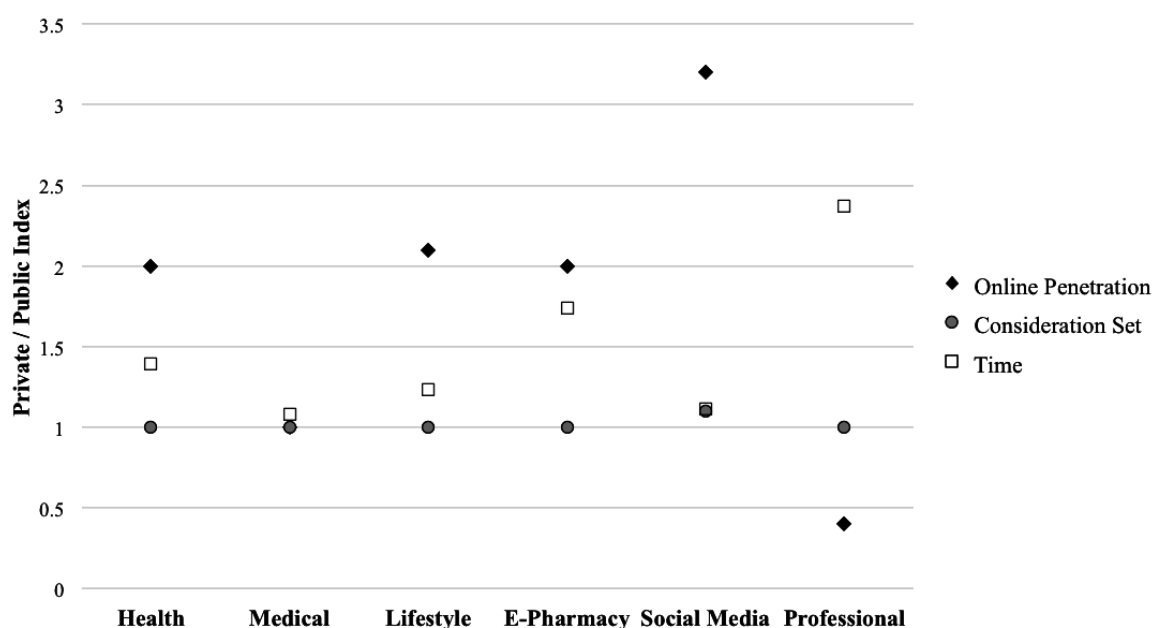
The analysis of single website and search showed that the majority of online users, between 75% and 98%, only look at one website within each category. This implies that they have a preferred, single source of information rather than searching across multiple websites for information. Those who look at two or more websites have relatively narrow search patterns. These results are consistent with other markets [9,35].

The consideration sets are almost identical for both markets. The most interesting result is that the search process is very narrow and it is shown in Figure 1 that even for the largest consideration set of 2.3, 95% of users only look at 2 or 3 websites, that is, only 5% of users conduct an extensive search process. This has important implications because it suggests that most users do not assess multiple websites and rely on a handful of very large, influential websites within each category. This assertion is supported by an inspection of the eHealth website sizes shown in Multimedia Appendix 1, measured by the number of unique visitors over a fixed period of time. The average time per website for each user is just a few minutes,

with private users spending more time on health and lifestyle categories. This indicates a shallow search process and raises important questions about the level of patient informedness.

A synthesis of the results is shown in Figure 3. Each score is a result of the private and public index ratio for each of the measures of online penetration, consideration set size, and the average time spent per website. Online penetration is a direct measure of overall adoption of the Internet; consideration set is a measure of the breadth of search; and the time spent per website is a measure of user interest, or engagement, in the website content. These results show that the private health care market has significantly higher adoption rates for eHealth websites in the health, lifestyle, e-pharmacy, and social media categories measured by online penetration and that they spend significantly more time on health, lifestyle, e-pharmacy, and professional websites. The consideration sets are almost identical for all categories. The medical category is the same for all measures, which suggests that other factors such as interaction with professional medical staff and medical conditions determine the nature of the search process for specific medical information that probably relates to a specific condition or ailment.

Figure 3. Online indices for online penetration, consideration set, and time per website, calculated by dividing private score by public score in each category.



In general, the private US market is more developed in terms of eHealth search than the public UK market. Given that both markets are quite similar in general e-commerce terms [1,44,47], the most plausible explanations are (1) that the private health care system incentivizes personal search into eHealth to encourage general health and well-being; and (2) free access to health care professionals is more widely available in the United Kingdom compared with the United States, including visits to physicians and health care support through telephone.

Limitations and Future Research

The United Kingdom and United States were chosen as representative countries because they are similar in terms of their demographic structure, economy, and Internet access and usage [1,44,47]. Crucially, the two countries share the English language, eliminating effects of language on search behavior. It is therefore plausible that the observed differences can be explained by the differences in funding of the health care systems, where US patients are personally responsible for organizing and paying for their own health care. Other factors

such as the availability and ease of access to physicians were also considered, and both of these factors require further research to understand the complex interactions between the online channel and face-to-face meetings between health care professionals and patients.

The analysis in this study was concerned with the whole of the digital population in each country. Previous research has identified the importance of user characteristics and more broadly the concept of eHealth literacy in influencing search behavior [39,49], and there is potential for developing the analysis in this study by analyzing and exploring the behavior of specific demographic groups and also to make further international comparisons.

Clickstream data provides very detailed information about search patterns, for example, the consideration set concept and the exact amount of time spent per website is captured, which cannot be measured as accurately or on such a large scale using a traditional survey of eHealth trends [5]. These types of results are important for overall eHealth policy making [50] and also have implications for eHealth design [51]. However, a limitation of clickstream data is that the results only measure the actual behavior of users and do not inform us of their motivations, that is, it is not known why they pursue a particular course of action. One way of compensating for this weakness is to conduct a parallel online survey of the panel members that is more qualitative in nature and this is an interesting area for future research.

Comparisons With Prior Work

The proposed taxonomy framework is similar to that of previous study [19], although we think it is important to have more granular categories to uncover important differences in search behavior, which has been supported by the empirical results that identify significant variations and also important similarities in the way that the US and UK health care populations access eHealth information. The diffusion of eHealth adoption is an important topic because it tells us about how the general population is accessing health and medical information [50,52-54]. Research using large surveys of eHealth use [5,53] provides very good general information about the use of particular websites and can be used to estimate the online adoption rates and factors that influence eHealth use within a population. However surveys cannot accurately measure detailed outcomes such as the consideration set concept or time spent per brand, and the approach used here is novel in this respect. The approach taken in this study has been to assess the demand side of eHealth usage combined with the categorization of the supply side of eHealth websites, and this is important because it leads to a better understanding of actual usage that can inform policy [50]. Previous research has shown that a country's health care system and a person's insurance status have little effect on patients' online search behavior for health information [1,55], which is not supported by our results.

Conclusions

The taxonomy of eHealth websites was developed using a combination of deductive and inductive methods and is a useful way of describing and categorizing eHealth websites. The online measurement framework is an important step toward a standard approach to measuring eHealth search behavior, similar to other standard health care measurement systems that are used in health surveys [22]. Online panel data provide a reliable source of data that can be used to conduct standard measurement between different segments of the population, between countries and also to enable longitudinal studies to assess important changes and trends over time.

There are significant statistical differences measured by *P* values and also effect sizes [46] for search penetration and time per website in the health and lifestyle categories. One possible explanation is that patients in the United States have a personal financial incentive because of the private health care system whereas in the United Kingdom, the service is free. Another explanation is that the availability of health care information through health care professionals is more widely available to all patients in the United Kingdom, regardless of income. This explanation is supported by data from the British Medical Association and the US Department of Health and Human Services, which shows that the average number of visits to a physician in the United Kingdom is 6 times per year, twice that of the United States [56,57]. In addition, there is an extensive telephone support system supported by the National Health Service (NHS) in the United Kingdom, which does not have a similar counterpart in the United States. The more general point here is that the availability of health care information through other channels, for example, health care professionals, television and printed media, may reduce the adoption of the Internet channel [58].

The allocation of search effort across the categories is very similar for the private and public markets and show that almost 90% of Internet activity is accounted for by the health, medical, and lifestyle categories. Social media focused eHealth websites are very small and users prefer authoritative content from professionally published websites rather than accessing information from their peer group.

In all categories of eHealth, between 75% and 98% of users access one website only, which suggests that users find a trusted source and remain loyal to it. Consideration sets are relatively small and fall within the range 2.1-2.3. These narrow search patterns may reflect the complexity of the information and perhaps risk-averse search behavior, that is, most searchers evaluate just two or three websites. The time spent per website is also relatively small, although higher in the US. This result suggests that the overall level of informedness will be low, given the low level of search effort measured by time per website and the fact that within each eHealth category most users access just one website and most searchers only visit two or three websites.

Acknowledgments

The authors would like to thank comScore for providing the online panel data used in this study. We would also like to thank Julia A Jacobs for her help in the composition of the dataset and Yu-Lun Liu for comments and advice on the statistics. The authors would also like to acknowledge the contribution of the reviewers with their general comments and feedback on the paper. In particular, we would like to thank one of the reviewers for their suggestion to interpret the results with reference to patient access to non-Internet channels for health information. We would also like to thank the second reviewer for their suggestions regarding methodology. Any errors in the logic or analysis of the arguments remain the sole responsibility of the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

eHeath website classification and unique visitors (represented in thousands; source: ComScore).

[[PDF File \(Adobe PDF File\), 35KB - jmir_v19i4e117_app1.pdf](#)]

References

1. Fox S. Pewinternet. Washington, DC: Pew Internet & American Life Project; 2011 May 12. The social life of health information URL: <http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/> [accessed 2017-03-31] [WebCite Cache ID 6pNZWtFNw]
2. Laugesen J, Hassanein K, Yuan Y. The impact of internet health information on patient compliance: a research model and an empirical study. *J Med Internet Res* 2015;17(6):e143 [FREE Full text] [doi: [10.2196/jmir.4333](https://doi.org/10.2196/jmir.4333)] [Medline: [26068214](https://pubmed.ncbi.nlm.nih.gov/26068214/)]
3. Shuyler KS, Knight KM. What are patients seeking when they turn to the Internet? Qualitative content analysis of questions asked by visitors to an orthopaedics Web site. *J Med Internet Res* 2003 Oct 10;5(4):e24 [FREE Full text] [doi: [10.2196/jmir.5.4.e24](https://doi.org/10.2196/jmir.5.4.e24)] [Medline: [14713652](https://pubmed.ncbi.nlm.nih.gov/14713652/)]
4. Eysenbach G, Kohler C. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. *AMIA Annu Symp Proc* 2003;225-229 [FREE Full text] [Medline: [14728167](https://pubmed.ncbi.nlm.nih.gov/14728167/)]
5. 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](https://doi.org/10.2196/jmir.1023)] [Medline: [19017584](https://pubmed.ncbi.nlm.nih.gov/19017584/)]
6. Adams S, de Bont A, Berg M. Looking for answers, constructing reliability: an exploration into how Dutch patients check web-based medical information. *Int J Med Inform* 2006 Jan;75(1):66-72. [doi: [10.1016/j.ijmedinf.2005.07.036](https://doi.org/10.1016/j.ijmedinf.2005.07.036)]
7. 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](https://doi.org/10.2196/jmir.1600)] [Medline: [21345783](https://pubmed.ncbi.nlm.nih.gov/21345783/)]
8. Seçkin G, Yeatts D, Hughes S, Hudson C, Bell V. Being an informed consumer of health information and assessment of electronic health literacy in a national sample of internet users: validity and reliability of the e-HLS instrument. *J Med Internet Res* 2016 Jul 11;18(7):e161 [FREE Full text] [doi: [10.2196/jmir.5496](https://doi.org/10.2196/jmir.5496)] [Medline: [27400726](https://pubmed.ncbi.nlm.nih.gov/27400726/)]
9. Johnson EJ, Moe WW, Fader PS, Bellman S, Lohse GL. On the depth and dynamics of online search behavior. *Manage Sci* 2004 Mar;50(3):299-308. [doi: [10.1287/mnsc.1040.0194](https://doi.org/10.1287/mnsc.1040.0194)]
10. Fiksdal AS, Kumbamu A, Jadhav AS, Cocos C, Nelsen LA, Pathak J, et al. Evaluating the process of online health information searching: a qualitative approach to exploring consumer perspectives. *J Med Internet Res* 2014;16(10):e224 [FREE Full text] [doi: [10.2196/jmir.3341](https://doi.org/10.2196/jmir.3341)] [Medline: [25348028](https://pubmed.ncbi.nlm.nih.gov/25348028/)]
11. Zhang Y. The effects of preference for information on consumers' online health information search behavior. *J Med Internet Res* 2013;15(11):e234 [FREE Full text] [doi: [10.2196/jmir.2783](https://doi.org/10.2196/jmir.2783)] [Medline: [24284061](https://pubmed.ncbi.nlm.nih.gov/24284061/)]
12. Mager A. Search engines matter: from educating users towards engaging with online health information practices. *Policy Internet* 2012 Jan;4(2):1-21 [FREE Full text] [doi: [10.1515/1944-2866.1166](https://doi.org/10.1515/1944-2866.1166)]
13. 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;17(12):e288 [FREE Full text] [doi: [10.2196/jmir.4345](https://doi.org/10.2196/jmir.4345)] [Medline: [26715363](https://pubmed.ncbi.nlm.nih.gov/26715363/)]
14. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? *J Am Med Assoc* 1998 Feb 25;279(8):611-614. [Medline: [9486757](https://pubmed.ncbi.nlm.nih.gov/9486757/)]
15. 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](https://doi.org/10.1016/j.socscimed.2007.01.012)] [Medline: [17328998](https://pubmed.ncbi.nlm.nih.gov/17328998/)]
16. Zielstorff RD. Controlled vocabularies for consumer health. *J Biomed Inform* 2003 Aug;36(4-5):326-333. [doi: [10.1016/j.jbi.2003.09.015](https://doi.org/10.1016/j.jbi.2003.09.015)]
17. Kontos EZ, Viswanath K. Cancer-related direct-to-consumer advertising: a critical review. *Nat Rev Cancer* 2011;11(2):142-150 [FREE Full text] [doi: [10.1038/nrc2999](https://doi.org/10.1038/nrc2999)]

18. Auton F. The advertising of pharmaceuticals direct to consumers: a critical review of the literature and debate. *Int J Advert* 2004;23(1):5-52. [doi: [10.1080/02650487.2004.11072871](https://doi.org/10.1080/02650487.2004.11072871)]
19. Di Giacomo P, Macerati R. Health websites in Italy: use, classification and international policy. *Med Inform Internet Med* 2002 Sep;27(3):153-160. [doi: [10.1080/1463923021000014112](https://doi.org/10.1080/1463923021000014112)] [Medline: [12507261](#)]
20. Pang PC, Chang S, Verspoor K, Pearce J. Designing health websites based on users' web-based information-seeking behaviors: a mixed-method observational study. *J Med Internet Res* 2016 Jun 06;18(6):e145 [FREE Full text] [doi: [10.2196/jmir.5661](https://doi.org/10.2196/jmir.5661)] [Medline: [27267955](#)]
21. 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](https://doi.org/10.2196/jmir.3117)] [Medline: [25048379](#)]
22. Stewart AL, Hays RD, Ware Jr JE. The MOS short-form general health survey. Reliability and validity in a patient population. *Med Care* 1988 Jul;26(7):724-735. [Medline: [3393032](#)]
23. Göritz AS, Reinhold N, Batinic B. Online Panels. In: *Online social sciences*. Seattle: Hogrefe & Huber Publishers; 2002:47.
24. Zhang J, Fang X, Liu Sheng OR. Online consumer search depth: theories and new findings. *J Manag Inf Syst* 2007 Jan 30;23(3):71-95. [doi: [10.2753/MIS0742-1222230304](https://doi.org/10.2753/MIS0742-1222230304)]
25. Chasin J, Goldstein L. Journalism. Reston, Virginia, United States: comScore; 2014 Mar. ComScore Media Metrix Description of Methodology for the United States: Unified Digital Measurement URL: <http://www.journalism.org/files/2014/03/comScore-Media-Matrix-Description-of-Methodology.pdf> [accessed 2017-03-31] [WebCite Cache ID 6pNba6cft]
26. Cook WA, Pettit RC. Websm.: Advertising Research Foundation; 2009 Apr 06. comScore Media Metrix U.S. Methodology. An ARF research review URL: http://www.websm.org/uploadi/editor/1366314535Cook_2009_comScore-RReview.pdf [accessed 2017-03-31] [WebCite Cache ID 6pNcJ5Qka]
27. ComScore. Comscore. Reston, Virginia, USA: comScore; 2015. ComScore Unified Digital MeasurementTM Methodology URL: http://www.comscore.com/Media/Files/Misc/comScore-Unified-Digital-Measurement-Methodology-PDF?cs_edgescape_cc=US [accessed 2017-03-31] [WebCite Cache ID 6pNfopTRo]
28. Dubowicz A, Schulz PJ. Medical information on the internet: a tool for measuring consumer perception of quality aspects. *Interact J Med Res* 2015;4(1):e8 [FREE Full text] [doi: [10.2196/ijmr.3144](https://doi.org/10.2196/ijmr.3144)] [Medline: [25835333](#)]
29. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *Br Med J* 2002;324(7337):573-577 [FREE Full text] [doi: [10.1136/bmj.324.7337.573](https://doi.org/10.1136/bmj.324.7337.573)] [Medline: [11884321](#)]
30. Chin MH. The patient's role in choice of medications: direct-to-consumer advertising and patient decision aids. *Yale J Health Policy Law Ethics* 2005;5(2):771-784. [Medline: [16052899](#)]
31. AlGhamdi KM, Moussa NA. Internet use by the public to search for health-related information. *Int J Med Inform* 2012 Jun;81(6):363-373. [doi: [10.1016/j.ijmedinf.2011.12.004](https://doi.org/10.1016/j.ijmedinf.2011.12.004)] [Medline: [22217800](#)]
32. Nickerson RC, Varshney U, Muntermann J. A method for taxonomy development and its application in information systems. *Eur J Inf Syst* 2012 Jun 19;22(3):336-359. [doi: [10.1057/ejis.2012.26](https://doi.org/10.1057/ejis.2012.26)]
33. Bucklin RC, Sismeiro C. Click here for internet insight: advances in clickstream data analysis in marketing. *J Interact Market* 2009 Feb;23(1):35-48. [doi: [10.1016/j.intmar.2008.10.004](https://doi.org/10.1016/j.intmar.2008.10.004)]
34. Bucklin RE, Latin JM, Ansari A, Gupta S, Bell D, Coupey E, et al. Choice and the Internet: from clickstream to research stream. *Mark Lett* 2002;13(3):245-258. [doi: [10.1023/A:1020231107662](https://doi.org/10.1023/A:1020231107662)]
35. Lohse GL, Bellman S, Johnson EJ. Consumer buying behavior on the Internet: findings from panel data. *J Interact Market* 2000 Jan;14(1):15-29. [doi: [10.1002/\(SICI\)1520-6653\(200024\)14:1<15::AID-DIR2>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1520-6653(200024)14:1<15::AID-DIR2>3.0.CO;2-C)]
36. Nelson JL, Webster JG. Audience currencies in the age of big data. *Int J Media Mana* 2016 Apr 06;18(1):9-24. [doi: [10.1080/14241277.2016.1166430](https://doi.org/10.1080/14241277.2016.1166430)]
37. Baltagi BH. *Econometric Analysis of Panel Data*. New York: John Wiley & Sons; 2014.
38. Callegaro M, Baker R, Bethlehem J, Göritz A, Krosnick JA, Lavrakas PJ, editors. *Online Panel Research: A Data Quality Perspective*. Chichester: John Wiley & Sons; 2014.
39. 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](https://doi.org/10.1016/j.socscimed.2004.02.020)] [Medline: [15312915](#)]
40. Hodgetts D, Bolam B, Stephens C. Mediation and the construction of contemporary understandings of health and lifestyle. *J Health Psychol* 2005 Jan;10(1):123-136. [doi: [10.1177/1359105305048559](https://doi.org/10.1177/1359105305048559)] [Medline: [15576504](#)]
41. Bessell TL, Anderson JN, Silagy CA, Sansom LN, Hiller JE. Surfing, self-medicating and safety: buying non-prescription and complementary medicines via the internet. *Qual Saf Health Care* 2003 Apr;12(2):88-92 [FREE Full text] [Medline: [12679503](#)]
42. Leckie GJ, Pettigrew KE, Sylvain C. Modeling the information seeking of professionals: a general model derived from research on engineers, health care professionals, and lawyers. *Libr Q* 1996 Apr;66(2):161-193. [doi: [10.1086/602864](https://doi.org/10.1086/602864)]
43. Kotler P, Wong V, Saunders J, Armstrong G. *Principles of marketing*. Upper Saddle River, N.J: Pearson/Prentice Hall; 2008.

44. ComScore. Comscore. Reston, Virginia, USA: comScore; 2016. Media Metrix: Key measures report, worldwide Internet population for US and UK URL: <http://www.comscore.com/Products/Audience-Analytics/Media-Metrix> [accessed 2017-03-31] [WebCite Cache ID 6pNjFDJQc]
45. Holland CP, Jacobs JA, Klein S. The role and impact of comparison websites on the consumer search process in the US and German airline markets. *Inf Technol Tourism* 2016 Jan 4;16(1):127-148. [doi: [10.1007/s40558-015-0037-9](https://doi.org/10.1007/s40558-015-0037-9)]
46. Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. *J Grad Med Educ* 2012 Sep;4(3):279-282. [doi: [10.4300/JGME-D-12-00156.1](https://doi.org/10.4300/JGME-D-12-00156.1)]
47. Reinecke P. ec.europa.eu. 2015. E-commerce statistics for individuals URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics_for_individuals [accessed 2017-03-31] [WebCite Cache ID 6pNgugVGt]
48. Ferlie EB, Shortell SM. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *Milbank Q* 2001;79(2):281-315 [FREE Full text] [Medline: [11439467](https://pubmed.ncbi.nlm.nih.gov/11439467/)]
49. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of health 1.0 and health 2.0 skills. *J Med Internet Res* 2017 Jan 24;19(1):e27 [FREE Full text] [doi: [10.2196/jmir.6709](https://doi.org/10.2196/jmir.6709)] [Medline: [28119275](https://pubmed.ncbi.nlm.nih.gov/28119275/)]
50. Khoja S, Durrani H, Nayani P, Fahim A. Scope of policy issues in eHealth: results from a structured literature review. *J Med Internet Res* 2012 Feb 17;14(1):e34 [FREE Full text] [doi: [10.2196/jmir.1633](https://doi.org/10.2196/jmir.1633)] [Medline: [22343270](https://pubmed.ncbi.nlm.nih.gov/22343270/)]
51. Van Velsen L, Wentzel J, Van Gemert-Pijnen JE. Designing eHealth that matters via a multidisciplinary requirements development approach. *JMIR Res Protoc* 2013;2(1):e21 [FREE Full text] [doi: [10.2196/resprot.2547](https://doi.org/10.2196/resprot.2547)] [Medline: [23796508](https://pubmed.ncbi.nlm.nih.gov/23796508/)]
52. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
53. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. *J Med Internet Res* 2009;11(1):e4 [FREE Full text] [doi: [10.2196/jmir.1035](https://doi.org/10.2196/jmir.1035)] [Medline: [19275980](https://pubmed.ncbi.nlm.nih.gov/19275980/)]
54. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](https://doi.org/10.2196/jmir.1672)] [Medline: [22155738](https://pubmed.ncbi.nlm.nih.gov/22155738/)]
55. Erdem SA. Marketing to health information seekers: a critical look at the principal assumptions. *Am J Health Sci* 2010;1(1):23-28 [FREE Full text] [doi: [10.19030/ajhs.v1i1.989](https://doi.org/10.19030/ajhs.v1i1.989)]
56. British Medical Association. BMA. 2014 Jul. General practice in the UK URL: https://www.bma.org.uk/-/media/files/pdfs/news%20views%20analysis/press%20briefings/pressbriefinggeneralpracticeintheuk_july2014_v2.pdf [accessed 2017-03-31] [WebCite Cache ID 6pNiDnXZi]
57. National Center for Health Statistics. CDC. Atlanta, GA, USA: US Department of Health and Human Services; 2013. National Ambulatory Medical Care Survey: 2013 State and National Summary Tables URL: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2013_namcs_web_tables.pdf [accessed 2017-03-31] [WebCite Cache ID 6pNj31IEN]
58. Jamal A, Khan SA, AlHumud A, Al-Duhyyim A, Alrashed M, Bin SF, et al. Association of online health information-seeking behavior and self-care activities among type 2 diabetic patients in Saudi Arabia. *J Med Internet Res* 2015;17(8):e196 [FREE Full text] [doi: [10.2196/jmir.4312](https://doi.org/10.2196/jmir.4312)] [Medline: [26268425](https://pubmed.ncbi.nlm.nih.gov/26268425/)]

Abbreviations

EU: European Union
HCM: health care market
OR: odds ratio
UK: United Kingdom
US: United States

Edited by G Eysenbach; submitted 04.10.16; peer-reviewed by K Lee, C Fincham; comments to author 14.01.17; revised version received 01.03.17; accepted 14.03.17; published 13.04.17.

Please cite as:

Schneider JA, Holland CP
 eHealth Search Patterns: A Comparison of Private and Public Health Care Markets Using Online Panel Data
J Med Internet Res 2017;19(4):e117
 URL: <http://www.jmir.org/2017/4/e117/>
 doi:[10.2196/jmir.6739](https://doi.org/10.2196/jmir.6739)
 PMID:[28408362](https://pubmed.ncbi.nlm.nih.gov/28408362/)

©Janina Anne Schneider, Christopher Patrick Holland. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

A Psychometric Analysis of the Italian Version of the eHealth Literacy Scale Using Item Response and Classical Test Theory Methods

Nicola Diviani^{1,2}, PhD; Alexandra Lelia Dima^{3,4}, PhD; Peter Johannes Schulz⁵, PhD

¹Department of Health Sciences & Health Policy, Faculty of Humanities and Social Sciences, University of Lucerne, Lucerne, Switzerland

²Person Centered Health Care & Health Communication Group, Human Functioning Unit, Swiss Paraplegic Research, Nottwil, Switzerland

³Amsterdam School of Communication Research, Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, Netherlands

⁴Health Services and Performance Research (HESPER EA 7425), University Claude Bernard Lyon 1, Lyon, France

⁵Institute of Communication and Health, Faculty of Communication Sciences, Università della Svizzera italiana, Lugano, Switzerland

Corresponding Author:

Nicola Diviani, PhD

Department of Health Sciences & Health Policy

Faculty of Humanities and Social Sciences

University of Lucerne

Frohbürgstrasse 3

Lucerne, 6002

Switzerland

Phone: 41 41 229 56 34

Fax: 41 41 229 50 01

Email: nicola.diviani@unilu.ch

Abstract

Background: The eHealth Literacy Scale (eHEALS) is a tool to assess consumers' comfort and skills in using information technologies for health. Although evidence exists of reliability and construct validity of the scale, less agreement exists on structural validity.

Objective: The aim of this study was to validate the Italian version of the eHealth Literacy Scale (I-eHEALS) in a community sample with a focus on its structural validity, by applying psychometric techniques that account for item difficulty.

Methods: Two Web-based surveys were conducted among a total of 296 people living in the Italian-speaking region of Switzerland (Ticino). After examining the latent variables underlying the observed variables of the Italian scale via principal component analysis (PCA), fit indices for two alternative models were calculated using confirmatory factor analysis (CFA). The scale structure was examined via parametric and nonparametric item response theory (IRT) analyses accounting for differences between items regarding the proportion of answers indicating high ability. Convergent validity was assessed by correlations with theoretically related constructs.

Results: CFA showed a suboptimal model fit for both models. IRT analyses confirmed all items measure a single dimension as intended. Reliability and construct validity of the final scale were also confirmed. The contrasting results of factor analysis (FA) and IRT analyses highlight the importance of considering differences in item difficulty when examining health literacy scales.

Conclusions: The findings support the reliability and validity of the translated scale and its use for assessing Italian-speaking consumers' eHealth literacy.

(*J Med Internet Res* 2017;19(4):e114) doi:[10.2196/jmir.6749](https://doi.org/10.2196/jmir.6749)

KEYWORDS

eHealth literacy; eHEALS; item response theory; classical test theory; validation; Italian

Introduction

Health Information on the Web

Following the advent of the Internet, health-related information is increasingly available to the public [1]. It has been estimated that almost 3 out of 4 Internet users worldwide have looked for health information on the Web [2,3]. Wrong or incomplete information could potentially have negative consequences, for instance, on the doctor-patient relationship, participation in screening programs, or adherence to treatments [4]. More attention needs therefore to be devoted to people's ability to interact with Web-based health information.

People's general ability to deal with health information has traditionally been defined as health literacy [5]. To assess health literacy skills in the electronic environment, Norman and Skinner [6,7] have introduced the concept and the measure of eHealth literacy, defined as "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem." According to its authors, the eHealth Literacy Scale (eHEALS) is a promising tool to assess consumers' comfort and skills in using information technology for health and to identify those who may benefit from referrals to eHealth interventions or resources within a clinical environment [7].

Translations of the eHealth Literacy Scale

eHEALS consists of 8 items measuring consumers' combined knowledge, comfort, and perceived skills related to finding, evaluating, and applying electronic health information to health problems. The scale was developed building on the concept of eHealth literacy [7]. According to its authors, eHealth literacy comprises six core skills, or literacies (traditional literacy, health literacy, information literacy, scientific literacy, media literacy, and computer literacy) which, following principles of the social cognitive theory and self-efficacy theory [8], are to be considered precursors of behavior change and skill development [6,7].

The authors of the scale have demonstrated the reliability and validity of its original English version [7]. Over the years, other studies have supported the reliability and validity of eHEALS, for instance, by showing that the scale correlates strongly with—but is distinct from—several scales measuring different aspects of health-related Internet use, such as health information seeking on the Web, attitudes toward the adoption of information and communication technologies (ICTs) for health purposes, use of Internet searching strategies, perceived outcomes of seeking health information by surfing the net, and use of Internet evaluation criteria [7,9,10]. Across these studies, no consistent association of eHEALS scores with the personal characteristics of the respondents, such as gender, education, or age was found. eHEALS has so far been translated and validated in Dutch [11], Japanese [12], Chinese [13], German [14], Spanish [15], Italian [16], Iranian [17], and Hebrew [10,18]. All the linguistic versions of the scale presented high internal consistency measured via Cronbach alpha. These results have generally been taken as an indication of the reliability of the scale in the different cultural contexts.

First Open Question: Population Validity

Most of the studies aimed at validating linguistic versions of eHEALS present at least two important limitations. First of all, whereas the English version of the scale has been applied in a variety of samples, validations have mostly been conducted among specific populations, for instance, students [13-16], young adults [17], patients [11], or seniors [19]. These samples only partly reflect the target population of the tool, that is, consumers of Web-based health information. To date, therefore, it is still not possible to draw general conclusions on the reliability and validity of eHEALS in broader samples.

Second Open Question: Item Difficulty

A second important limitation of past validation studies resides in the widespread reliance on classical test theory (CTT) and factor analysis (FA) only. Like traditional health literacy and other ability tests, the items of eHEALS refer to skills of varying difficulty that may belong to a single eHealth literacy continuum (see [Multimedia Appendix 1](#) for an overview of the items). For instance, whereas knowing how to find health-related information using the Internet (item 1) could be considered a basic skill, being able to distinguish good and bad health information found on the Internet (item 7) could be considered a more advanced skill; as such, a respondent's agreement with item 7 may indicate more intense confidence in their own eHealth literacy than their agreement with item 1. Differences in item difficulty lead to different probabilities of their different response options being endorsed by respondents. This might be the reason behind the different conclusions on the factorial structure of the scale drawn by different authors over the years. Soellner and colleagues [20], for instance, used confirmatory factor analysis (CFA) to compare the 1-factor model based on Norman and Skinner's [7] analyses with a 2-factor model specified a priori based on the content and wording of the items of the scale and on own previous research [14]. The results of their analyses indicated a better fit for the 2-factor model, supporting the division into two subscales: *Information Seeking* (items 1-5 and 8) and *Information Appraisal* (items 6 and 7). CFA was also applied by Neter et al [18] to the Hebrew translation of the scale. Their analyses confirmed the better fit of a 2-factor solution, but two factors were found to include different items (Factor 1: items 1, 2, and 4; Factor 2: items 3 and 5-8). More recently, the exploratory factor analysis (EFA) using principal components analysis (PCA) of an Iranian translation of the scale also suggested a 2-factor structure which groups items 1 and 2 in the first factor and all the other items in a second factor [17]. All the other translations of eHEALS confirmed the 1-factor structure proposed by the authors of the scale, although van der Vaart and colleagues [11] reported in their EFA using PCA, a second component with an eigenvalue of 1.1, which could support the existence of a second dimension.

Psychometrics literature acknowledges that, if items vary in difficulty, PCA might produce a multidimensional solution that groups together items of similar difficulties; although CFA is considered an informative test of structural validity if item properties are known and acceptable, item response theory (IRT) methods are recommended for examining dimensionality in this context [21,22]. A recent psychometric analysis of the original

version of eHEALS used PCA to test unidimensionality although it presented a rather comprehensive parametric IRT exploration of item properties [23]. Yet, scale dimensionality can be appropriately tested within the IRT framework together with several other item properties. Two distinct approaches are available and can be compared for a better understanding of the concept. Nonparametric item response theory (NIRT) (ie, Mokken scale analysis, MSA) allows testing fit to a measurement model arguably most appropriate for eHealth literacy. The concept refers to relative differences between individuals; that is, a person who knows both to find information and to assess its quality is described as having higher literacy than a person who can find information without being able to assess it, but the difference between these two persons is not quantitative in nature. For such concepts, MSA would be a first choice, as it allows to investigate whether an item set measures ordinal differences between respondents regarding a latent trait (ie, ordinal measurement) [23,24]. By comparison, parametric IRT methods aim at a precise quantification of differences. For ordinal items as in eHEALS, the rating scale model (RSM) represents a more stringent set of requirements which, if met at item and respondent level, would represent proof of optimal measurement quality in terms of precision and parsimony, also described as “fundamental measurement” [25]. By testing both IRT models on the Italian version of the eHealth Literacy Scale (I-eHEALS), we can understand in more detail its psychometric properties, what inferences it can support, and what avenues of further development can be pursued for this operationalization of eHealth literacy.

Aims of the Study

As an attempt to get new insights on the two main open questions on eHEALS outlined above, this paper reports on the translation and validation in a population sample of the I-eHEALS. In addition to CTT and FA, we applied MSA to examine dimensionality and model fit, and employed parametric IRT methods to reproduce and extend prior explorations of eHEALS item properties [23,26].

Methods

Overview

In order to explore the psychometric properties of I-eHEALS, a Web-based survey was conducted among a sample of individuals living in the Italian-speaking region of Switzerland. This population is very close to the Italian population from a sociocultural point of view and also as regards health information seeking activities on the Web, as they have access to the same information.

Procedure and Participants

Data were collected through Web-based self-administered questionnaires in two surveys conducted within a larger project in Summer 2013 (Study 1) and Spring 2015 (Study 2). Participants for both surveys were recruited through advertisements placed in the waiting room of a medical private practice, at a local university, as well as in a regional Web-based newspaper. The choice of using different channels for recruitment had the objective to ensure diversity within the

samples regarding age and educational background. The advertisements contained information about the study, contact details of the research team, and a link to the questionnaire. Participants could take part in the survey only once. All participants who completed the survey and agreed to provide contact details (email address or phone number) were entered into a prize draw to win one of three €25 coupons from a local grocery store. A total of 296 individuals ($N_{\text{Study1}}=117$, $N_{\text{Study2}}=179$) aged between 16 and 71 years (mean age 37.37, SD 13.776) comprised the final sample. The sample was predominantly female (193/296, 65.2%), and almost half of the respondents had at least some university education (129/296, 43.6%). The remaining respondents had either a high school diploma (82/296, 27.7%) or a vocational training certificate (55/296, 18.6%).

Whereas almost 9 out of 10 respondents (257/296, 86.8%) reported using the Internet every day, the majority of them (210/296, 70.9%) reported using it for health-related information less than once a week. No significant differences in Web-based health information seeking were observed between Study 1 and Study 2 ($P=.44$).

Participants to Study 1 (mean 33.81, SD 10.466) were slightly younger compared with those in Study 2 (mean 39.77, SD 15.170); $t_{288}=-3.697$, $P<.001$, $d=.46$; however, the two samples did not differ as regards gender ($\chi^2=4.5$, $P=.104$) and educational level ($\chi^2=.7$, $P=.88$).

According to the Swiss Federal Act on Research involving Human Beings (Human Research Act [HRA], September 30, 2011), research not concerning diseases or the structure or the function of the human body does not need formal approval from an ethical review board. All participants were informed about the nature and the aims of the study before enrollment and could decide to withdraw their consent to study participation at any time.

Instrument and Measures

The main section of the surveys was devoted to the 8 items of the I-eHealth Literacy Scale (Multimedia Appendix 1). Like in the English version of the scale, participants were asked to rate their agreement with the statements on a 5-point Likert scale ranging from *Strongly disagree* to *Strongly agree* [7]. The scale underwent a rigorous forward and backward translation process conducted in accordance with the World Health Organization guidelines [27]. In a first step, an Italian-speaking translator fluent in English and knowledgeable of the English-speaking culture translated the items into Italian. In a second step, the items were translated back to English by an independent translator, whose mother tongue was English and who had no knowledge of the questionnaire. The resulting items were compared with the original items by the two translators and the research team to identify possible conceptual differences. Additionally, in order to fully take into account possible cultural differences, in-depth interviews were conducted among 13 individuals considered representative of the target population. The sample of the in-depth interview was composed of 4 men and 9 women aged between 17 and 61 years, with varying levels of education and with different Internet usage habits. All

participants to the interviews were instructed to think aloud during the completion of the questionnaire and to highlight problematic points. The whole process led to some minor changes in wording and confirmed the clarity and comprehensibility of I-eHEALS.

Only in Study 1, data were also collected about the respondents' experiences with and attitudes toward health information seeking on the Web. In particular, data were collected about frequency of Web-based health information seeking, trust in the Internet as a source of health information, attitudes toward the adoption of ICTs for health purposes (2 items, $r=.692$, $P<.001$), use of Internet searching strategies (5 items, Cronbach alpha=.674) [10], perceived outcomes of seeking health information by surfing the net (9 items, Cronbach alpha=.937) [10,28], use of Internet evaluation criteria (5 items, Cronbach alpha=.879) [29], and predisposition toward eHealth in general (2 items, $r=.600$, $P<.001$) [7]. All these constructs are known to be positively related to eHealth literacy and were used to assess the convergent validity of the Italian scale. An overview of the scales used in the study is presented in [Multimedia Appendix 2](#).

Finally, data about selected sociodemographic characteristics of the participants were collected. These included gender, age, educational level, and general and health-related Internet use.

Data Analysis

Statistical analyses were conducted using IBM SPSS Statistics 21.0, R statistical software (R Foundation for Statistical Computing), and Winsteps software (Winsteps, Beaverton, Oregon).

Item Characteristics and Exploratory Factor Analysis

First, item characteristics were described. Factorability of the 8 I-eHEALS items was examined by computing inter-item correlations (all the items should correlate at least .3 with at least one other item), the Kaiser-Meyer-Olkin measure of sampling adequacy (recommended value $>.6$), and Bartlett's test of sphericity (should be significant) [30]. PCA was subsequently conducted to examine the latent variables underlying the observed variables of the Italian scale [31] following previously used methods [11].

Confirmatory Factor Analysis

In a second step, in order to compare the unidimensional solution (Model I) proposed by Norman and Skinner [7] with the 2-factor solution identified by Soellner et al [14] and emerged during PCA of our own data (Model II), model fit indices for the two models were calculated using CFA. Model fit was tested using chi-square tests [32], and the following model fit indices and cutoff values: comparative fit index (CFI) >0.95 , Tucker-Lewis index (TLI) >0.95 , root mean squared error of approximation (RMSEA) <0.06 , and standardized root mean squared residual (SRMR) <0.09 [33]. The two (nested) models were compared using the chi-square difference test (*Anova* function; *Lavaan* package [34]).

Item Response Theory Analyses

Subsequently, NIRT analyses (MSA) were performed using the *Mokken* package in R [35] to examine the structure of the scale

taking into account variations in item difficulty (ie, differences between items in the proportion of answers indicating high ability). MSA examines whether an item set orders respondents accurately on a continuum representing a latent trait by testing unidimensionality (whether items can be located on a single latent continuum in terms of probabilities of respondents endorsing response formats with higher scores), monotonicity (whether the probability of obtaining high scores on an item does not decrease as latent trait scores increase), and local independence (whether associations between items are explained only by their relationship with the construct). If these conditions are met, the items fit the monotone homogeneity model (MHM), and can thus be considered a scale. In the case of polytomous items (as for I-eHEALS), if they also meet a further condition, invariant item ordering (IIO) (ie, items show the same ordering of difficulty across different levels of the latent), the scale allows the identification of "person-free" hierarchy of item difficulty. Therefore, the scale can be used for comparing subgroups regarding their position on the latent trait [36]. Unidimensionality was tested by examining homogeneity—indicating the degree of association between all items (H), and between each item and the item set (H_i)—and by performing an automated item selection procedure (AISP), which is the bottom-up item clustering algorithm performed for increasing homogeneity thresholds [37]. Recommended thresholds for homogeneity (range 0-1) are 0.3 to 0.4 (weak), 0.4 to 0.5 (medium), and over 0.5 (good). Local independence, monotonicity, and IIO were tested via *check.ca*, *check.monotonicity*, and *check.iio* functions; output was examined for significant violations of these assumptions [24,35]. The minimum size of the restscores group (*minsize*) was set at 30 because of the small size of our sample [24]. Person fit was assessed by computing the number of Guttman errors per participant [38].

Additionally, we performed Rasch analyses in line with prior explorations by Nguyen and colleagues [23] and guidelines by Tennant and Conaghan [26] using the Winsteps software. Fit to the RSM was examined, as the 8 items use the same response scale and RSM is a more parsimonious model for this format [39]. We examined item and person infit and outfit against an acceptable mean squares range of 0.6-1.4 and standardized fit statistics of ± 2.0 (Wright and Linacre [40]). Two criteria for good item rating structure were examined: 10 or more observations in each rating category; and outfit mean-squares <2.0 for each category. The hierarchy of item difficulty and the match between person ability and item difficulty (scale targeting) were explored graphically. Person reliability (adequate values $>.85$) and person separation (>2.5) were computed. Differential item functioning (DIF) was examined for differences in item difficulty between groups against a threshold of >0.5 logits for gender, age (dichotomized using median split), education level (college versus no college education), and source of data (Study 1 or 2).

Classical Test Theory Analyses

Reliability of the final scale was assessed using Cronbach alpha [41]. Bivariate correlations and independent samples *t* test were used to assess differences in mean I-eHEALS scores related to gender, age, educational levels, and frequency of Internet use.

Convergent validity of the scale was assessed by computing Pearson correlations between I-eHEALS and other constructs which have been shown to be positively correlated with eHealth literacy in past research, such as attitudes toward eHealth or perceived outcomes of Web-based health information seeking.

Results

Items Characteristics

Participants scored on average 26.65 (SD 6.276) on I-eHEALS. No differences were found among Study 1 (mean 27.21, SD

6.083) and Study 2 (mean 26.27, SD 6.388) participants; $t_{294}=1.261$, $P=.21$. Average scores on the individual I-eHEALS items ranged between 2.75 (SD 1.146, item 8) and 3.62 (SD 0.960, item 2) on a 1 to 5 scale, thus indicating considerable variation in item difficulty. Interitem correlations ranged from $r=.309$ ($P<.001$) to $r=.800$ ($P<.001$). All items except one (*Strongly Disagree* category for item 2, 8 observations) had at least 10 observations for each category. More details on items characteristics and inter-item correlations are presented in [Table 1](#).

Table 1. Descriptive statistics and inter-item correlations for the Italian version of the eHealth Literacy Scale (I-eHEALS) items.

Item	Mean (SD ^a)	Skew	Kurt	Inter-item correlations							
				1	2	3	4	5	6	7	8
1	3.56 (0.996)	-0.59	-0.10	1							
2	3.62 (0.960)	-0.62	0.06	.800	1						
3	3.23 (0.969)	-0.12	-0.45	.601	.614	1					
4	3.40 (0.982)	-0.45	-0.29	.717	.692	.661	1				
5	3.52 (0.978)	-0.60	0.11	.578	.603	.519	.579	1			
6	3.16 (1.171)	-0.28	-0.84	.386	.378	.353	.438	.468	1		
7	3.41 (1.107)	-0.43	-0.49	.356	.372	.309	.406	.407	.719	1	
8	2.75 (1.146)	0.12	-0.88	.445	.461	.450	.488	.513	.529	.530	1

^aSD: standard deviation.

Exploratory and Confirmatory Factor Analyses

All the items correlated at least .3 with at least one other item, suggesting reasonable factorability ([Table 1](#)). The Kaiser-Meyer-Olkin measure of sampling adequacy was .879, above the commonly recommended value of .6, and Bartlett test of sphericity was significant ($\chi^2_{28}=1368.7$, $P<.001$). FA was thus deemed to be suitable with all 8 items. A total of 13 multivariate outliers were identified (Mahalanobis distance >26.125 chi-square threshold for $df=8$, $P=.01$); these were kept

in the dataset to replicate procedures of published factor analyses.

PCA suggested a 2-factor solution with a first factor explaining 57.7% of the variance in I-eHEALS scores, and a second factor explaining an additional 14.9% of the variance (see [Table 2](#) for details). All items presented high factor loadings on Factor 1 (range=.651 to .834), whereas two items presented high factor loadings also on Factor 2 (item 6=.585, item 7=.626). This 2-factor solution mirrors the one proposed by Soellner and colleagues [14].

Table 2. Principal components analysis of the Italian version of the eHealth Literacy Scale (I-eHEALS) items. Factor loadings $<.4$ are not displayed.

Item	Factor 1	Factor 2
I-eHEALS1	.821	
I-eHEALS2	.827	
I-eHEALS3	.751	
I-eHEALS4	.834	
I-eHEALS5	.773	
I-eHEALS6	.683	.585
I-eHEALS7	.651	.626
I-eHEALS8	.717	
Eigenvalue	4.619	1.196
Cumulative explained variance	57.7%	72.7%

CFA was run for two different models: the 1-factor model proposed by Norman and Skinner (Model I) and the 2-factor model proposed by Soellner et al and suggested by our own

PCA (Model II). The comparison of the two models showed better fit for the 2-factorial model ([Table 3](#)). This was confirmed

by the individual model indices and by chi-square differences of 144.8 (df=1; $P<.001$) for Model I versus Model II.

Although Model II appeared to be the one better fitting our data, chi-square tests and fit indices (with the exception of SRMR) indicated a suboptimal model fit for both models.

Table 3. Confirmatory factor analysis of two models of the Italian version of the eHealth Literacy Scale (I-eHEALS).

Model	Chi-square (df) ^a	CFI ^b	TLI ^c	RMSEA ^d	SRMR ^e	AIC ^f	BIC ^g
Model I 1 factor (Norman and Skinner)	247.8 (20)	0.833	0.766	0.196	0.098	5767.609	5826.655
Model II 2 factors (Soellner et al and PCA ^h)	102.9 (19)	0.938	0.909	0.122	0.069	5624.841	5687.577

^aChi-square difference: Model I versus Model II, $\chi^2_1=144.77$; $P<.001$, N=296.

^bCFI: comparative fit index.

^cTLI: Tucker-Lewis index.

^dRMSEA: root mean squared error of approximation.

^eSRMR: standardized root mean squared residual.

^fAIC: Akaike information criterion.

^gBIC: Bayesian information criterion.

^hPCA: principal component analysis.

Nonparametric Item Response Theory: Unidimensionality, Local Independence, Monotonicity, and Invariant Item Ordering

The H_i values of all the items of I-eHEALS and the summary H coefficient of the scale ($H=0.553$, SE 0.032) were above the lower cutoff point of 0.3 (see Table 4). These results confirmed that the I-eHEALS scale can be considered unidimensional, and that all the items measure a single underlying construct as intended.

Exploration of the scale unidimensionality with increasing homogeneity thresholds via *AISP* indicated that at homogeneity

threshold levels of 0.30 to 0.45, all items belonged to the same scale; whereas at a threshold of 0.50, items 6 and 7 clustered together in a separate scale.

Local independence and monotonicity tests suggested no significant violations of these two criteria for any of the items in the dataset, thus confirming that no conditional associations are present between the items except those due to the latent dimension, and that the probability of endorsing response options indicating higher ability increases monotonically for all items as respondents' level of eHealth literacy increases (Figure 1).

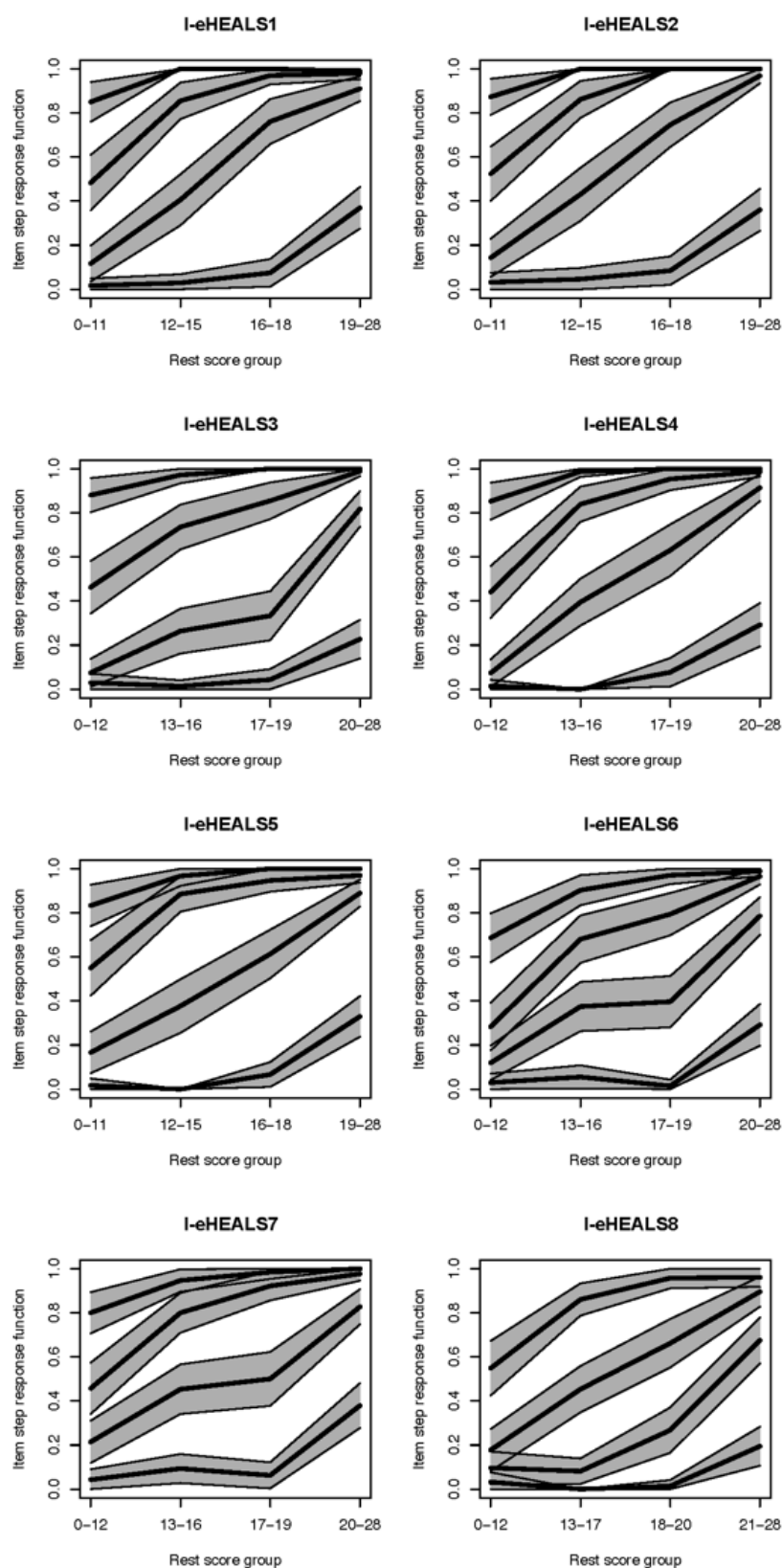
Table 4. Loevinger's scalability coefficients for Italian version of the eHealth Literacy Scale (I-eHEALS) items.

Item	H_i^a	SE ^b
I-eHEALS1	0.585	0.036
I-eHEALS2	0.599	0.033
I-eHEALS3	0.546	0.044
I-eHEALS4	0.604	0.034
I-eHEALS5	0.560	0.039
I-eHEALS6	0.516	0.043
I-eHEALS7	0.486	0.042
I-eHEALS8	0.541	0.040
Scale	H^c	SE
I-eHEALS scale	0.553	0.541

^a H_i : item homogeneity.

^bSE: standard error.

^cH: scale homogeneity.

Figure 1. Italian version of the eHealth Literacy Scale (I-eHEALS) item step response functions.

The results of IIO assessment indicated no significant violations of this assumption. Thus, the items showed the same order of difficulty across levels of the latent construct, with item 2 as easiest and item 8 as most difficult.

A total of 18 participants had a number of Guttman errors at the extreme high end of the distribution (higher than 1.5 interquartile ranges above the third quartile), and were considered outliers.

Parametric Item Response Theory: Item Properties Within the Rasch Framework

The item infit and outfit and standardized t scores are shown in Table 5. All item fit mean squares were within the accepted range. Underfit (both mean squares >1.4 and standardized fit statistics >2) was identified in 28 persons (9.5%), and 45

(15.2%) overfitted the model (both mean squares $<.60$ and standardized fit statistics <-2) according to infit values. These results were largely consistent with results for outfit values.

All categories had 10 or more observations, except the SD category for item 2 (8 observations). Outfit mean-squares for each rating category were within the accepted range.

Table 5. Italian version of the eHealth Literacy Scale (I-eHEALS) item infit and outfit and standardized t scores.

Item	Measure	Model SE ^a	Infit MSQ ^b	ZSTD ^c	Outfit MSQ	ZSTD
I-eHEALS1	-0.49	0.09	0.8411	-1.9392	0.8458	-1.8092
I-eHEALS2	-0.62	0.09	0.782	-2.7092	0.7468	-3.0693
I-eHEALS3	0.22	0.08	0.9237	-0.9291	0.9461	-0.6291
I-eHEALS4	-0.12	0.09	0.7068	-3.8793	0.698	-3.9193
I-eHEALS5	-0.4	0.09	0.8903	-1.3091	0.8549	-1.7091
I-eHEALS6	0.38	0.08	1.3131	3.5213	1.3256	3.5913
I-eHEALS7	-0.15	0.09	1.37	3.9814	1.3018	3.2313
I-eHEALS8	1.18	0.08	1.1172	1.4411	1.1209	1.4511

^aSE: standard error.

^bMSQ: mean-square.

^cZSTD: z-standardized.

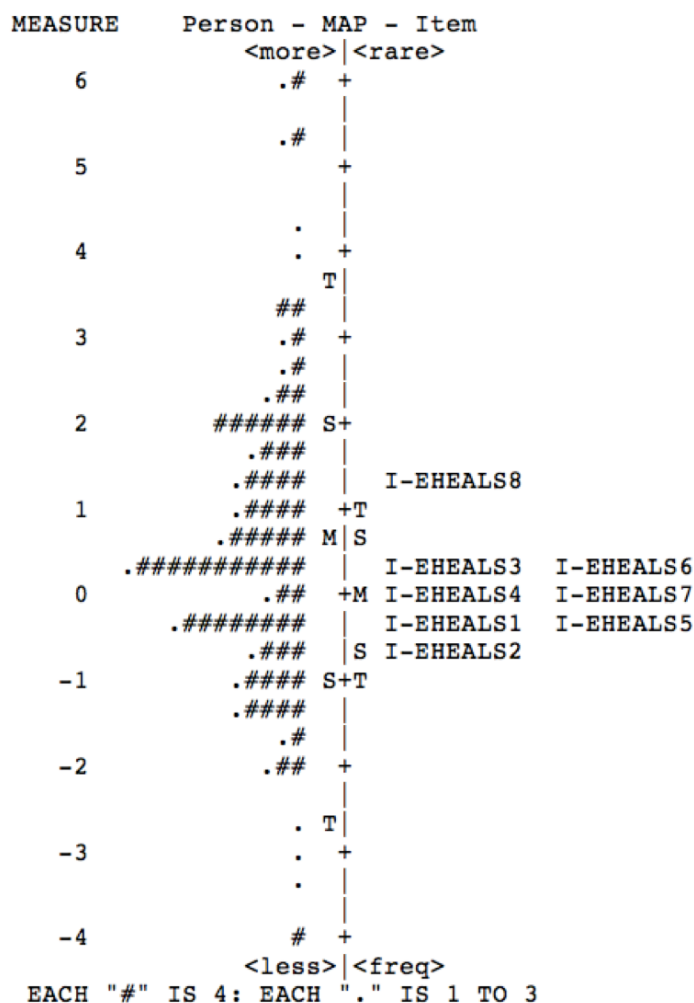
The hierarchy of item difficulty (from the easiest to most difficult—item 2 to item 8) and targeting of items and persons are shown in Figure 2. Most participants were located at above average levels of the eHealth literacy latent, whereas items and item category thresholds were located predominantly close to average values. Thus, I-eHEALS was less able to measure respondents with extreme levels of eHealth literacy. In total, 5 maximum scores and 3 minimum scores were identified, indicating limited ceiling and floor effects.

The real person reliability was .87 (person separation 2.57), indicating good ability to distinguish between respondents of different ability levels.

DIF was identified for item 8 by gender (difficulty higher by 0.59 logits for women; $P<.001$), and for item 7 depending on the source of data (difficulty higher by 0.56 logits in Study 1 than in Study 2; $P=.002$). No differences in item difficulty were present between respondents aged below 33 years or above 33 years, and between participants with college versus no college education.

Classical Test Theory: Reliability and Validity

The final I-eHEALS (mean 26.64, SD 6.276) had excellent reliability (Cronbach alpha=.891). Pearson correlations and t test to assess differences in I-eHEALS scores related to gender, age, educational level, and frequency of Internet use were performed. None of the respondents' characteristics under investigation was found to be significantly associated with I-eHEALS scores (Table 6). Pearson correlations between respondents' I-eHEALS scores and scores on other theoretically correlated constructs showed positive and significant correlations with health information seeking on the Web ($r=.434$, $P<.001$), trust in the Internet as a source of health information ($r=.251$, $P=.006$), attitudes toward the adoption of the ICTs for health purposes ($r=.479$, $P<.001$), eHealth predisposition ($r=.377$, $P<.001$), use of Internet searching strategies ($r=.453$, $P<.001$), perceived outcomes of seeking health information by surfing the net ($r=.577$, $P<.001$), and use of Internet evaluation criteria ($r=.331$, $P<.001$).

Figure 2. Hierarchy of Italian version of the eHealth Literacy Scale (I-eHEALS) item difficulty and targeting of items and persons.**Table 6.** Association of Italian version of the eHealth Literacy Scale (I-eHEALS) scores with theoretically relevant variables (N=117).

Characteristics	I-eHEALS	
	<i>d</i> ^a	<i>P</i>
Gender	.202	.28
	<i>r</i> ^b	<i>P</i>
Age	.076	.42
Educational level	.066	.48
Internet use	.051	.58
Web-based health information seeking	.434	<.001
Trust in the Internet as a source of health information	.251	.006
Attitudes toward the adoption of ICTs ^c for health purposes	.479	<.001
Perceived outcomes	.577	<.001
eHealth predisposition	.377	<.001
Use of searching strategies	.453	<.001
Use of evaluation criteria	.331	<.001

^a*d*: Cohen *d* effect size.^b*r*: Pearson correlation coefficient.^cICT: information and communication technology.

Discussion

Principal Findings

The main aim of this study was to validate I-eHEALS. As previous validation studies did not agree on the factor structure of the scale, particular attention was devoted to the investigation of this aspect. In addition to CTT and factor analyses, IRT techniques were therefore applied to take into account the fact that the items of the scale might differ in terms of difficulty, as it has been recommended for constructs measuring abilities [42]. I-eHEALS scale scores were used to examine group differences and associations with theoretically related concepts.

Although health literacy tools have been examined with other statistical techniques [43,44], most measure developments and validations in the field still rely on less adequate FA methods [45]. Even though some recent studies have already applied IRT to the original eHEALS [23,46], our study was the first to apply these techniques to investigate the structure of a translation of the scale. Also, in contrast with other validations which were conducted in students or in patients' samples, we used a general population sample. Following the results of preliminary exploratory and confirmatory factor analyses, the scale appeared to have a 2-factor structure. This solution had already been proposed in the past by Soellner and colleagues [14], and the existence of a second dimension was also reported for the Dutch eHEALS version [11]. Borrowing terminology widely used in conceptualizations of traditional health literacy [47], the two dimensions suggested by PCA seemed to refer to *functional eHealth literacy* skills (items 1-5 and 8) and *critical eHealth literacy* skills (items 6 and 7), respectively. However, when CFA was performed to compare the fit of model resulting from our PCA and proposed by Soellner and colleagues [14] and of the single-factor model originally proposed by the authors of the scale [7], neither model showed an adequate fit to the data.

The dimensionality of the scale was therefore subsequently assessed using nonparametric and parametric IRT methods, which take into account the fact that the items might differ in terms of difficulty. Such an approach has been recommended for constructs measuring abilities like eHEALS [42]. Mokken analyses showed that the Italian eHEALS version can be considered unidimensional, and that all the items measure a single underlying construct with good homogeneity, in line with what was originally proposed by the authors of the scale [7]. Moreover, our analyses showed that the data fit the MHM and also meet the additional assumption of IIO. There are three key implications of this result for the applicability of the scale (see [24] for a theoretical overview). First, I-eHEALS items can be used to order respondents with respect to their latent eHealth literacy levels based on the scale score, thus justifying the use of mean scores for further analyses. Simply examining reliability via Cronbach alpha is not by itself sufficient to allow the use of mean or sum scores [48], and a confirmation of unidimensionality and monotonicity is necessary before considering its use as an indicator of reliability [49]. Second, fitting the MHM model implies that the test is able to order respondents on the latent measurement continuum in a similar way if different subsets of items are used (and thus achieving

item-free measurement). This suggests that the items are a good starting point for developing a larger item pool, from which alternative questionnaire versions can be developed, for instance, for repeated assessments in longitudinal studies. Third, IIO implies that the items target eHealth literacy skills that form the same hierarchy for all respondents (ie, ordering holds at individual level as well). This allows the use of I-eHEALS to compare subgroups of citizens with different levels of eHealth literacy, which is a common aim of health literacy research [50].

The stricter Rasch analyses additionally allowed us to conclude that I-eHEALS has a good ability to distinguish between respondents of different ability levels and that only a few differences in item difficulty were present between male and female respondents, whereas no such differences were found for younger versus older participants or between participants with college versus no college education. Rasch analyses also indicated that the scale is less able to measure respondents with extreme levels of eHealth literacy. Interestingly, the order of difficulty of our items was different from that identified in other studies using IRT on eHEALS [23,46], thus suggesting that some personal characteristics might play a role in the kind of eHealth literacy tasks people perceive as more or less demanding. Participants in our sample rated item 2 as the easiest and item 8 as the most difficult. The Health Science college students in the study by Nguyen et al [23] perceived item 4 to be the easiest item and item 5 to be the hardest item. Within the same study, participants recruited from Amazon MTurk rated item 7 as the easiest item and item 6 as the most difficult one. As acknowledged by Nguyen et al, these differences could be attributed to the demographic makeup of each sample group: Health Science students may be more familiar with the location of health resources on the Internet, whereas tech-savvy and highly educated MTurkers might have higher perceptions of their ability to distinguish high-quality health information versus low-quality health information. We used a general population sample (participants were neither health nor technology experts) and it is therefore reasonable that our participants perceived different tasks as being more or less demanding. We strongly encourage future research to investigate these aspects in more depth.

Consistently with other translated versions of eHEALS, no significant correlations of I-eHEALS scores with respondents' characteristics were found. Although a rigorous test of this relationship (or lack thereof) would require a more diverse sample as regards age and education, the absence of an association with the traditional determinants of health literacy might be seen as a further indication that eHEALS (and its translations) is not able to capture actual skills. This hypothesis would be supported by a study conducted in two Dutch populations by van der Vaart and colleagues [11]. The authors found no association between scores on eHEALS and an actual Internet performance test, questioning the ability of the instrument to adequately capture the phenomenon under investigation. In this view—as it has been argued in the past—eHEALS could be more realistically described as a measure of self-efficacy in the electronic health information context [11,51]. As suggested by Frisch and colleagues [52], this is a common shortcoming of self-reported measures of

health literacy. In our view, however, this does not undermine the value of eHEALS. According to one of the eHEALS authors, the nonsignificant correlation could be related to the fact that the scale in its present form does not capture the skills related to the use of social media, which have become more and more important in the last few years. In this perspective, eHEALS can still be considered a valid tool for assessing competency with Web 1.0 technologies [53]. Higher I-eHEALS scores were indeed shown to be significantly associated with more frequent Web-based health information seeking, higher trust in the Internet as a source of health information, more positive attitudes toward the adoption of ICTs for health purposes, higher eHealth predisposition, and more positive perceived outcomes of seeking health information by surfing the net. These associations suggest that the scale can safely be used to assess consumers' perceived comfort and skills in using information technology for health. This would be useful to identify those who may be keener to participate in eHealth interventions or use eHealth resources within a clinical environment and those who are in need of more support.

Limitations

Three limitations of this study have to be acknowledged. First, our sample was younger, more educated, and included a higher percentage of women compared with the general population of the Italian-speaking region of Switzerland [54]. If we consider, however, that these are the characteristics that are usually associated with health information seeking on the Web (eg, [4]), we believe that our sample is suitable to provide us with an adequate snapshot of our population of interest. Yet, being a convenience sample, it cannot be considered truly representative of the population, thus limiting the generalizability of our results. Second, our sample was relatively small compared with other validations of translated version of eHEALS. However, NIRT can cope with small sample sizes better than other statistical

techniques [21]. We are therefore confident that sample size had a limited impact on our results, particularly given the small number of items investigated. Finally, we did not include a measure of actual ability to perform eHealth literacy tasks. However, as our goal was to provide a good set of items for investigating eHealth literacy in Italian-speaking populations, this was outside the scope of this study. Nevertheless, we join other scholars in acknowledging the need for more research specifically aimed at further investigating the link between actual and perceived ability to perform eHealth literacy tasks—as it was done by van der Vaart and colleagues [11] in the Netherlands. Only after doing that, it will be possible to fully capture the complexity of the phenomenon under investigation.

Conclusions and Practice Implications

The study confirmed that I-eHEALS is a reliable and valid tool to assess Italian-speaking consumers' perceived comfort and skills in using information technology for health. A previous validation of I-eHEALS among students had already proven the suitability of the scale among that specific population [16]. The sample used in our study allows us to extend this conclusion to the general population. I-eHEALS can therefore safely be used by public health officials and health care providers to identify those who are most ready to take part in eHealth interventions or to use Web-based resources within both clinical and nonclinical environments, as well as those who would need more support. Moreover, compared with the previous Italian validation, our IRT analysis was also able to highlight several strengths of the scale, for instance, its unidimensionality which justifies the calculation of a total mean score of all the items. Moreover, it indicated future directions in eHealth literacy assessment, such as the importance of wording in item development, and the possibility of extending the item pool and developing alternative versions.

Acknowledgments

The authors wish to thank Sara Palazzo for her help in the forward and backward translation of eHEALS scale into Italian, the cognitive testing of the instrument, and the data collection. Study 1 was supported by internal funding of the Institute of Communication and Health of the University of Lugano. Study 2 was supported by a personal grant awarded to the first author by the Swiss National Science Foundation (grant number P2TIP1_148345). The funding sources had no role in study design, data collection, analysis, interpretation, or writing of the report.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The Italian version of the eHealth Literacy Scale (I-eHEALS).

[PDF File (Adobe PDF File), 27KB - [jmir_v19i4e114_app1.pdf](#)]

Multimedia Appendix 2

Overview of the scales used in the study.

[PDF File (Adobe PDF File), 29KB - [jmir_v19i4e114_app2.pdf](#)]

References

1. Viswanath K. Science and society: the communications revolution and cancer control. *Nat Rev Cancer* 2005 Oct;5(10):828-835. [doi: [10.1038/nrc1718](https://doi.org/10.1038/nrc1718)] [Medline: [16195753](https://pubmed.ncbi.nlm.nih.gov/16195753/)]
2. European Commission. European citizen's digital health literacy. 2014. URL: http://ec.europa.eu/public_opinion/flash/fl_404_en.pdf [accessed 2017-03-27] [WebCite Cache ID 6pHGDuQj2]
3. Fox S, Duggan M. Health Online 2013. 2013. URL: <http://bibliobase.sermis.pt:8008/BiblioNET/Upload/PDF5/003820.pdf> [accessed 2017-03-27] [WebCite Cache ID 6pHGJQjc0]
4. 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 [FREE Full text] [Medline: [11780707](https://pubmed.ncbi.nlm.nih.gov/11780707/)]
5. Nutbeam D. Health promotion glossary. *Health Promot* 1986 May;1(1):113-127. [Medline: [10318625](https://pubmed.ncbi.nlm.nih.gov/10318625/)]
6. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006 Jun;8(2):e9 [FREE Full text] [doi: [10.2196/jmir.8.2.e9](https://doi.org/10.2196/jmir.8.2.e9)] [Medline: [16867972](https://pubmed.ncbi.nlm.nih.gov/16867972/)]
7. 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](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
8. Bandura A. Self-efficacy: the exercise of control. New York: W.H. Freeman; 1997.
9. 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](https://doi.org/10.2196/jmir.2645)] [Medline: [23639979](https://pubmed.ncbi.nlm.nih.gov/23639979/)]
10. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. *J Med Internet Res* 2012 Jan;14(1):e19 [FREE Full text] [doi: [10.2196/jmir.1619](https://doi.org/10.2196/jmir.1619)] [Medline: [22357448](https://pubmed.ncbi.nlm.nih.gov/22357448/)]
11. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth literacy scale (eHEALS) measure what it intends to measure? Validation of a dutch version of the eHEALS in two adult populations. *Journal of Medical Internet Research*. *J Med Internet Res* 2011;13(4):e86. [Medline: [22071338](https://pubmed.ncbi.nlm.nih.gov/22071338/)]
12. Mitsutake S, Shibata A, Ishii K, Okazaki K, Oka K. [Developing Japanese version of the eHealth Literacy Scale (eHEALS)]. *Nihon Koshu Eisei Zasshi* 2011 May;58(5):361-371. [Medline: [21905612](https://pubmed.ncbi.nlm.nih.gov/21905612/)]
13. Koo M, Norman CD, Chang H. Psychometric evaluation of a chinese version of the eHealth literacy scale (eHEALS) in school age children. *Int Electron J Health Educ* 2012;15:29-36.
14. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement: German translation of the eHEALS. *Journal of Media Psychology Theories Methods and Applications* 2014 Jan;26(1):29-38. [doi: [10.1027/1864-1105/a000104](https://doi.org/10.1027/1864-1105/a000104)]
15. Paramio Pérez G, Almagro BJ, Hernando Gómez A, Aguaded Gómez JI. [Validation of the eHealth Literacy Scale (eHEALS) in Spanish University Students]. *Rev Esp Salud Publica* 2015;89(3):329-338 [FREE Full text] [doi: [10.4321/S1135-57272015000300010](https://doi.org/10.4321/S1135-57272015000300010)] [Medline: [26388346](https://pubmed.ncbi.nlm.nih.gov/26388346/)]
16. De Caro W, Corvo E, Marucci AR, Mitello L, Lancia L, Sansoni J. eHealth literacy scale: a nursing analysis and Italian validation. *Stud Health Technol Inform* 2016;225:949. [Medline: [27332425](https://pubmed.ncbi.nlm.nih.gov/27332425/)]
17. Bazm S, Mirzaei M, Fallahzadeh H, Bazm R. Validity and reliability of Iranian version of eHealth literacy scale. *Journal of Community Health Research* 2016;5(2):121-130 [FREE Full text]
18. Neter E, Brainin E, Baron-Epel O. The dimensionality of health literacy and eHealth literacy. *European Health Psychologist* 2015;17(6):275-280.
19. Chung SY, Nahm ES. Testing reliability and validity of the eHealth Literacy Scale (eHEALS) for older adults recruited online. *Comput Inform Nurs* 2015 Apr;33(4):150-156. [doi: [10.1097/CIN.0000000000000146](https://doi.org/10.1097/CIN.0000000000000146)] [Medline: [25783223](https://pubmed.ncbi.nlm.nih.gov/25783223/)]
20. Soellner R, Huber S, Lenartz N, Rudinger G. [Facets of health competence: an expert consultation: project health competence]. *Zeitschrift für Pädagogik* 2010;56:104-114.
21. van Schuur WH. Mokken scale analysis: between the Guttman scale and parametric item response theory. *Polit Anal* 2017 Jan 4;11(02):139-163. [doi: [10.1093/pan/mpg002](https://doi.org/10.1093/pan/mpg002)]
22. Christensen KB, Engelhard G, Salzberger Jr T. Ask the experts: Rasch vs. factor analysis. *Rasch Measurement Transactions* 2012;26(3):1373-1378 [FREE Full text]
23. 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](https://doi.org/10.2196/publichealth.4967)] [Medline: [27244771](https://pubmed.ncbi.nlm.nih.gov/27244771/)]
24. Stochl J, Jones PB, Croudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Med Res Methodol* 2012;12:74 [FREE Full text] [doi: [10.1186/1471-2288-12-74](https://doi.org/10.1186/1471-2288-12-74)] [Medline: [22686586](https://pubmed.ncbi.nlm.nih.gov/22686586/)]
25. Bond T, Fox CM. Applying the Rasch model: Fundamental measurement in the human sciences. London: Routledge; 2015.
26. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 2007 Dec 15;57(8):1358-1362 [FREE Full text] [doi: [10.1002/art.23108](https://doi.org/10.1002/art.23108)] [Medline: [18050173](https://pubmed.ncbi.nlm.nih.gov/18050173/)]
27. World Health Organization. WHO. Geneva: World Health Organisation; 2010. Process of translation and adaptation of instruments URL: http://www.who.int/substance_abuse/research_tools/translation/en/ [accessed 2017-03-27] [WebCite Cache ID 6pHJa711B]

28. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. *J Am Med Assoc* 2003 May 14;289(18):2400-2406. [doi: [10.1001/jama.289.18.2400](https://doi.org/10.1001/jama.289.18.2400)] [Medline: [12746364](https://pubmed.ncbi.nlm.nih.gov/12746364/)]
29. Barnes MD, Penrod C, Neiger BL, Merrill RM, Thackeray R, Eggett DL, et al. Measuring the relevance of evaluation criteria among health information seekers in the internet. *J Health Psychol* 2003 Jan;8(1):71-82. [doi: [10.1177/1359105303008001436](https://doi.org/10.1177/1359105303008001436)] [Medline: [22113902](https://pubmed.ncbi.nlm.nih.gov/22113902/)]
30. Dziuban CD, Shirkey EC. When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin* 1974;81(6):358-361. [doi: [10.1037/h0036316](https://doi.org/10.1037/h0036316)]
31. Park HS, Dailey R, Lemus D. The use of exploratory factor analysis and principal components analysis in communication research. *Human Comm Res* 2002 Oct;28(4):562-577. [doi: [10.1111/j.1468-2958.2002.tb00824.x](https://doi.org/10.1111/j.1468-2958.2002.tb00824.x)]
32. Thompson B. Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association; 2004.
33. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999 Jan;6(1):1-55. [doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)]
34. Rosseel Y. Lavaan: an R Package for Structural Equation Modeling. *J Stat Soft* 2012;48(2):1-36. [doi: [10.18637/jss.v048.i02](https://doi.org/10.18637/jss.v048.i02)]
35. Van der Ark LA. Mokken Scale Analysis in R. *J Stat Soft* 2007;20(11):1-19. [doi: [10.18637/jss.v020.i11](https://doi.org/10.18637/jss.v020.i11)]
36. Sijtsma K, Meijer RR, van der Ark LA. Mokken scale analysis as time goes by: an update for scaling practitioners. *Pers Individ Dif* 2011 Jan;50(1):31-37. [doi: [10.1016/j.paid.2010.08.016](https://doi.org/10.1016/j.paid.2010.08.016)]
37. Hemker BT, Sijtsma K, Molenaar IW. Selection of unidimensional scales from a multidimensional item bank in the polytomous mokken IRT model. *Appl Psychol Meas* 1995 Dec 01;19(4):337-352. [doi: [10.1177/014662169501900404](https://doi.org/10.1177/014662169501900404)]
38. Meijer RR, Niessen AS, Tendeiro JN. A practical guide to check the consistency of item response patterns in clinical research through person-fit statistics: examples and a computer program. *Assessment* 2016 Feb;23(1):52-62. [doi: [10.1177/1073191115577800](https://doi.org/10.1177/1073191115577800)] [Medline: [25804439](https://pubmed.ncbi.nlm.nih.gov/25804439/)]
39. Linacre J. Comparing Partial Credit Models (PCM) and Rating Scale Models (RSM). *Rasch Measurement Transactions* 2000;14(3):768 [FREE Full text]
40. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measurements Transactions* 1994;8(3):370 [FREE Full text]
41. Kline P. A handbook of test construction: Introduction to psychometric design. London: Methuen; 1986.
42. Embretson SE, Hershberger SL, editors. The new rules of measurement: what every psychologist and educator should know. Mahwah, NJ: L Erlbaum Associates; 1999.
43. Bann CM, McCormack LA, Berkman ND, Squiers LB. The health literacy skills instrument: A 10-item short form. *J Health Commun* 2012;17(Suppl 3):191-202. [doi: [10.1080/10810730.2012.718042](https://doi.org/10.1080/10810730.2012.718042)] [Medline: [23030570](https://pubmed.ncbi.nlm.nih.gov/23030570/)]
44. McCormack L, Bann C, Squiers L, Berkman ND, Squire C, Schillinger D, et al. Measuring health literacy: a pilot study of a new skills-based instrument. *J Health Commun* 2010;15(Suppl 2):51-71. [doi: [10.1080/10810730.2010.499987](https://doi.org/10.1080/10810730.2010.499987)] [Medline: [20845193](https://pubmed.ncbi.nlm.nih.gov/20845193/)]
45. Nguyen TH, Paasche-Orlow MK, Kim MT, Han H, Chan KS. Modern measurement approaches to health literacy scale development and refinement: overview, current uses, and next steps. *J Health Commun* 2015;20(Suppl 2):112-115. [doi: [10.1080/10810730.2015.1073408](https://doi.org/10.1080/10810730.2015.1073408)] [Medline: [26513038](https://pubmed.ncbi.nlm.nih.gov/26513038/)]
46. 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 Feb;100(2):320-326. [doi: [10.1016/j.pec.2016.09.008](https://doi.org/10.1016/j.pec.2016.09.008)] [Medline: [27658660](https://pubmed.ncbi.nlm.nih.gov/27658660/)]
47. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int* 2000;15(3):259-267. [doi: [10.1093/heapro/15.3.259](https://doi.org/10.1093/heapro/15.3.259)]
48. Sijtsma K. On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika* 2009 Mar;74(1):107-120 [FREE Full text] [doi: [10.1007/s11336-008-9101-0](https://doi.org/10.1007/s11336-008-9101-0)] [Medline: [20037639](https://pubmed.ncbi.nlm.nih.gov/20037639/)]
49. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol* 1993;78(1):98-104. [doi: [10.1037/0021-9010.78.1.98](https://doi.org/10.1037/0021-9010.78.1.98)]
50. Nguyen TH, Park H, Han H, Chan KS, Paasche-Orlow MK, Haun J, et al. State of the science of health literacy measures: validity implications for minority populations. *Patient Educ Couns* 2015 Jul 29;1492-1512 [FREE Full text] [doi: [10.1016/j.pec.2015.07.013](https://doi.org/10.1016/j.pec.2015.07.013)] [Medline: [26275841](https://pubmed.ncbi.nlm.nih.gov/26275841/)]
51. Diviani N, van den Putte B, Meppelink CS, van Weert JC. Exploring the role of health literacy in the evaluation of online health information: insights from a mixed-methods study. *Patient Educ Couns* 2016 Jun;99(6):1017-1025. [doi: [10.1016/j.pec.2016.01.007](https://doi.org/10.1016/j.pec.2016.01.007)] [Medline: [26817407](https://pubmed.ncbi.nlm.nih.gov/26817407/)]
52. Frisch A, Camerini L, Diviani N, Schulz PJ. Defining and measuring health literacy: how can we profit from other literacy domains? *Health Promot Int* 2012 Mar;27(1):117-126. [doi: [10.1093/heapro/dar043](https://doi.org/10.1093/heapro/dar043)] [Medline: [21724626](https://pubmed.ncbi.nlm.nih.gov/21724626/)]
53. Norman CD. eHealth literacy 2.0: problems and opportunities with an evolving concept. *J Med Internet Res* 2011 Dec;13(4):e125 [FREE Full text] [doi: [10.2196/jmir.2035](https://doi.org/10.2196/jmir.2035)] [Medline: [22193243](https://pubmed.ncbi.nlm.nih.gov/22193243/)]
54. Ufficio di Statistica. Annuario Statistico Ticinese. 2015. URL: http://www3.ti.ch/DFE/DR/USTAT/allegati/volume/161441annuario_2015_20150312.pdf [accessed 2017-03-27] [WebCite Cache ID 6pHI5syfd]

Abbreviations

AIC: Akaike information criterion
AISP: automatic item selection procedure
BIC: Bayesian information criterion
CFI: comparative fit index
CFA: confirmatory factor analysis
CTT: classical test theory
DIF: differential item functioning
DMM: Double Monotonicity Model
EFA: exploratory factor analysis
eHEALS: eHealth Literacy Scale
FA: factor analysis
HRI: Human Research Act
ICT: information and communication technology
I-eHEALS: Italian version of the eHealth Literacy Scale
IIO: invariant item ordering
IRT: item response theory
MHM: monotone homogeneity model
MSA: Mokken scale analysis
NIRT: nonparametric item response theory
PCA: principal component analysis
RMSEA: root mean squared error of approximation
RSM: rating scale model
SRMR: standardized root mean squared residual
TLI: Tucker-Lewis index

Edited by G Eysenbach; submitted 06.10.16; peer-reviewed by E Neter, L Sudbury-Riley, D Walker; comments to author 10.11.16; revised version received 02.12.16; accepted 11.02.17; published 11.04.17.

Please cite as:

Diviani N, Dima AL, Schulz PJ

A Psychometric Analysis of the Italian Version of the eHealth Literacy Scale Using Item Response and Classical Test Theory Methods
J Med Internet Res 2017;19(4):e114

URL: <http://www.jmir.org/2017/4/e114/>

doi: [10.2196/jmir.6749](https://doi.org/10.2196/jmir.6749)

PMID: [28400356](https://pubmed.ncbi.nlm.nih.gov/28400356/)

©Nicola Diviani, Alexandra Lelia Dima, Peter Johannes Schulz. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 11.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Impact of Information and Communication Technologies on Nursing Care: Results of an Overview of Systematic Reviews

Geneviève Rouleau^{1,2}, RN, PhD(c); Marie-Pierre Gagnon^{1,3}, PhD; José Côté^{2,4}, RN, PhD; Julie Payne-Gagnon³, MA; Emilie Hudson⁵, RN, BSc; Carl-Ardy Dubois⁴, PhD

¹Faculty of Nursing Sciences, Université Laval, Quebec, QC, Canada

²Research Center of the Centre Hospitalier de l'Université de Montréal, Research Chair in Innovative Nursing Practices, Montreal, QC, Canada

³Research Centre of the Centre Hospitalier Universitaire de Québec-Université Laval, Quebec, QC, Canada

⁴Faculty of Nursing Sciences, Université de Montréal, Montreal, QC, Canada

⁵School of Nursing, McGill University, Montreal, QC, Canada

Corresponding Author:

Marie-Pierre Gagnon, PhD

Faculty of Nursing Sciences

Université Laval

Pavillon Ferdinand-Vandry

1050 Avenue de la Médecine

Quebec, QC, G1V 0A6

Canada

Phone: 1 418 525 4444 ext 53169

Fax: 1 418 525 4194

Email: marie-pierre.gagnon@fsi.ulaval.ca

Abstract

Background: Information and communication technologies (ICTs) are becoming an impetus for quality health care delivery by nurses. The use of ICTs by nurses can impact their practice, modifying the ways in which they plan, provide, document, and review clinical care.

Objective: An overview of systematic reviews was conducted to develop a broad picture of the dimensions and indicators of nursing care that have the potential to be influenced by the use of ICTs.

Methods: Quantitative, mixed-method, and qualitative reviews that aimed to evaluate the influence of four eHealth domains (eg, management, computerized decision support systems [CDSSs], communication, and information systems) on nursing care were included. We used the nursing care performance framework (NCPF) as an extraction grid and analytical tool. This model illustrates how the interplay between nursing resources and the nursing services can produce changes in patient conditions. The primary outcomes included nurses' practice environment, nursing processes, professional satisfaction, and nursing-sensitive outcomes. The secondary outcomes included satisfaction or dissatisfaction with ICTs according to nurses' and patients' perspectives. Reviews published in English, French, or Spanish from January 1, 1995 to January 15, 2015, were considered.

Results: A total of 5515 titles or abstracts were assessed for eligibility and full-text papers of 72 articles were retrieved for detailed evaluation. It was found that 22 reviews published between 2002 and 2015 met the eligibility criteria. Many nursing care themes (ie, indicators) were influenced by the use of ICTs, including time management; time spent on patient care; documentation time; information quality and access; quality of documentation; knowledge updating and utilization; nurse autonomy; intra and interprofessional collaboration; nurses' competencies and skills; nurse-patient relationship; assessment, care planning, and evaluation; teaching of patients and families; communication and care coordination; perspectives of the quality of care provided; nurses and patients satisfaction or dissatisfaction with ICTs; patient comfort and quality of life related to care; empowerment; and functional status.

Conclusions: The findings led to the identification of 19 indicators related to nursing care that are impacted by the use of ICTs. To the best of our knowledge, this was the first attempt to apply NCPF in the ICTs' context. This broad representation could be kept in mind when it will be the time to plan and to implement emerging ICTs in health care settings.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews: CRD42014014762; http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014014762 (Archived by WebCite at <http://www.webcitation.org/6plhMLBZh>)

(*J Med Internet Res* 2017;19(4):e122) doi:[10.2196/jmir.6686](https://doi.org/10.2196/jmir.6686)

KEYWORDS

information and communication technology; eHealth; telehealth; nursing care; review, overview of systematic review

Introduction

Background

The use of information and communication technologies (ICTs) for health, referred to as eHealth [1,2] represent a means to support health care delivery [3]. These technologies change how nurses plan, deliver, document, and review clinical care; this will only continue as technology advances. The processes whereby nurses receive and review diagnostic information, make clinical decisions, communicate and socialize with patients and their relatives, and implement clinical interventions will be fundamentally modified with further integration of ICTs into nursing practice [4,5].

There is a wide range of ICTs used for supporting and providing health care. Mair et al [6] suggested four general domains of eHealth that include a variety of ICTs: management systems, communication systems, computerized decision support systems (CDSSs), and information systems. Management systems allow for the acquisition, storage, transmission, and display of administrative or clinical activities related to patients, such as electronic health records (EHRs) or electronic medical records (EMRs). Communication systems can be used for diagnostic, management, counseling, educational, or support purposes. They can be implemented to facilitate communication between health professionals or between health professionals and patients. There are a wide range of communication systems, varying from email and mobile phones to telemedicine and telecare systems. CDSSs are automated systems accessible from various devices, such as computer, mobile phone, or personal digital assistants (PDAs). They support decision-making for health professionals and assist them in practicing within clinical guidelines and care pathways. Information systems, such as Web-based resources and eHealth portals, refer to the use of Internet technology to access health-related information sources.

To support complex and diversified practices and interventions in nursing, myriad ICTs can be adopted, though not without challenges. Some ICTs, such as EHRs and computerized nursing care plans, facilitate access to patient information and help to document and plan nursing care [7]. However, with the use of these technologies, nurses are expected to change the way they document patient care by shifting from paper-based records to electronic systems. The features (eg, copy and paste, electronic interface, drop-down menus) of electronic nursing documentation may affect critical thinking and accuracy of documentation [8]. Telehealth technologies are another example, which include a wide range of ICTs such as remote patient monitoring, videoconferencing, and computer-mediated communications [9]. In the case of remote patient monitoring (telemonitoring), nurses must be able to process a large quantity

of data from the system (eg, vital signs, symptoms) and then use clinical decision skills to respond properly to each patient's condition [10]. In order to discern cues within the interactions via technological modalities, specific communication skills remain essential, that is, active listening, facilitating conversation, questioning, redirecting, and verifying [11-13].

ICTs are becoming an impetus for quality health care delivery by nurses. It is thus relevant to study the role of nurses in the clinical use of ICTs [3] as well as the impact of ICTs on nursing practices [14]. The use of any type of ICT to provide direct or indirect care to patients may transform nurses' day-to-day practice [3]. In some systematic reviews, different types of ICTs have been reviewed, for instance, EHRs [15], nursing computerized records systems [16], or CDSSs [17]. In general, nursing practice or nursing care was not well-defined in those reviews, and there was no conceptual framework enabling reflection on the way ICTs could influence indicators of nursing care. To overcome this gap, we used a broad and comprehensive conceptualization of nursing care based on the nursing care performance framework (NCPF) [18] to embrace a multidimensional perspective of nursing care. The NCPF is composed of three distinct but interrelated subsystems: nursing resources, nursing services, and patients' conditions. It is defined as "the capacity demonstrated by an organization or an organizational unit to acquire the needed nursing resources and use them in a sustainable manner to produce nursing services that effectively improve patients' conditions ([18], p.6)."

However, an integrated body of knowledge was lacking with respect to the effects of ICTs on nursing care, because of the heterogeneity of ICTs used in the literature as well as the poor conceptualization of nursing care. We conducted an overview of systematic reviews to develop a broad picture of the indicators of nursing care that have the potential to be enhanced or constrained by the use of ICTs. The use of an overview is an interesting starting point from which to compare and contrast outcomes of separate reviews [19] regarding the positive, negative, and neutral effects of ICTs on nursing care.

Objectives

We conducted an overview of systematic reviews to systematically summarize the evidence that comes from qualitative, quantitative, and mixed-method systematic reviews regarding the effects of ICTs on nursing care.

Nursing Care Performance Framework

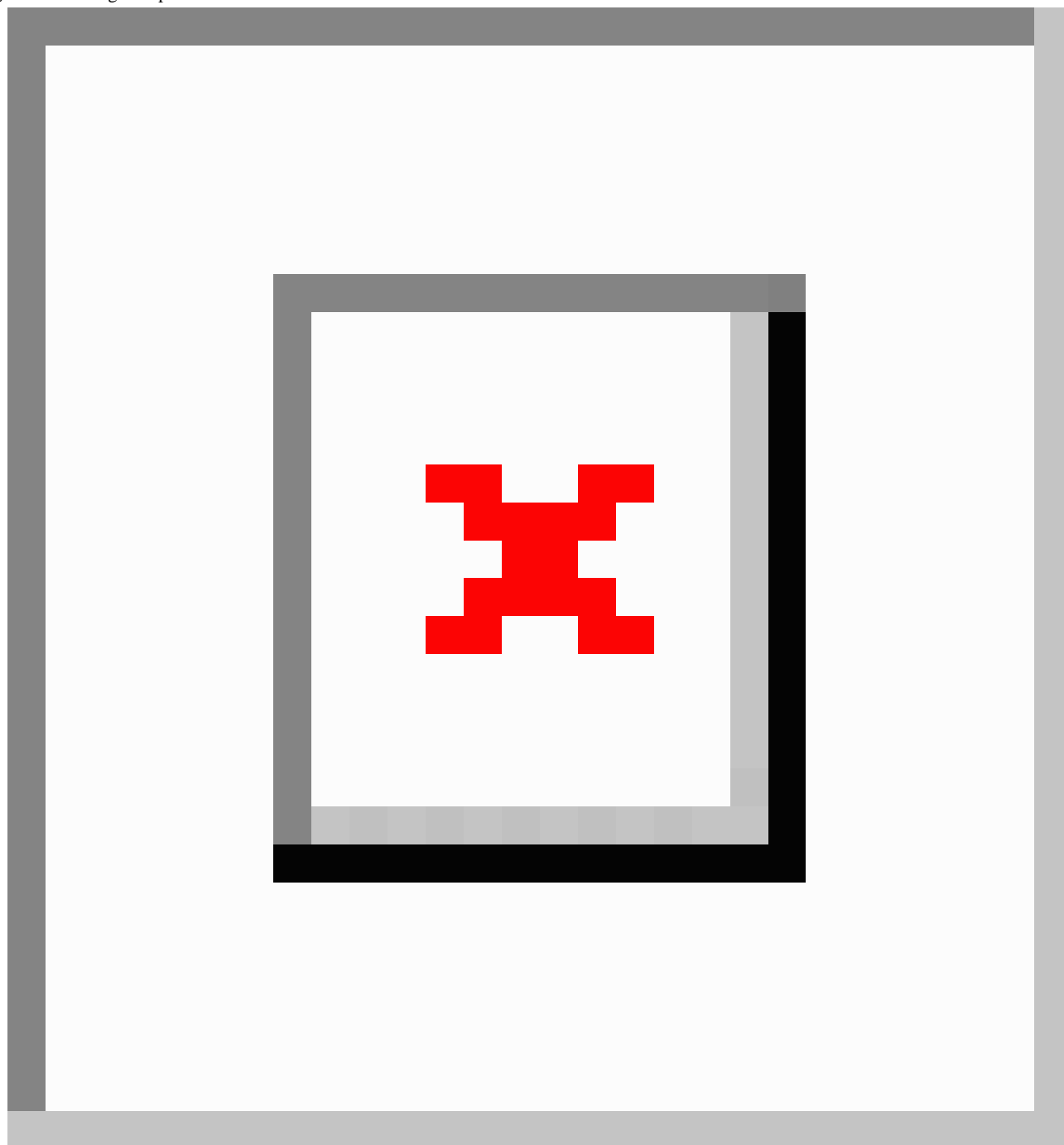
In order to illustrate how ICTs interventions influence nursing care and impact health outcomes, an organizational model was used [18]. The NCPF represents a synthesis of the most recent developments in the field and is part of leading initiatives aiming

to conceptualize nursing care performance. Conceptualization of nursing care performance is based on a system perspective that builds on system theory [20], Donabedian's earlier works on health care organization [21], and Parsons' theory of social action [22].

This model, illustrated in Figure 1, is composed of 14 dimensions and 51 indicators and shows how the interplay of three nursing subsystems (resources, processes or services, and patients' outcomes) can operate to achieve three key functions: (1) acquiring, deploying, and maintaining nursing resources; (2) transforming nursing resources into nursing services; and (3) producing changes in patients' conditions in response to the nursing services provided ("nursing-sensitive outcomes"). The first function refers to the human and material resources needed to provide effective nursing care, such as nursing staff supply, working conditions, staff maintenance, and economic sustainability. The second function encompasses nurses' practice environments (eg, nurse autonomy; collaboration), nursing processes (eg, assessment, care planning, and evaluation; problems and symptom management), nurses' professional satisfaction, and patient experience. The desirable end result of the interactions between nursing staff and nursing processes is to improve patients' conditions. The third function is then described as the positive changes that can be detected among patients (also called "nursing-sensitive outcomes").

The 51 indicators capture the content currently supported by the scientific literature and cover all major areas of nursing care performance. More than a simple list of indicators, the NCPF provides an integrative and systemic framework that has been used in recent studies to analyze various dimensions of nursing care [23,24]. The NCPF has been used, for example, to structure a scoping review undertaken to identify indicators that are sensitive to ambulatory nursing [23]. The results showed the capacity of NCPF to be extended and applied to ambulatory nursing care and furthermore, five new indicators have been added to the framework. The authors of the NCPF have suggested that further studies should be conducted to assess the implementation of the framework in different contexts of nursing care [18]. This overview constitutes a first attempt to use and apply the NCPF to structure and analyze the indicators of nursing care that are influenced by ICTs. We expect that using the NCPF will confirm existing indicators, add new indicators specific to the context of ICTs, and eventually modify existing indicators.

In this overview, our main interest was to extract data related to nurses. For instance, if results of a systematic review were exclusively on patient outcomes without describing nursing resources, services, or processes, the review was excluded. However, we considered nursing sensitive outcomes (ie, patients' outcomes) as long as they could be related to ICTs use by nurses.

Figure 1. Nursing care performance framework.

Methods

Overview and Eligibility Criteria

The protocol of this overview has been registered on PROSPERO (CRD42014014762) and published elsewhere [25]. We followed the Cochrane Collaboration methodology [26] and other relevant works in this domain [19,27] to develop the overview. The scope was formulated using PICOS (participants, interventions, comparisons, outcomes, study design) [28,29]. All types of qualitative, mixed-method, and quantitative reviews, published in French, English, or Spanish from January 1, 1995 and that aimed to evaluate the influence of ICTs (four eHealth domains) used by nurses on nursing care were eligible. The inclusion of reviews using multiple methodological approaches

is justified by the possibility of broadening the understanding of the impact of ICTs on nursing care. The populations of interest were registered nurses (RN), nurses in training, nursing students, or patients receiving care from qualified RN through the medium of ICTs. The interventions targeted were the use of ICTs covered in the four eHealth domains suggested by Mair et al [6]: (1) management systems; (2) communication systems; (3) CDSSs; and (4) information systems. The following ICTs were excluded: (1) nurse management systems, which are purely administrative and designed for the management of human resources and working conditions (eg, scheduling) and nursing staff maintenance (such as retention); (2) educational systems, for example, e-learning initiatives used for the training of nursing students, unless they are applied to direct patient care;

and (3) telephone systems, because according to most definitions of ICTs [30,31], they are not digital technologies and cannot support the electronic capture, storage, processing, and exchange

of information. Further details of the inclusion criteria for the selection of systematic reviews are described in [Table 1](#).

Table 1. Inclusion criteria for the selection of systematic reviews.

Criteria	Description of inclusion criteria
Type of reviews	All types of qualitative, mixed-method, and quantitative reviews that aimed to evaluate the influence of ICTs ^a (four eHealth domains) used by nurses on nursing care, which stated a methodology (a “Methods” section) with explicit eligibility criteria, had systematic research strategies to identify selected reviews and provided a systematic presentation and summary of the characteristics and outcomes of the included reviews [28].
Publication type	Reviews published in French, English, or Spanish from January 1, 1995.
Population	RN ^b , nurses in training, nursing students, or patients receiving care from qualified RN through the medium of ICTs.
Intervention: ICTs covered by four eHealth domains	Four eHealth domains were considered in the overview [6]: management systems, communication systems, computerized decision support systems, and information systems. ICTs embody all digital technologies that support the electronic capture, storage, processing, and exchange of information, in order to promote health, prevent illness, treat disease, manage chronic illness, and so on [30,31].
Management systems	Management systems are computer-based systems used for acquiring, storing, transmitting, and displaying patient administrative or health information from different sources. They can support administrative or clinical activities. Electronic health records (EHRs) and personal health records (PHRs) are examples of management systems.
Communication systems	Telecommunication systems are employed when users are distant in space and/or time. This kind of communication takes place in a synchronous or an asynchronous way, between health professionals, or between health professionals and patients or caregivers. It involves a targeted sharing of information between specific individuals, or individuals who play distinct roles for diagnostic, management, counseling, educational, or support purposes. There are a wide range of communication systems, from email and mobile phones to telemedicine and telecare systems.
Computerized decision support systems (CDSSs)	Refer to an automated computer-based system that aims to support health professionals in practicing within clinical guidelines and care pathways. These systems are usually operated in real-time and involve decision support that comes from artificial intelligence (eg, a software program).
Information systems	Are defined by the use of Internet technology to attain access to different information resources, such as health and lifestyle information. The information remains general, and it is not tailored to specific individual needs. Web-based resources and eHealth portals for retrieving information are some types of information systems.
Comparisons	Usual care, any other ICT, and other types of interventions.
Outcomes	The primary outcomes included nursing resources, nurses’ practice environment, nursing processes or scope of practice, professional satisfaction, and nursing-sensitive outcomes (eg, patient outcomes, such as risk outcomes and safety, patient comfort, and quality of life related to care). The secondary outcomes included nurses’ and patients’ satisfaction or dissatisfaction with ICTs.

^aICTs: information and communication technologies.

^bRN: registered nurse.

Search Strategy

A medical librarian developed and conducted the search strategies, drawing on other reviews of similar topics and using well-established search filters where appropriate. We searched publications in English, French, or Spanish in the following electronic databases from January 1, 1995: Cochrane Database of Systematic Reviews (until January 15, 2015); Epistemonikos (until December 25, 2014); PubMed (until December 8, 2014); Embase (until January 7, 2015); Web of Science (until January 9, 2015); and Cumulative Index to Nursing and Allied Health Literature (CINAHL) (until December 25, 2014).

Structured search strategies were developed using the thesaurus terms of each database (eg, Medical subject heading (MeSH) for PubMed) and using free text, targeting the “title” and “abstract” fields. The strategies were then adapted to the other databases. The results of each database search were collected in a single reference database, and duplicate citations were removed. The specific search strategies for databases are presented in [Multimedia Appendix 1](#).

Selection of Reviews

Two reviewers (GR, JPG) independently screened the title and abstract of the papers in order to assess their eligibility. References that did not meet the preestablished inclusion criteria were excluded. Full-text copies of publications were retrieved

and were assessed by the same two reviewers. Any discrepancies were resolved through discussion. A third reviewer was available for arbitration when consensus was not reached.

Data Extraction and Management

Three reviewers (GR, JPG, and EH) were involved in the data extraction and management process. Information on each review was independently extracted by two of the reviewers. Any disagreement arising during the data extraction process was discussed among the two reviewers. The third reviewer was involved in case of disagreement.

Characteristics of included reviews were extracted and summarized: objectives, type of review, number of included studies, search dates, population, setting, eHealth domain, types of general and specific ICTs, examples of included interventions, comparisons, primary and secondary outcomes, review limitations, and authors' conclusions. A data extraction form was developed based on the NCPF [18] and the dimensions of the actual scope of nursing practice [32]. The data extraction grid was modified during the extraction process by adding dimensions or categories of results. To facilitate teamwork between the three reviewers (GR, JPG, and EH) in performing the data extraction, we used a shared file in Google Sheets. Reviewers communicated with each other through Google Sheets and added comments on the extraction when needed. The three reviewers reviewed the completed data extraction grid to eliminate discrepancies and errors.

Methodological Quality Assessment of Included Reviews

The three reviewers (GR, JPG, and EH) were involved in the methodological quality assessment of the reviews that met the eligibility criteria, using the assessment of multiple systematic reviews (AMSTAR) tool [33,34]. Two reviewers assessed each review independently, and disagreements were discussed. The third reviewer was available for arbitration when needed. AMSTAR is an 11-item checklist from which reviewers assign one point when the criterion is met. AMSTAR items provide an assessment of methodological criteria such as the comprehensiveness of the search strategy and whether the quality of included studies was evaluated and accounted for [35]. AMSTAR characterizes quality at three levels: 8-11 is high quality (ie, minor or no methodological limitations), 4-7 is medium quality (ie, moderate methodological limitations), and 0-3 is low quality (ie, major methodological limitations) [36].

In this overview, we included different types of systematic reviews, that is, quantitative reviews (randomized and nonrandomized designs), mixed-method synthesis reviews, and qualitative reviews. AMSTAR is used primarily for quantitative reviews using randomized controlled trial (RCT) design. When undertaking an overview, challenges encountered are the assessment of limitations (risk of bias) as well as the quality of evidence in systematic reviews [37,38]. There were no reporting guidelines on assessing methodological quality of mixed-method and qualitative reviews at the time of the overview. We decided to apply AMSTAR to all reviews in order to use the same criteria for quality assessment, although this had limitations (ie,

inappropriateness of applying some criteria to mixed-method and qualitative reviews).

Data Synthesis

A statistical meta-analysis of outcomes was not possible because the included studies were too heterogeneous. We therefore conducted a narrative synthesis, which is defined as an approach of summarizing and explaining outcomes from multiple studies by employing the use of words and text [39]. The core characteristic of a narrative synthesis is the adoption of a "textual approach to the process of synthesis to 'tell the story' of the outcomes from the included studies" [39]. We categorized the reviews into subgroups according to the type of intervention and their effects (positive, negative, or no effect) on a specific dimension of nursing care (eg, practice environment, nursing processes, professional satisfaction, and nursing-sensitive outcomes).

Results

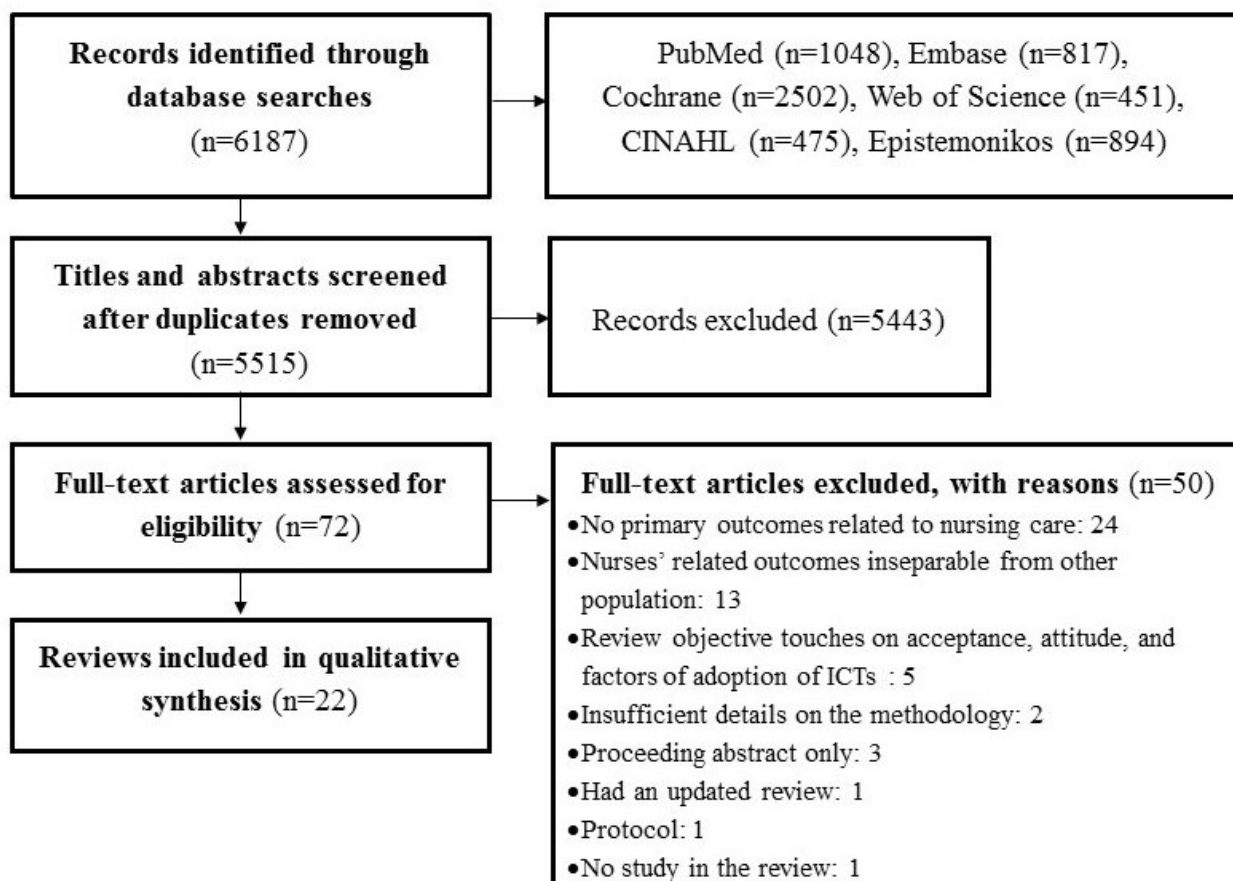
Description of the Reviews

A total of 6187 titles or abstracts were identified. After removing duplicate references, 5515 titles or abstracts were assessed for eligibility. Full-text papers of 72 articles were retrieved for detailed evaluation. It was found that 22 reviews published between 2002 and 2015 met the eligibility criteria. The list of these included reviews is presented in [Multimedia Appendix 2](#). Twelve reviews used a mixed-method synthesis approach, nine used a quantitative approach, and one used a qualitative approach (meta-ethnography). Fifty reviews were mainly excluded because they did not present primary outcomes related to nursing care (n=24), or because outcomes related to nurses were indistinguishable from other populations (n=13). In [Multimedia Appendix 3](#), details are provided regarding the primary reasons for exclusion and the full references of the excluded articles. The preferred reporting items for systematic reviews and meta-analyses (PRISMA) study flow diagram [40] are illustrated in [Figure 2](#) to show the overall process of review selection.

The general characteristics (ie, type of reviews, search dates, target population, and health care settings) of the included reviews are presented in [Multimedia Appendix 4](#). The review objectives, limitations, and main conclusions are synthesized in [Multimedia Appendix 5](#). The eHealth domains covered were management systems (n=14), communication systems (n=7), and CDSSs (n=10). No reviews dealt with information systems. Five reviews included more than one eHealth domain [3,41-44]. Articles reviewing management systems included the following ICTs: electronic medical or health or patient records, computer-based nursing records or computerized nursing care planning, and regional health care information system. The ICTs covered in the communication systems were email, mobile phone, bedside communication tool or bedside terminals, iPod technology to assist in educational conferences, and telemedicine or telehealth with the use of videophone or videoconferencing. The CDSSs covered were medication management technology—e-prescribing, electronic medication administration record systems, computerized provider order entry (CPOE), bar-code medication administration (BCMA)—and PDAs.

These eHealth services can be categorized as belonging to more than one domain [6], depending on their components. Details about eHealth domains, examples of included interventions, and comparisons are presented in [Multimedia Appendix 6](#).

Figure 2. The preferred reporting items for systematic reviews and meta-analyses (PRISMA) study flow diagram. ICT: information and communication technology.



Assessment of Review Quality

The AMSTAR tool was used to assess the methodological quality of all reviews. Four reviews, mostly quantitative ones, were high quality (scores: 8-9); nine were medium quality (scores between 4 and 7), and nine scored low quality (between 0 and 3). AMSTAR scoring for each review is presented in [Table 2](#). We adapted the interpretation of two criteria (#7 and #9) of the AMSTAR tool to assess the quality of mixed-method and qualitative reviews. For the criteria 7—reporting and assessment of scientific quality of the included reviews—we answered “yes” if authors mentioned having assessed and documented quality of quantitative reviews, and if they acknowledged clearly the difficulty of assessing qualitative or mixed-methods reviews. For criteria 9, entailing the inappropriateness of methods used to combine findings, we answered “yes,” based on the decision rules developed by Kitsiou et al [45]: “reviews’ authors made a statement regarding

the inappropriateness of pooling data (eg, highlighted issues about heterogeneity or variability between the studies), that is, the authors summarized and synthesized the available evidence narratively according to a defined analysis plan and/or using appropriate qualitative methods and techniques (eg, construction of common rubrics, content analysis, tabulation, groupings, and clustering).” Regarding criteria 10, about the assessment of publication bias, it seems that empirical evidence on this topic in qualitative research is very limited [46]. We presume that this is the same reality regarding the mixed-method reviews.

Dimensions of Nursing Care That Are Influenced by Information and Communication Technologies

The results (see [Figure 3](#)) will be presented in association with the NCPF: the function, the dimension, and the theme (which correspond or not to a particular indicator in the framework). [Table 3](#) presents the frequency of extracted data per dimensions, themes, and ICTs.

Table 2. Assessment of multiple systematic reviews (AMSTAR) scoring.

References	Type of reviews or designs	AMSTAR score
Free [42]	Quantitative (RCT ^a)	9 (high)
Mador [47]	Quantitative (various designs)	9 (high)
Urquhart [16]	Cochrane review—quantitative (RCT+1 other design)	8 (high)
McKibbin [43]	Mixed	8 (high)
Nieuwlaat [48]	Quantitative (RCT)	7 (medium)
Mickan [49]	Quantitative (RCT)	6 (medium)
Finkelstein [41]	Mixed	6 (medium)
Randell [50]	Quantitative (RCT)	5 (medium)
Georgiou [51]	Quantitative (various designs)	5 (medium)
Dowding [52]	Quantitative (various designs)	5 (medium)
Poissant [53]	Quantitative (various designs)	4 (medium)
Husebo [54]	Mixed (integrative)	4 (medium)
Jones [55]	Mixed (integrative)	4 (medium)
Meißner [56]	Qualitative (meta-ethnography)	3 (low)
Bowles [44]	Mixed	3 (low)
Anderson [17]	Mixed	3 (low)
Maeenpa [57]	Mixed	2 (low)
NGuyen [58]	Mixed	2 (low)
Stevenson [15]	Mixed	2 (low)
Bartoli [59]	Mixed	1 (low)
Carrington [3]	Mixed	1 (low)
Kelley [60]	Mixed (integrative)	0 (low)

^aRCT: randomized controlled trial.

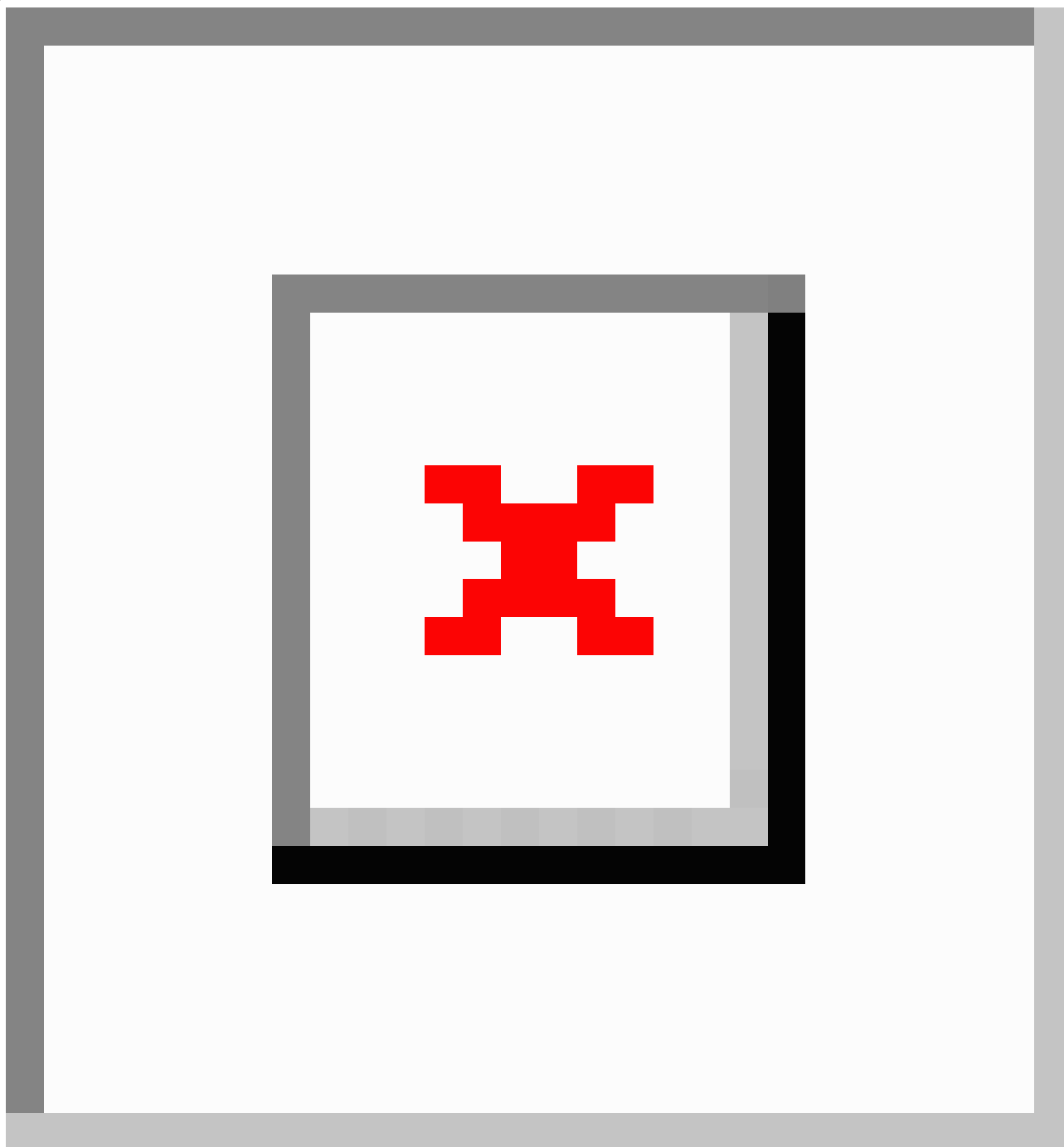
Figure 3. Presentation of results.

Table 3. Frequency extracted data.

Dimension	Themes (Number of reviews) (Types of eHealth domain)	Positive effects of ICTs ^a	Negative effects of ICTs	No effect	Total
Time and efficiency		20	17	7	44
	Time management (4) (MS ^b , CS ^c , CDSS ^d)	2	1	1	4
	Time spent for patient care (7) (MS, CS, CDSS)	4	5	3	12
	Documentation time (7) (MS ^e)	14	11	3	28
Nurses' practice environment		19	5	1	25
	Knowledge updating and utilization (3) (CS, CDSS)	3	0	1	4
	Information quality and access (5) (MS ^f , CDSS)	11	2	0	13
	Nurse autonomy (1) (CS ^e)	1	0	0	1
	Intra and interprofessional collaboration (6) (MS ^f , CS, CDSS)	4	3	0	7
Nursing processes		30	12	3	45
	Nurses competencies-skills (4) (MS, CDSS)	9	1	1	11
	Nurse-patient relationship (3) (CS ^e)	4	0	0	4
	Quality of documentation (7) (MS ^f , CS)	6	4	1	11
	Assessment, care planning, and evaluation (10) (MS, CS, CDSS)	13	8	2	23
	Teaching of patients and families (4) (CS ^f , CDSS)	5	0	0	5
	Communication and care coordination (2) (CS, MS)	2	0	0	2
Professional satisfaction		29	18	1	48
	Nurses' perspectives of the quality of care provided (6) (MS, CS, CDSS)	15	2	0	17
	Satisfaction or dissatisfaction of nurses using ICTs (10) MS, CS, CDSS)	14	16	1	31
Nursing sensitive outcomes		28	5	5	38
	Patient comfort and quality of life related to care (7) (CS, CDSS)	7	0	1	8
	Empowerment (4) (CS ^f , MS)	6	0	1	7
	Functional status (3) (CS ^e)	3	0	1	4
	Satisfaction or dissatisfaction of patients using ICTs (5) (CS, MS)	12	5	2	19

^aICTs: information and communication technologies.^bMS: management systems.^cCS: communications systems.^dCDSSs: computerized decision support systems.^eOne eHealth domain covered exclusively a particular theme.^fMajority of one eHealth domain covered a particular theme.

Function 1: Acquiring, Deploying, and Maintaining Resources

Time and Efficiency

Overall, 11 reviews [15,16,43,44,47,48,51,53,56,58,60] had results related to time: time management (time consumed or time saved resulting the use of ICTs); time spent for patient care; and documentation time.

Time Management

Four reviews [43,44,48,53] targeting CDSSs, communication, and management systems had findings on “time management” in a general way: one review showed no effect [48], another

showed negative effects [43], and two reported positive effects [44,53]. In Nieuwlaat et al's [48] review, results demonstrated that nurses perceived that conventional care compared with CDSSs were equally time-consuming (no effect). The other review reported that reminder systems were “time-consuming” [43]. The results in the Poissant et al [53] review revealed that the use of EHRs has been shown to reduce the time devoted to the verbal transmission of information at the end-of-shift. Consequently, this caused a change in the workflow, which may have been a strong incentive for nurses to become efficient users of the system. In the Bowles and Baugh [44] review, the effect of telehomecare was also reported positively in terms on “saving time.”

Time Spent for Patient Care

Almost one-third of the reviews (7/23) [16,43,47,51,56,58,60] outlined positive [16,47,56,58] and negative effects [43,51,56,60], as well as no effect [43,51,58] of CDSSs, management systems, and communication systems on time spent for patient care. Nurses are sometimes concerned that using electronic nursing documentation or the BCMA for documenting and for administering medication might take away or reduce time for patient care [43,56]. Conversely, other reviews including communication systems (eg, telehomecare) and management systems (eg, EHRs) found that time spent for patient care has significantly improved [16,47,56,58] and particularly, nurses using EHRs spent more time with patients in assessment, education, and communication [58].

Documentation Time

Nurse documentation time was reported in seven reviews [15,16,43,47,53,56,60] touching on management systems, such as EHRs, e-prescribing system, and critical care information system (CCIS). Effects of these ICTs on documentation time were mixed within and across the reviews: six reviews demonstrated positive effects [15,16,47,53,56,60], six demonstrated negative effects [15,16,43,47,53,60], and three demonstrated no effect [43,47,60]. Negative results showed that nurses spent more time documenting when they used management systems and the positive results showed the contrary: documentation time diminished with ICTs. The time saved for documenting was sometimes reallocated for patient care and had positive outcome on the improvement of health care [15]. Otherwise, when the task of documenting took much more time, nurses had less time to spend with patients [60].

Function 2: Transforming Resources into Services

Nurses' Practice Environment

Knowledge Updating and Utilization

Three reviews found positive effects [17,44,54] of CDSSs and communication systems on knowledge updating and utilization, whereas one review found no effect [17]. CDSSs are useful tools to increase knowledge and information use, and translate outcomes from research into practice by improving nurses' compliance with established guidelines [17]. The potential of communication systems (eg, telehomecare or telehealth) to transfer nursing knowledge was also reported [44,54].

Information Quality and Access

The eHealth domain that was the most covered in relation to information quality and access was management systems, covered in four reviews [43,57,58,60], followed by CDSSs in two reviews [17,43]. One review documented the improvement of information quality as perceived by doctors and nurses after the implementation of EHRs [58], and the results of five reviews highlighted information access [17,43,57,58,60]. Management systems and CDSSs had positive impact in three reviews [43,57,60] on information access regarding patient issues, clinical data, medication information or profile, and other information (policies, guidelines, drug resources, patient files). Nurse practitioners felt that CDSSs could assist them with patient care when data is easily accessible with the use of the

technology [43]. The negative impact were pointed out in two reviews [58,60] that cited the results of the same primary study [61], that is, nurses could not retrieve the information perceived as essential for patient care within the electronic nursing documentation system.

Nurse Autonomy

Only one review mentioned nurse autonomy as a positive effect. In this review [59], nurses were expected to handle most cases autonomously and to refer to doctors only in exceptional cases when using the tele-triage system designed to monitor chronic heart failure patients remotely.

Intra- and Interprofessional Collaboration

Four reviews highlighted positive [17,42,52,59] effects regarding intra- and interprofessional collaboration, one showed negative effect [43] and one reported no effect [60] with the use of CDSSs [17], communications [42,59], and management systems [43,52,60]. Reviews including CDSSs reported improved communication between members of the interdisciplinary team [17], such as between nurses and surgeons [42], better and more trustworthy relationships between nurses and doctors by using telehomecare systems [59], and more frequent collaboration between members of the health care team when using management systems (ie, clinical dashboards) [52]. In one review, results showed that electronic nursing documentation systems negatively affected collaborative working relationships between nurses and physicians [60].

Nursing Processes

Nurses' Competencies and Skills

Four reviews that encompassed CDSSs and management systems showed that they had a positive influence on these domains of nurses' competencies and skills: decision support or decision-making [17,43,56], observation skills [56], clinical judgment [17,56], and critical thinking [60]. Additionally, due to some features of CDSSs and management systems (eg, readability of data, remote accessibility of data, better quality of patients' records, presence of reminders, or automatic alerts), these ICTs supported clinical judgment and decision-making [43,56]. Conversely, some features of the ICTs not previously available on paper, such as copy and paste, drop-down menus, and check boxes, affected the nurses' capacity to employ critical thinking regarding their patients [60]. Finally, the results presented in Anderson and Willson [17] review showed no effect of CDSSs on the knowledge or clinical decision-making of nurses associated with pressure ulcer prevention.

Quality of Documentation

Positive effects on documentation quality were highlighted in six reviews [3,15,43,56,58,60]; five on these reviews encompassed management systems. Negative effects were reported in three reviews [15,56,60], and another review documented no effect [41]. Results from the Stevenson et al [15] review: nurses reported that EHRs did not reflect their practice and reported that it was "incapable of capturing much of what they believed was crucial in nursing care." With regards to psychological care, nurses also reported issues with fitting complex caring practice into systems that are not intended to

accommodate it, for example, when providing emotional and psychological support. Since EPRs lack sensitivity, they cannot capture “the being there stuff,” for example, caring for a dying patient by sitting on their bedside and holding their hand. Two reviews [15,56] stated that when the quality of documentation is improved, quality of care and patient safety can be fostered since it allows a complete overview of the patient’s situation [56].

Nurse-Patient Relationship

In three reviews, use of communication systems (virtual visits using videophones, telehomecare, telehealth) positively impacted nurse-patient relationship. Reviews mentioned the potential of ICTs to provide a pathway for communication [55], create new types of bonds with patients [44], establish trust through the videoconference system, and create a sense of connection (from the patients’ perspective) [54].

Assessment, Care Planning, and Evaluation

Impact of CDSSs, management, and communication systems were mixed, that is, positive effects were mentioned in seven reviews [15,17,41,44,49,55,56], negative effects were mentioned in five reviews [15,16,41,42,54], and no effect was documented in two reviews [16,55]. For example, a handheld computer-based support system for preference-based care planning led to a higher consistency between patient preferences and nursing care plan priorities [17]. An “email intervention” cited in the Finkelstein et al [41] review led to a more comprehensive heart failure and medication adherence assessment by nurses being recorded. EHRs contain templates that guide nurses for assessment and help them identify problems [56]. The mixed review by Stevenson et al [15] revealed negative impact of EHRs regarding poor care plans updates, the difficulty of individualizing care plans within the systems, and the difficulty of capturing a broad picture of the patient within the electronic personal record. Similarly, the Urquhart et al [16] review showed that computerized nursing care planning compared with manual planning led to (1) no effect between groups regarding planning; and (2) negative effects, because planned tasks were not carried out as expected for nurses using ICTs.

Teaching of Patients and Families

Four reviews reported teaching benefits: three with the use of communication systems [41,54,55] and one with CDSSs [43]. For example, virtual visits simplified teaching and information sharing with patients and thus became a way to transfer knowledge [54]. Also, patients had clearer instructions on discharge and on their medication administration at home as reported by nurse practitioners [43].

Communication and Care Coordination

Two reviews found that communication systems had positive impact on delivering continuous and coordinated care, on the prevention of preventing relapses into poor health [54], and on improving communication about resident care [56].

Professional Satisfaction

Nurses’ Perspectives of the Quality of Care Provided

In six reviews, positive effects [41,43,44,54-56] of CDSSs, management, and communication systems were reported:

improvement of quality of care and patient safety; nurses’ perceptions that BCMA reduce medication errors and improve medication administration processes [43]; and the provision of comprehensive and adaptive care related to the patients’ needs with the help of telehealth used with elders [55]. In four reviews [15,43,56,58], negative results were discussed: EHRs do not improve patient care as perceived by nurses [58]; and patients do not receive necessary care because the quality of residents’ records is lacking [56].

Satisfaction or Dissatisfaction of Nurses Using ICTs

The results in ten reviews, targeting the three eHealth domains, found that nurse satisfaction was mixed: nine reviews reported positive effects [17,41,43,44,48,54,56,58,60], eight reported negative effects [15,43,44,48,54,56,58,60], and one reported no effect [43]. Results pertained to overall acceptance of ICTs and their satisfaction was described in general ways, such as “nurses were satisfied with ICTs.” There were also elements associated with ICTs, such as system navigability (eg, complexity, ease of use, user-friendliness, and flexibility), nurses’ attitudes, concerns about patients’ privacy, and perceived benefits or inconveniences. Some nurses found EHRs to be irrelevant for practice [58].

Function 3: Producing Changes in Patients’ Condition

Nursing-Sensitive Outcomes

Patient Comfort and Quality of Life Related to Care

The positive effects of CDSSs and communication systems on comfort and quality of life related to care [3,17,41,44,50,54,55] were described in terms of patient outcomes: fewer number of wetting occurrences [17], reduction of malnourished patients [3,50], the reduction of pain and anxiety [44], better quality of life [41], and lower burden related to care [55]. One review reported little improvement on quality of care with the use of telehomecare [44].

Empowerment

Four reviews [16,41,44,54] highlighted empowerment as a positive effect of communication systems. One management system showed no effect [16]. Some examples of positive impact include diabetic patients, who felt that the telehomecare empowered them [44] and had positive results in terms of diabetes management with an eHealth application [41]. One review also cited videoconferences for conducting nursing virtual visits as tools to increase patients’ abilities to manage self-care [54].

Functional Status

In three reviews [41,54,55], the results regarding the effects of communication systems on functional status (eg, physical, cognitive, psychosocial functional capacity) were discussed in a positive way. Computer use (in a telehealth context) and elders’ self-esteem have been positively associated [55]. In another review [54], the results showed that communication systems (eg, virtual visits using videoconference) decreased loneliness and melancholia, enhanced psychosocial and social activity, and aided memory among home-dwelling elders. In the Finkelstein et al [41] review, the results revealed that the

health status of patients among groups did not differ with the use of communication systems.

Satisfaction or Dissatisfaction of Patients of Using ICTs

Patients' satisfaction with ICTs was documented in five reviews that demonstrated positive effects [41,44,54,55,58], three that showed negative effects [54,55,58], and two that showed no effect [55,58]. Patient results indicated their degree of satisfaction or dissatisfaction with ICTs; their acceptance, acceptability, and receptiveness of their usage of ICTs; and their appreciation for being able to schedule videoconferences about topics of their choice [44,54,58]. The results were presented in terms of usefulness (or uselessness); perceived and actual benefits or advantages, such as accessibility and flexibility [54]; ease of use, usability, complexity; and the degree to which the ICTs were well-designed and functioned fully [41,55,58]. Some patients were confident in using ICTs [44], whereas others were concerned about the confidentiality of their health information [58]. Results from the Husebo and Storm [54] review indicate that patients who had visual contact with nurses through communication systems felt cared for and perceived a sense of connection.

Summary Description of eHealth Domains Related to Specific Themes

On the basis of the content of Table 3, we propose a summary description of which eHealth domains cover specific themes of nursing care.

Management Systems

The only eHealth domain reported to influence the documentation time was management systems, such as electronic nursing documentation [60], CCIS [47], CPOE, eMAR [43], and EHRs [53]. The other themes reported with these systems were time spent on patient care; time management; information quality and access, intra and interprofessional collaboration; quality of documentation; nurses' competencies and skills; assessment, care planning, and evaluation; nurses' perspectives of the quality of care provided; empowerment; and satisfaction or dissatisfaction of nurses and patients using ICTs.

Communication Systems

Communication systems was the only eHealth domain found to be applicable to the themes of nurse-patient relationship, autonomy for nurses in their role, and patients' functional status. These themes were also discussed related to communication systems: teaching patients and families, knowledge update and utilization; intra and interprofessional collaboration; quality of documentation; assessment, care planning, and evaluation; communication and care coordination; nurses' perspectives of the quality of care provided; satisfaction or dissatisfaction of nurses and patients using ICTs; patient comfort and quality of life related to care; and patients' empowerment.

Computerized Decision Support Systems (CDSSs)

CDSSs are mentioned in nurses' practice environment dimension (3/4): knowledge updating and utilization; information quality and access; and intra and interprofessional collaboration. Regarding the nurses' competencies and skills, CDSSs are involved with decision-making processes. Some other themes

discussed in relation to CDSSs are assessment, care planning, and evaluation; teaching of patients and families; nurses' perspectives of the quality of care provided; satisfaction or dissatisfaction of nurses and patients using ICTs; and patient comfort and quality of life related to care.

Discussion

Summary of Main Results

This overview allowed a broad understanding of the dimensions of nursing care influenced by using ICTs for providing care. Regarding the primary outcomes of interest, the themes that were most frequently reported are documentation time; assessment, care planning, and evaluation; nurses' perspective of the quality of care provided; information quality and access; and time spent for patient care. For secondary outcomes, satisfaction or dissatisfaction of nurses and patients using ICTs was frequently mentioned.

Discussion of Results With Respect to the First Function of the NCPF

In relation to the first function of the NCPF (*acquiring, deploying, and maintaining nursing resources*), many reviews outlined outcomes linked to "time." The use of ICTs affected time management, time spent for patient care, and documentation time. This theme could also refer to a dimension of the NCPF called *maintenance and economic sustainability of the nursing staff* [18]. Sustainability refers to the importance of having quality resources at the lowest cost. This dimension highlights productivity and the necessity to optimize the outputs produced from a given set of inputs; in other words, to minimize the amount of nursing tasks, materials, and equipment without sacrificing the quality of nursing services. The "time" dimension can be understood in terms of how ICTs can impact staff, productivity, optimization of the staff's time management, and resources utilization. We do believe that time is an interesting outcome related to the resources of the overall structure (nursing staff), but it does not reflect directly on how ICTs can transform or support what nurses do (nursing activities or interventions) within their actual scope of practice. Considering our results, we do not believe that further research should focus on "time" in order to better understand the effects of ICTs on nursing care (and specifically, on nursing processes).

This review did not explore other dimensions and indicators related to the first function of the NCPF, such as nursing staff supply. These dimensions include quantity and quality indicators. As an example, it would be interesting to explore whether the availability of ICTs in specific health care settings impacts the quantity of nurses needed to perform nursing services.

Another relevant topic would be to probe whether ICTs act as facilitator or motivator to enhance nurses' working conditions, or serve as a barrier that inhibits them. To what extent can ICTs create favorable conditions that attract nurses and reinforce stability in the workforce? A systematic review was undertaken on the effect of ICTs on retention and recruitment of health care professionals [62]. The results revealed that, in 9 out of the 13 studies, ICT use demonstrated a positive, though often indirect,

influence on recruitment and retention. The influence of ICTs on retention of nurses was also examined in a qualitative study [63]. The results highlighted various impact of ICTs on nurse retention (ie, little or no impact, unclear impact, or indirect positive impact).

Discussion of Results With Respect to the Second Function of the NCPF

The three dimensions corresponding to the second function, *transforming nursing resources into nursing services*, are nurses' practice environment, nursing processes, and professional satisfaction. The themes "knowledge updating and utilization" and "communication and care coordination" were not explicitly described in the NCPF and we used them from the instrument of "actual scope of nursing practice" [32]. The indicator "scope of practice" is included in the nursing processes of the NCPF, but there are no explicit underlying subindicators.

The "information quality and access" theme was analyzed as an effect of ICTs on nurses' practice environment. In other words, ICTs are seen as a potential way to support nursing work by allowing them to get access to various sources of information and clinical data. The theme "quality of documentation" is not part of the nurses' practice environment because it is linked to what nurses do as activities.

The capacity of nurses to deliver nursing interventions is intimately and consistently linked with organizational processes that capture the nursing practice context and mediate its outcomes [64,65]. These processes, defined as interventions, support nursing work and sustain a professional environment [66]. We hypothesize that, if nurses have access to a comprehensive set of information about patients, this would trickle down on nursing processes, such as quality of documentation, assessment, care planning, and evaluation. It would also impact communication and care coordination to benefit patient outcomes.

A surprising result is the following: only one review mentioned nurse autonomy in relation to the use of ICTs [59]. It would be interesting to know more about questions such as: How can we define "autonomy" in a context in which nurses use or are exposed to ICTs to provide nursing care? How can ICTs support or influence nurse autonomy? Can ICTs be a required training tool in nurses' practice environments to support their own autonomy?

The NCPF model reflects the deployment of nurses' full scope of practice, including assessment, planning, and evaluation; problem and symptom management; health promotion and illness prevention; care coordination; and discharge planning, which are conceptualized through interventions and processes in the model.

From a health care provider perspective, these processes grasp the technical elements of care and reflect the extent to which staff are capable of using and mobilizing their competencies to deploy their entire scope of practice. These processes demonstrate the capability of nurses to engage the needs of patients [18]. Our results show that, in reference to the processes described in the NCPF, few such processes have been described in the studies included in this overview. However, assessment,

care planning, and evaluation are the most cited themes in the nursing processes dimension, followed by teaching of patients and families and, finally, by communication and care coordination. Despite these outcomes, it would be helpful to conduct primary studies on how ICTs could influence or support other nursing processes, such as problem and symptom management, health promotion and illness prevention, and discharge planning.

Nurses' professional satisfaction is conceived as the result of nursing processes. Our results revealed two facets of this satisfaction: nurses' perspective of the quality of care provided and nurses' satisfaction or dissatisfaction using ICTs. The NCPF included additional indicators that were not mentioned in the included reviews to capture the nurses' professional satisfaction: having the time to do their job and the enjoyment derived from it.

Discussion of Results With Respect to the Third Function of the NCPF

We believe that nursing-sensitive outcomes, which are the "patient outcomes," are underrepresented in our overview because of our inclusion criteria that focused on reviews of the impact or effect of ICTs on nursing resources and services. Thus, patient outcomes were only considered if nursing outcomes were reported. This means that we included patients' outcomes as primary outcomes as long as they fell within the usage of ICTs by nurses, and then, when outcomes related to the second function of the NCPF (nursing services and processes) were reported. Dubois and colleagues [67] undertook a systematic work including three literature reviews to identify the priority indicators in evaluating the nursing contribution to quality of care. The results revealed that the most frequently examined nursing sensitive outcomes are pressure ulcers, medication administration errors, urinary infections by catheter, and falls. These indicators are located in the "risk outcomes and safety" dimension of the NCPF. Despite this, there are several systematic reviews on the effects of ICTs on patients' outcomes [68-71]. However, these reviews do not necessarily explore the impact of ICTs on nursing services and processes (second function of the NCPF) when considering patients' outcomes.

Strengths and Potential Biases

There are many strengths of this overview. First, it employed a comprehensive search strategy, which was developed and implemented by a medical librarian. Second, data extraction and quality assessment were conducted by three reviewers working independently. Third, the data extraction process was done with the use of the NCPF, which supported the organization and the analysis of results. This framework supported reflection on the way ICTs could influence specific aspects of nursing care. Some new, redefined, or adapted dimensions and indicators have been suggested in the framework: time management, time spent for patient care and documentation time, information quality and access, quality of documentation, knowledge updating and utilization as part of the nurses' practice environment, communication and care coordination, and nurse and patient satisfaction or dissatisfaction regarding their use of ICTs. Fourth, one of the authors of the NCPF (CAD) challenged the analysis and interpretation of the

results. Some debriefing meetings were held to discuss the way the themes were presented related to their organization in the NCPF (under specific subsystems, functions, dimensions, and indicators).

There are also limitations to this overview. First, as mentioned by other authors [27,72], we were limited by the information provided by the review authors. The granularity of details available was limited and some information was lacking regarding both the description of ICTs (eg, their features, components, contexts of use, and area of practice) as well as findings regarding the dimensions of nursing care influenced by ICTs. Therefore, it was not possible to make significant conclusions about how a specific ICT influenced one or many indicators (themes) of nursing care, and it was challenging to categorize these extracted findings (impact of ICTs) within the NCPF. A comprehensive description of interventions (ICTs) would have been helpful. Further research could be done to gain knowledge about how a specific ICT used in a certain area of practice can impact on one or many dimensions and indicators of nursing care.

Third, the nature of the topic was not easy to capture in the reported data of systematic reviews. It was difficult to establish if nurses experienced changes in their practice with the use of ICTs, or if instead they believed that ICTs would change their practice and work environment without really experiencing these transformations. Some outcomes related to the use of ICTs are reported in terms of “barriers.” However, it is not always clear if it is a barrier to use ICTs or an effect or impact of having used them. Systematic reviews on the determinants of nurses’ acceptance and use of ICTs are plenty [31,73-75], but do not inform on the real effects of ICTs on nursing practice.

Fourth, we used AMSTAR to assess methodological quality of qualitative and mixed-method reviews even if this tool was not developed for types of reviews other than quantitative using mainly RCT designs. The results of this work should be interpreted with caution. Although it provides a broad perspective on the phenomenon of interest, the main shortcoming of a review of systematic reviews is the heterogeneity in terms of population, interventions (types of ICTs), types of reviews, and the variety of outcomes, which might lead to the possibility of biased conclusions. For further research and methodological development in this domain, we strongly recommend a consolidated tool to evaluate the quality of different types of reviews on a common scale. The results of the assessment of methodological quality of mixed-method and qualitative reviews must be interpreted with caution, considering that AMSTAR is not used and designed for that purpose. In fact, some criteria do not fit the specificities of other types of reviews because there are no gold standards or guidelines allowing us to perform this task. Consequently, mixed-method and qualitative reviews started with a lower score, which cannot lead to a judgment about the likely bias and methodological limitations inherent in the majority of reviews summarized in Table 2.

Finally, this overview draws a picture of the reality of ICTs that covered a period extended from 2002 to the start of 2015. The

emerging or novel ICTs that have been published from 2015 until now could not be captured.

Differences Between Protocol and Overview

As stated in the protocol [25], one of the objectives was to explore whether specific categories of ICTs (management systems, communication systems, CDSSs, or information systems) could have an impact on nursing care. As mentioned earlier, the heterogeneity of reviews and the lack of granularity regarding extracted data or information were some reasons why we could not pursue the initial objective.

When we planned this overview, we were particularly interested in the dimensions of nursing care inherent to the second and the third function of the NCPF, which are nurses’ practice environment, nursing processes, professional satisfaction (second function or subsystem), and nursing-sensitive outcomes (third function or subsystem). Throughout the data extraction process, we realized that some outcomes, particularly those related to the time and efficiency, were frequently mentioned. We then decided to extract these results based on their frequency and their impact on the nursing care.

Authors’ Conclusions

To the best of our knowledge, this is the first attempt to draw a broad understanding and a schematization of specific dimensions and indicators of nursing care influenced by ICTs. Using the NCPF was useful to illustrate the way ICTs can impact 3 subsystems (nursing resources, nursing services or processes, and nursing sensitive outcomes or patients’ outcomes), 5 dimensions, and 19 themes corresponding to the NCPF indicators. Findings of this overview are a good starting point from which we could deepen our conceptualization on the way nursing care system performance can be affected by ICTs. According to a systemic perspective, it is plausible to believe that the adoption and implementation of ICTs in the nursing care system must be addressed under a multidimensional perspective, considering that the 3 subsystems are interrelated. If nurses use ICTs to support their interventions, and the impact of such ICTs are positive or negative on the work they do, this could possibly reverberate on patient outcomes. We have to keep this broad representation in mind when it will be the time to plan and to implement emerging ICTs in health care settings.

Takeaway Messages

Using the NCPF was relevant to draw a broad, multidimensional, and a system-based perspective on the dimensions and indicators of nursing care that can be impacted by ICTs.

ICTs have a mixed impact on 19 indicators related to nursing care: documentation time, time spent for patient care, time management, knowledge updating and utilization, information quality and access, nurse autonomy, intra and interprofessional collaboration, nurses competencies-skills, nurse-patient relationship, quality of documentation, assessment, care planning and evaluation, teaching of patients and families, communication and care coordination, nurses’ perspectives of the quality of care provided, patient comfort and quality of life related to care, empowerment, functional status, and satisfaction or dissatisfaction of nurses and patients using ICTs.

Management systems, including, for instance, electronic nursing documentation system, CCIS, CPOE, eMAR, and EHRs, have been discussed exclusively with the theme “documentation time” (in the included reviews).

Communication systems have been described exclusively regarding nurse-patient relationship, autonomy for nurses in their role, and patients’ functional status (eg, physical, cognitive, and psychosocial functional capacity).

Acknowledgments

This research is funded thanks to a Knowledge Transfer (KT) Canada Student Research Stipend, awarded to GR, an initiative funded by the Canadian Institutes of Health Research (CIHR; grant #88368). MPG holds the Tier 2 Canadian Research Chair in Technologies and Practices in Health and JC holds the Research Chair in Innovative Nursing Practices. Special thanks to William Witteman, who planned and performed the search strategy, and Suzana Anjos for the linguistic revision of this manuscript. We acknowledge the precious work of the three reviewers: D Dowding, JM Carrington, and P Johansson.

Authors' Contributions

GR conceived and designed the overview with input from MPG and JC. GR informed the search strategy and performed the search, with the help of a health librarian. GR, JPG, and EH were responsible for data extraction. GR, MPG, JC, JPG, EH, and CAD have been involved in data analysis and interpretation of results. GR, MPG, JC, JPG, EH, and CAD were engaged in the drafting of this manuscript and they all read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategies.

[[PDF File \(Adobe PDF File\), 377KB - jmir_v19i4e122_app1.pdf](#)]

Multimedia Appendix 2

List of included reviews.

[[PDF File \(Adobe PDF File\), 122KB - jmir_v19i4e122_app2.pdf](#)]

Multimedia Appendix 3

Excluded articles and reasons for exclusion.

[[PDF File \(Adobe PDF File\), 198KB - jmir_v19i4e122_app3.pdf](#)]

Multimedia Appendix 4

General characteristics of included reviews.

[[PDF File \(Adobe PDF File\), 163KB - jmir_v19i4e122_app4.pdf](#)]

Multimedia Appendix 5

Review objectives, limitations, and main conclusions.

[[PDF File \(Adobe PDF File\), 268KB - jmir_v19i4e122_app5.pdf](#)]

Multimedia Appendix 6

eHealth domains, interventions, and comparisons.

[[PDF File \(Adobe PDF File\), 318KB - jmir_v19i4e122_app6.pdf](#)]

References

1. Eng T. The e-Health Landscape: A Terrain Map of Emerging Information and Communication Technologies in Health and Health Care. Princeton, NJ: The Robert Wood Johnson Foundation; 2001.

2. World Health Organization. 2016. e-Health URL: <http://www.who.int/topics/ehealth/en/> [accessed 2017-01-21] [WebCite Cache ID 6ngZcO7Y5]
3. Carrington JM, Tias VL. Nursing informatics year in review. *Nurs Adm Q* 2013;37(2):136-143. [doi: [10.1097/NAQ.0b013e3182869deb](https://doi.org/10.1097/NAQ.0b013e3182869deb)] [Medline: [23454993](#)]
4. Institute of Medicine. *The Future of Nursing: Leading Change, Advancing Health*. Washington, DC: The National Academies press; 2011.
5. Monteiro AP. Cyborgs, biotechnologies, and informatics in health care - new paradigms in nursing sciences. *Nurs Philos* 2016 Jan;17(1):19-27. [doi: [10.1111/nup.12088](https://doi.org/10.1111/nup.12088)] [Medline: [25919864](#)]
6. Mair F, May C, Murray E, Finch T, Anderson G, O'Donnell C, et al. NETSCC. 2009. Understanding the implementation and integration of e-Health services URL: http://www.netssc.ac.uk/hsdr/files/project/SDO_FR_08-1602-135_V01.pdf [accessed 2016-09-21] [WebCite Cache ID 6kge8Bkve]
7. Sullivan DH. Technological advances in nursing care delivery. *Nurs Clin North Am* 2015 Dec;50(4):663-677. [doi: [10.1016/j.cnur.2015.07.005](https://doi.org/10.1016/j.cnur.2015.07.005)] [Medline: [26596655](#)]
8. Kossman SP, Scheidenhelm SL. Nurses' perceptions of the impact of electronic health records on work and patient outcomes. *Comput Inform Nurs* 2008;26(2):69-77. [doi: [10.1097/01.NCN.0000304775.40531.67](https://doi.org/10.1097/01.NCN.0000304775.40531.67)] [Medline: [18317257](#)]
9. Nagel DA, Penner JL. Conceptualizing telehealth in nursing practice: advancing a conceptual model to fill a virtual gap. *J Holist Nurs* 2016 Mar;34(1):91-104. [doi: [10.1177/0898010115580236](https://doi.org/10.1177/0898010115580236)] [Medline: [25858897](#)]
10. Barrett D. The role of telemonitoring in caring for older people with long-term conditions. *Nurs Older People* 2012 Sep;24(7):21-25. [doi: [10.7748/nop2012.09.24.7.21.c9257](https://doi.org/10.7748/nop2012.09.24.7.21.c9257)] [Medline: [23008916](#)]
11. Greenberg ME. A comprehensive model of the process of telephone nursing. *J Adv Nurs* 2009 Dec;65(12):2621-2629. [doi: [10.1111/j.1365-2648.2009.05132.x](https://doi.org/10.1111/j.1365-2648.2009.05132.x)] [Medline: [19941546](#)]
12. Purc-Stephenson RJ, Thrasher C. Nurses' experiences with telephone triage and advice: a meta-ethnography. *J Adv Nurs* 2010 Mar;66(3):482-494. [doi: [10.1111/j.1365-2648.2010.05275.x](https://doi.org/10.1111/j.1365-2648.2010.05275.x)] [Medline: [20423383](#)]
13. Romero YM, Angelo M, Muñoz GL. Imaginative construction of care: the nursing professional experience in a remote care service. *Rev Lat Am Enfermagem* 2012;20(4):693-700 [FREE Full text] [Medline: [22990154](#)]
14. de Veer AJ, Fleuren MA, Bekkema N, Francke AL. Successful implementation of new technologies in nursing care: a questionnaire survey of nurse-users. *BMC Med Inform Decis Mak* 2011;11:67 [FREE Full text] [doi: [10.1186/1472-6947-11-67](https://doi.org/10.1186/1472-6947-11-67)] [Medline: [22032728](#)]
15. Stevenson JE, Nilsson GC, Petersson GI, Johansson PE. Nurses' experience of using electronic patient records in everyday practice in acute/inpatient ward settings: a literature review. *Health Informatics J* 2010 Mar;16(1):63-72. [doi: [10.1177/1460458209345901](https://doi.org/10.1177/1460458209345901)] [Medline: [20413414](#)]
16. Urquhart C, Currell R, Grant MJ, Hardiker NR. Nursing record systems: effects on nursing practice and healthcare outcomes. *Cochrane Database Syst Rev* 2009 Jan 21(1):CD002099. [doi: [10.1002/14651858.CD002099.pub2](https://doi.org/10.1002/14651858.CD002099.pub2)] [Medline: [19160206](#)]
17. Anderson JA, Willson P. Clinical decision support systems in nursing: synthesis of the science for evidence-based practice. *Comput Inform Nurs* 2008;26(3):151-158. [doi: [10.1097/01.NCN.0000304783.72811.8e](https://doi.org/10.1097/01.NCN.0000304783.72811.8e)] [Medline: [18438151](#)]
18. Dubois C, D'Amour D, Pomey M, Girard F, Brault I. Conceptualizing performance of nursing care as a prerequisite for better measurement: a systematic and interpretive review. *BMC Nurs* 2013 Mar 07;12:7 [FREE Full text] [doi: [10.1186/1472-6955-12-7](https://doi.org/10.1186/1472-6955-12-7)] [Medline: [23496961](#)]
19. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol* 2011;11(1):15 [FREE Full text] [doi: [10.1186/1471-2288-11-15](https://doi.org/10.1186/1471-2288-11-15)] [Medline: [21291558](#)]
20. Reid P, Grossman JH, Fanjiang G. *Building a better delivery system: a new engineering/health care partnership*. Washington, DC: National Academies Press; 2005.
21. Donabedian A. Models for organizing the delivery of personal health services criteria for evaluating them. *Milbank Mem Fund Q* 1972;50(4):103-154.
22. Parsons T. *Structure and Process in Modern Societies*. New York, NY: Free Press; 1960.
23. Rapin J, D'Amour D, Dubois C. Indicators for evaluating the performance and quality of care of ambulatory care nurses. *Nurs Res Pract* 2015;2015:861239 [FREE Full text] [doi: [10.1155/2015/861239](https://doi.org/10.1155/2015/861239)] [Medline: [26380108](#)]
24. VanFosson CA, Jones TL, Yoder LH. Unfinished nursing care: an important performance measure for nursing care systems. *Nurs Outlook* 2016;64(2):124-136. [doi: [10.1016/j.outlook.2015.12.010](https://doi.org/10.1016/j.outlook.2015.12.010)] [Medline: [26850334](#)]
25. Rouleau G, Gagnon M, Côté J. Impact of information and communication technologies on nursing care: an overview of systematic reviews (protocol). *Syst Rev* 2015 May 23;4:75 [FREE Full text] [doi: [10.1186/s13643-015-0062-y](https://doi.org/10.1186/s13643-015-0062-y)] [Medline: [26002726](#)]
26. Becker L, Oxman A. *Cochrane Handbook for Systematic Reviews of Interventions*. 2011. Chapter 22: Overviews of reviews URL: http://handbook.cochrane.org/chapter_22/22_overviews_of_reviews.htm [accessed 2016-09-20] [WebCite Cache ID 6keSUnVWK]
27. Worswick J, Wayne SC, Bennett R, Fiander M, Mayhew A, Weir MC, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? *Syst Rev* 2013;2:26 [FREE Full text] [doi: [10.1186/2046-4053-2-26](https://doi.org/10.1186/2046-4053-2-26)] [Medline: [23647654](#)]

28. O'Connor D, Green S, Higgins J. Cochrane Handbook for Systematic Reviews of Interventions. 2011. Chapter 5: Defining the review question and developing criteria for including studies URL: http://handbook.cochrane.org/chapter_5/5_defining_the_review_question_and_developing_criteria_for.htm [accessed 2016-09-21] [WebCite Cache ID 6kgi5pz4l]
29. Centre for reviews dissemination. 2009. Systematic reviews: CRD's guidance for undertaking reviews in health care URL: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [accessed 2017-01-27] [WebCite Cache ID 6npTBjkDu]
30. Bashshur RL, Shannon GW, Krupinski EA, Grigsby J, Kvedar JC, Weinstein RS, et al. National telemedicine initiatives: essential to healthcare reform. *Telemed J E Health* 2009;15(6):600-610. [doi: [10.1089/tmj.2009.9960](https://doi.org/10.1089/tmj.2009.9960)] [Medline: [19534591](https://pubmed.ncbi.nlm.nih.gov/19534591/)]
31. 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](https://doi.org/10.1007/s10916-010-9473-4)] [Medline: [20703721](https://pubmed.ncbi.nlm.nih.gov/20703721/)]
32. D'Amour D, Dubois C, Déry J, Clarke S, Tchouaket E, Blais R, et al. Measuring actual scope of nursing practice: a new tool for nurse leaders. *J Nurs Adm* 2012 May;42(5):248-255. [doi: [10.1097/NNA.0b013e31824337f4](https://doi.org/10.1097/NNA.0b013e31824337f4)] [Medline: [22525288](https://pubmed.ncbi.nlm.nih.gov/22525288/)]
33. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10 [FREE Full text] [doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10)] [Medline: [17302989](https://pubmed.ncbi.nlm.nih.gov/17302989/)]
34. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009 Oct;62(10):1013-1020. [doi: [10.1016/j.jclinepi.2008.10.009](https://doi.org/10.1016/j.jclinepi.2008.10.009)] [Medline: [19230606](https://pubmed.ncbi.nlm.nih.gov/19230606/)]
35. Weir MC, Ryan R, Mayhew A, Worswick J, Santesso N, Lowe D, et al. The Rx for change database: a first-in-class tool for optimal prescribing and medicines use. *Implement Sci* 2010;5:89 [FREE Full text] [doi: [10.1186/1748-5908-5-89](https://doi.org/10.1186/1748-5908-5-89)] [Medline: [21087505](https://pubmed.ncbi.nlm.nih.gov/21087505/)]
36. Sharif MO, Janjua-Sharif FN, Sharif FN, Ali H, Ahmed F. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. *Oral Health Dent Manag* 2013 Mar;12(1):9-16. [Medline: [23474576](https://pubmed.ncbi.nlm.nih.gov/23474576/)]
37. Lunny C, Brennan SE, McDonald S, McKenzie JE. Evidence map of studies evaluating methods for conducting, interpreting and reporting overviews of systematic reviews of interventions: rationale and design. *Syst Rev* 2016 Jan 06;5:4 [FREE Full text] [doi: [10.1186/s13643-015-0178-0](https://doi.org/10.1186/s13643-015-0178-0)] [Medline: [26739283](https://pubmed.ncbi.nlm.nih.gov/26739283/)]
38. Pieper D, Buechter R, Jerinic P, Eikermann M. Overviews of reviews often have limited rigor: a systematic review. *J Clin Epidemiol* 2012 Dec;65(12):1267-1273. [doi: [10.1016/j.jclinepi.2012.06.015](https://doi.org/10.1016/j.jclinepi.2012.06.015)] [Medline: [22959594](https://pubmed.ncbi.nlm.nih.gov/22959594/)]
39. Popay J, Roberts H, Sowden A. Lancaster. 2006. Guidance on the conduct of narrative synthesis in systematic reviews: final report URL: <http://www.lancaster.ac.uk/shm/research/nssr/research/dissemination/publications.php> [accessed 2016-09-21] [WebCite Cache ID 6kgid2HMT]
40. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: [10.1371/journal.pmed.1000100](https://doi.org/10.1371/journal.pmed.1000100)] [Medline: [19621070](https://pubmed.ncbi.nlm.nih.gov/19621070/)]
41. Finkelstein J, Knight A, Marinopoulos S, Gibbons MC, Berger Z, Aboumatar H, et al. Enabling patient-centered care through health information technology. *Evid Rep Technol Assess (Full Rep)* 2012 Jun(206):1-1531. [Medline: [24422882](https://pubmed.ncbi.nlm.nih.gov/24422882/)]
42. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. *PLoS Med* 2013 Jan;10(1):e1001363 [FREE Full text] [doi: [10.1371/journal.pmed.1001363](https://doi.org/10.1371/journal.pmed.1001363)] [Medline: [23458994](https://pubmed.ncbi.nlm.nih.gov/23458994/)]
43. McKibbon KA, Lokker C, Handler SM, Dolovich LR, Holbrook AM, O'Reilly D, et al. Enabling medication management through health information technology (Health IT). *Evid Rep Technol Assess (Full Rep)* 2011 Apr(201):1-951. [Medline: [23126642](https://pubmed.ncbi.nlm.nih.gov/23126642/)]
44. Bowles KH, Baugh AC. Applying research evidence to optimize telehomecare. *J Cardiovasc Nurs* 2007;22(1):5-15 [FREE Full text] [Medline: [17224692](https://pubmed.ncbi.nlm.nih.gov/17224692/)]
45. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. *J Med Internet Res* 2015;17(3):e63 [FREE Full text] [doi: [10.2196/jmir.4174](https://doi.org/10.2196/jmir.4174)] [Medline: [25768664](https://pubmed.ncbi.nlm.nih.gov/25768664/)]
46. Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gülmezoglu M, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med* 2015 Oct;12(10):e1001895 [FREE Full text] [doi: [10.1371/journal.pmed.1001895](https://doi.org/10.1371/journal.pmed.1001895)] [Medline: [26506244](https://pubmed.ncbi.nlm.nih.gov/26506244/)]
47. Mador RL, Shaw NT. The impact of a Critical Care Information System (CCIS) on time spent charting and in direct patient care by staff in the ICU: a review of the literature. *Int J Med Inform* 2009 Jul;78(7):435-445. [doi: [10.1016/j.jiminf.2009.01.002](https://doi.org/10.1016/j.jiminf.2009.01.002)] [Medline: [19261544](https://pubmed.ncbi.nlm.nih.gov/19261544/)]
48. Nieuwlaat R, Connolly SJ, Mackay JA, Weise-Kelly L, Navarro T, Wilczynski NL, CCDSS Systematic Review Team. Computerized clinical decision support systems for therapeutic drug monitoring and dosing: a decision-maker-researcher partnership systematic review. *Implement Sci* 2011;6:90 [FREE Full text] [doi: [10.1186/1748-5908-6-90](https://doi.org/10.1186/1748-5908-6-90)] [Medline: [21824384](https://pubmed.ncbi.nlm.nih.gov/21824384/)]
49. Mickan S, Atherton H, Roberts NW, Heneghan C, Tilson JK. Use of handheld computers in clinical practice: a systematic review. *BMC Med Inform Decis Mak* 2014;14:56 [FREE Full text] [doi: [10.1186/1472-6947-14-56](https://doi.org/10.1186/1472-6947-14-56)] [Medline: [24998515](https://pubmed.ncbi.nlm.nih.gov/24998515/)]

50. Randell R, Mitchell N, Dowding D, Cullum N, Thompson C. Effects of computerized decision support systems on nursing performance and patient outcomes: a systematic review. *J Health Serv Res Policy* 2007 Oct;12(4):242-249. [doi: [10.1258/135581907782101543](https://doi.org/10.1258/135581907782101543)] [Medline: [17925077](https://pubmed.ncbi.nlm.nih.gov/17925077/)]
51. Georgiou A, Prgomet M, Paoloni R, Creswick N, Hordern A, Walter S, et al. The effect of computerized provider order entry systems on clinical care and work processes in emergency departments: a systematic review of the quantitative literature. *Ann Emerg Med* 2013 Jun;61(6):644-653.e16. [doi: [10.1016/j.annemergmed.2013.01.028](https://doi.org/10.1016/j.annemergmed.2013.01.028)] [Medline: [23548404](https://pubmed.ncbi.nlm.nih.gov/23548404/)]
52. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: review of the literature. *Int J Med Inform* 2015 Feb;84(2):87-100. [doi: [10.1016/j.ijmedinf.2014.10.001](https://doi.org/10.1016/j.ijmedinf.2014.10.001)] [Medline: [25453274](https://pubmed.ncbi.nlm.nih.gov/25453274/)]
53. Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. *J Am Med Inform Assoc* 2005 Oct;12(5):505-516 [FREE Full text] [doi: [10.1197/jamia.M1700](https://doi.org/10.1197/jamia.M1700)] [Medline: [15905487](https://pubmed.ncbi.nlm.nih.gov/15905487/)]
54. Husebø Anne Marie Lunde AML, Storm M. Virtual visits in home health care for older adults. *The Scientific World Journal* 2014;2014:689873 [FREE Full text] [doi: [10.1155/2014/689873](https://doi.org/10.1155/2014/689873)] [Medline: [25506616](https://pubmed.ncbi.nlm.nih.gov/25506616/)]
55. Jones JF, Brennan PF. Telehealth interventions to improve clinical nursing of elders. *Annu Rev Nurs Res* 2002;20:293-322. [Medline: [12092513](https://pubmed.ncbi.nlm.nih.gov/12092513/)]
56. Meißner A, Schnepf W. Staff experiences within the implementation of computer-based nursing records in residential aged care facilities: a systematic review and synthesis of qualitative research. *BMC Med Inform Decis Mak* 2014;14:54 [FREE Full text] [doi: [10.1186/1472-6947-14-54](https://doi.org/10.1186/1472-6947-14-54)] [Medline: [24947420](https://pubmed.ncbi.nlm.nih.gov/24947420/)]
57. Mäenpää T, Suominen T, Asikainen P, Maass M, Rostila I. The outcomes of regional healthcare information systems in health care: a review of the research literature. *Int J Med Inform* 2009 Nov;78(11):757-771. [doi: [10.1016/j.ijmedinf.2009.07.001](https://doi.org/10.1016/j.ijmedinf.2009.07.001)] [Medline: [19656719](https://pubmed.ncbi.nlm.nih.gov/19656719/)]
58. Nguyen L, Bellucci E, Nguyen LT. Electronic health records implementation: an evaluation of information system impact and contingency factors. *Int J Med Inform* 2014 Nov;83(11):779-796. [doi: [10.1016/j.ijmedinf.2014.06.011](https://doi.org/10.1016/j.ijmedinf.2014.06.011)] [Medline: [25085286](https://pubmed.ncbi.nlm.nih.gov/25085286/)]
59. Bartoli L, Zanaboni P, Masella C, Ursini N. Systematic review of telemedicine services for patients affected by chronic obstructive pulmonary disease (COPD). *Telemed J E Health* 2009 Nov;15(9):877-883. [doi: [10.1089/tmj.2009.0044](https://doi.org/10.1089/tmj.2009.0044)] [Medline: [19919194](https://pubmed.ncbi.nlm.nih.gov/19919194/)]
60. Kelley TF, Brandon DH, Docherty SL. Electronic nursing documentation as a strategy to improve quality of patient care. *J Nurs Scholarsh* 2011 Jun;43(2):154-162. [doi: [10.1111/j.1547-5069.2011.01397.x](https://doi.org/10.1111/j.1547-5069.2011.01397.x)] [Medline: [21605319](https://pubmed.ncbi.nlm.nih.gov/21605319/)]
61. Darbyshire P. 'Rage against the machine?': nurses' and midwives' experiences of using computerized patient information systems for clinical information. *J Clin Nurs* 2006 Feb;13(1):17-25. [doi: [10.1111/j.1365-2702.2006.01180.x](https://doi.org/10.1111/j.1365-2702.2006.01180.x)] [Medline: [16422744](https://pubmed.ncbi.nlm.nih.gov/16422744/)]
62. Gagnon M, Pollender H, Trépanier A, Dupl  a E, Ly BA. Supporting health professionals through information and communication technologies: a systematic review of the effects of information and communication technologies on recruitment and retention. *Telemed J E Health* 2011 May;17(4):269-274. [doi: [10.1089/tmj.2010.0175](https://doi.org/10.1089/tmj.2010.0175)] [Medline: [21476881](https://pubmed.ncbi.nlm.nih.gov/21476881/)]
63. Gagnon M, Breton E, Par   G, Courcy F, C     J, Tr  panier A, et al. [The impact of information and communication technologies on nurse retention]. *Sante Publique* 2013;25(3):305-313. [Medline: [24007906](https://pubmed.ncbi.nlm.nih.gov/24007906/)]
64. Aiken LH, Clarke SP, Sloane DM. Hospital restructuring: does it adversely affect care and outcomes? *J Health Hum Serv Adm* 2001;23(4):416-442. [Medline: [11924307](https://pubmed.ncbi.nlm.nih.gov/11924307/)]
65. Sidani S, Doran DM, Mitchell PH. A theory-driven approach to evaluating quality of nursing care. *J Nurs Scholarsh* 2004;36(1):60-65. [Medline: [15098420](https://pubmed.ncbi.nlm.nih.gov/15098420/)]
66. Lake ET. Development of the practice environment scale of the nursing work index. *Res Nurs Health* 2002 Jun;25(3):176-188. [doi: [10.1002/nur.10032](https://doi.org/10.1002/nur.10032)] [Medline: [12015780](https://pubmed.ncbi.nlm.nih.gov/12015780/)]
67. Dubois C, D'Amour D, Brault I. Indicateurs prioritaires pour   valuer la contribution infirmi  re    la qualit   des soins : revue syst  matique des   crits. Montr  al: Secr  tariat international des infirmi  res et infirmiers de l'espace francophone (SIDIIEF); 2015.
68. Agboola SO, Ju W, Elfiky A, Kvedar JC, Jethwani K. The effect of technology-based interventions on pain, depression, and quality of life in patients with cancer: a systematic review of randomized controlled trials. *J Med Internet Res* 2015;17(3):e65 [FREE Full text] [doi: [10.2196/jmir.4009](https://doi.org/10.2196/jmir.4009)] [Medline: [25793945](https://pubmed.ncbi.nlm.nih.gov/25793945/)]
69. Lima IC, Galv  o MTG, Alexandre HO, Lima FET, Ara  jo TL. Information and communication technologies for adherence to antiretroviral treatment in adults with HIV/AIDS. *Int J Med Inform* 2016 Aug;92:54-61. [doi: [10.1016/j.ijmedinf.2016.04.013](https://doi.org/10.1016/j.ijmedinf.2016.04.013)] [Medline: [27318071](https://pubmed.ncbi.nlm.nih.gov/27318071/)]
70. Singh TP, Vangaveti VN, Kennedy RL, Malabu UH. Role of telehealth in diabetic foot ulcer management - a systematic review. *Aust J Rural Health* 2016 Aug;24(4):224-229. [doi: [10.1111/ajr.12284](https://doi.org/10.1111/ajr.12284)] [Medline: [27098271](https://pubmed.ncbi.nlm.nih.gov/27098271/)]
71. Raaijmakers LC, Pouwels S, Berghuis KA, Nienhuijs SW. Technology-based interventions in the treatment of overweight and obesity: a systematic review. *Appetite* 2015 Dec;95:138-151. [doi: [10.1016/j.appet.2015.07.008](https://doi.org/10.1016/j.appet.2015.07.008)] [Medline: [26165415](https://pubmed.ncbi.nlm.nih.gov/26165415/)]
72. Cheung A, Weir M, Mayhew A, Kozloff N, Brown K, Grimshaw J. Overview of systematic reviews of the effectiveness of reminders in improving healthcare professional behavior. *Syst Rev* 2012;1:36 [FREE Full text] [doi: [10.1186/2046-4053-1-36](https://doi.org/10.1186/2046-4053-1-36)] [Medline: [22898173](https://pubmed.ncbi.nlm.nih.gov/22898173/)]

73. Gephart S, Carrington JM, Finley B. A systematic review of nurses' experiences with unintended consequences when using the electronic health record. *Nurs Adm Q* 2015;39(4):345-356. [doi: [10.1097/NAQ.0000000000000119](https://doi.org/10.1097/NAQ.0000000000000119)] [Medline: [26340247](#)]
74. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, et al. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. *BMC Med* 2011;9:46 [FREE Full text] [doi: [10.1186/1741-7015-9-46](https://doi.org/10.1186/1741-7015-9-46)] [Medline: [21524315](#)]
75. Young LB, Chan PS, Cram P. Staff acceptance of tele-ICU coverage: a systematic review. *Chest* 2011 Feb;139(2):279-288 [FREE Full text] [doi: [10.1378/chest.10-1795](https://doi.org/10.1378/chest.10-1795)] [Medline: [21051386](#)]

Abbreviations

ADE: adverse drug event
ADL: activities of daily living
AMSTAR: assessment of multiple systematic reviews
APN: advanced practice nurse
BCMA: bar-coded medication administration
CCIS: critical care information system
CDSSs: computerized decision support systems
CINAHL: Cumulative Index to Nursing and Allied Health Literature
COPD: chronic obstructive pulmonary disease
CPIS: computerized patient information systems
CPOE: computerized provider order entry
CS: communication systems
D-RHIS: disease-specific regional health care information systems
ECG: electrocardiogram
ED: emergency department
EHR: electronic health record
eMAR: electronic medication administration record systems
EMR: electronic medical record
EPR: electronic personal record
ES-NIS: expert system nursing information system
EWS: early warning score
HCPs: health care providers
HIT: health information technology
ICTs: information and communication technologies
ICU: intensive care unit
I-RHIS: integrated regional health care information systems
IT: information technology
LOS: length of stay
MeSH: Medical subject heading
MMIT: medication management health information technology
MS: management systems
NCPE: nursing care performance framework
NHS: National Health Service
ORA: organization risk analyzer
PCC: patient-centered care
PDA: personal digital assistant
PHR: personal health record
PICOs: participants, interventions, comparisons, outcomes and studies
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCT: randomized controlled trial
RHIO: regional health care information organization
RHIS: regional health care information system
RN: registered nurse
SMS: short message service
TDMD: therapeutic drug monitoring and dosing
WAP: wireless application protocol

Edited by G Eysenbach; submitted 23.09.16; peer-reviewed by D Dowding, JM Carrington, P Johansson; comments to author 29.12.16; revised version received 27.01.17; accepted 05.03.17; published 25.04.17.

Please cite as:

Rouleau G, Gagnon MP, Côté J, Payne-Gagnon J, Hudson E, Dubois CA

Impact of Information and Communication Technologies on Nursing Care: Results of an Overview of Systematic Reviews

J Med Internet Res 2017;19(4):e122

URL: <http://www.jmir.org/2017/4/e122/>

doi: [10.2196/jmir.6686](https://doi.org/10.2196/jmir.6686)

PMID: [28442454](https://pubmed.ncbi.nlm.nih.gov/28442454/)

©Geneviève Rouleau, Marie-Pierre Gagnon, José Côté, Julie Payne-Gagnon, Emilie Hudson, Carl-Ardy Dubois. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 25.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Web-Based Medical Appointment Systems: A Systematic Review

Peng Zhao¹, MSc; Illhoi Yoo^{1,2}, PhD; Jaie Lavoie³, PharmD, MS; Beau James Lavoie⁴, PharmD, MS; Eduardo Simoes^{1,2}, MSc, DLSHTM, MPH, MD

¹Informatics Institute, University of Missouri, Columbia, MO, United States

²Department of Health Management and Informatics, School of Medicine, University of Missouri, Columbia, MO, United States

³Vizient, Center for Advanced Analytics & Informatics, Chicago, IL, United States

⁴Trinity Health, Livonia, MI, United States

Corresponding Author:

Illhoi Yoo, PhD

Department of Health Management and Informatics

School of Medicine

University of Missouri

Clinical Support and Education Building (DC006.00)

Five Hospital Dr

Columbia, MO, 65212

United States

Phone: 1 5738827642

Fax: 1 5738826158

Email: yooil@health.missouri.edu

Abstract

Background: Health care is changing with a new emphasis on patient-centeredness. Fundamental to this transformation is the increasing recognition of patients' role in health care delivery and design. Medical appointment scheduling, as the starting point of most non-urgent health care services, is undergoing major developments to support active involvement of patients. By using the Internet as a medium, patients are given more freedom in decision making about their preferences for the appointments and have improved access.

Objective: The purpose of this study was to identify the benefits and barriers to implement Web-based medical scheduling discussed in the literature as well as the unmet needs under the current health care environment.

Methods: In February 2017, MEDLINE was searched through PubMed to identify articles relating to the impacts of Web-based appointment scheduling.

Results: A total of 36 articles discussing 21 Web-based appointment systems were selected for this review. Most of the practices have positive changes in some metrics after adopting Web-based scheduling, such as reduced no-show rate, decreased staff labor, decreased waiting time, and improved satisfaction, and so on. Cost, flexibility, safety, and integrity are major reasons discouraging providers from switching to Web-based scheduling. Patients' reluctance to adopt Web-based appointment scheduling is mainly influenced by their past experiences using computers and the Internet as well as their communication preferences.

Conclusions: Overall, the literature suggests a growing trend for the adoption of Web-based appointment systems. The findings of this review suggest that there are benefits to a variety of patient outcomes from Web-based scheduling interventions with the need for further studies.

(*J Med Internet Res* 2017;19(4):e134) doi:[10.2196/jmir.6747](https://doi.org/10.2196/jmir.6747)

KEYWORDS

appointments and schedules; Internet; smartphone; patient-centered care; no-show patients; hospital information systems

Introduction

Background of Web-Based Appointment System

Traditionally, medical appointments have been made with schedulers over the telephone or in person. These methods are based on verbal communications with real people and allow for maximum flexibility in complicated situations [1]. However, because these traditional methods require the intervention of schedulers, the ability to get a timely appointment is not only limited by the availability of appointment slots, but also by the schedulers and phone lines [2,3]. Patients' satisfaction with appointment booking is influenced by their ability to book at the right time with the right health service providers [4].

The Internet has recently emerged as another means to make appointments. Web-based appointment scheduling has been a popular research topic. Several studies conducted satisfaction surveys and found that Web-based appointment scheduling is an extremely important feature, and most patients would use the service again [2,5-7].

There are two major types of Web-based medical appointment services, medical scheduling software as a service (SaaS) and proprietary Web-based scheduling systems. Medical scheduling SaaS has gained increasing prominence in recent years. These appointment systems are not built up by health care practices themselves, but are provided and maintained by health IT companies such as ZocDoc and InQuicker on a paid subscription basis [8]. The appointment services are cloud-based and can be integrated into health care providers' own management systems. The other type of appointment service is proprietary appointment systems, which are integrated into patient portals on providers' websites [9]. A patient portal is a secured Web-based service that allows patients to access their health information and communicate with their health care providers at any time [10]. In the United States, the growth of patient portals has largely been spurred by meaningful use (MU) requirements [11] because of the federal incentive program for adoption of electronic health records. To meet the requirements of MU and receive its incentives, the portal should be actively used by both the practice and patients [12].

There are two modes of Web-based appointment systems, asynchronous and real-time. In the asynchronous mode, appointments are requested through emails or electronic forms on providers' website, and then manually processed by schedulers. In the real-time mode, patients can directly interact with providers' scheduling management systems [3,13]. Although the asynchronous Web-based appointment systems also use the Internet as a medium, they basically replicate the process of telephone-based appointment scheduling [13]. Under

the asynchronous mode, if an appointment is requested outside of a provider's business hours, it will not be processed until schedulers return to work. Normally, Web-based appointment requests are put in the same queue as phone-call appointments, and are thus limited by the backlog of phone calls in the queue [14].

Aims of the Study

Despite the increasing adoption of Web-based appointment systems, their potential benefits are yet to be systematically studied. The purpose of this review was to examine the current body of literature about Web-based medical appointment systems, specifically in regard to their potential benefits to patients and providers. We also want to identify the most effective services or components of them and explore the benefits and barriers of implementation. It is not the intention of this work to review the literature regarding fundamental theories of medical scheduling or system design, which have been studied and reviewed by Cayirli et al [15] and Gupta et al [16]. To the best of our knowledge, this study is the first systematic literature review of the impacts of implementing Web-based medical scheduling systems.

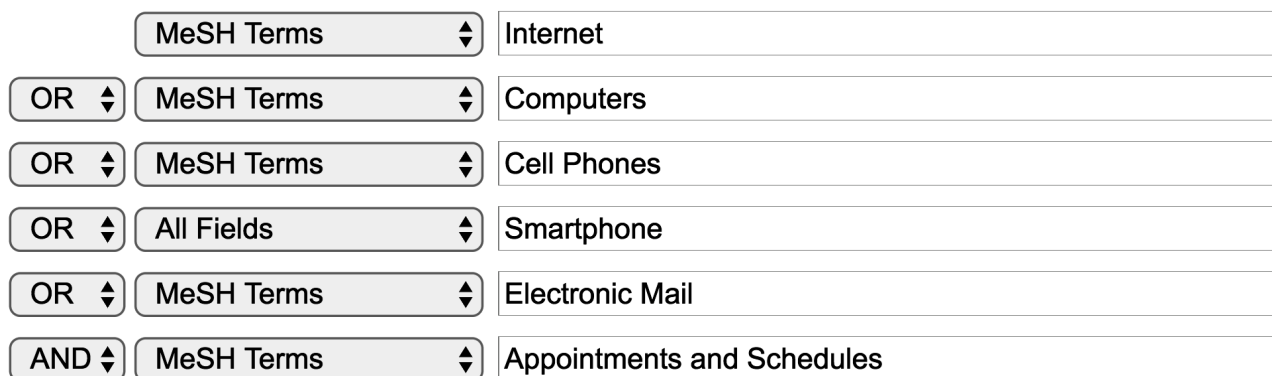
Methods

Data Source

In this study, we present a systematic literature review of Web-based medical appointment systems following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement for systematic reviews [17].

A literature search was performed in MEDLINE using PubMed to identify pertinent articles relating to the impacts of Web-based appointment scheduling. The MeSH terms used in the search included "Internet," "computers," "cell phones," "electronic mail," and "appointments and schedules." "Smartphone" used to be an entry term for "cell phones," and it became a MeSH Descriptor in 2016. To include articles indexed by "smartphone" after 2016 and articles involving smartphones before 2016, "smartphone" was included in the search without any restrictions. Figure 1 shows the logical relationships among the search keywords and their restrictions in the search builder of PubMed.

The literature search was initially performed in April 2016. Since then, in order to make this literature review up-to-date (by including new articles), we regularly conducted literature searches with the same search keywords. Our last literature search was carried out in late February 2017.

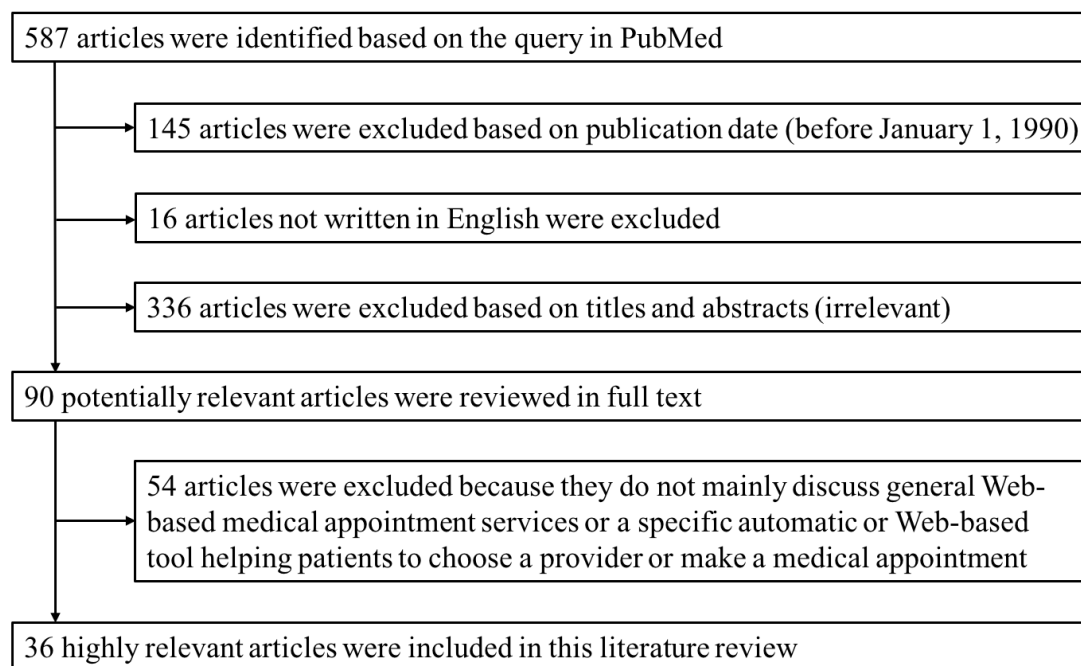
Figure 1. Logical relationships between the search keywords in the search builder of PubMed.

Inclusion and Exclusion Criteria

In this study, articles published only after January 1, 1990, were included, because articles published earlier than this time were unlikely to be relevant to Web-based appointments. We only included articles mainly discussing general Web-based medical appointment services or a specific automated or Web-based tool that assisted patients in choosing a provider or making a medical appointment. The exclusion criteria were systems that solely discussed email- or phone-based appointment reminders and systems not designed for use by patients. Articles not written in English were excluded too.

Study Selection

The process of identifying eligible articles is shown in [Figure 2](#). The initial query returned 587 articles, which were then filtered by publication date and language. 145 articles were excluded because they were published before January 1, 1990. Also, 16 non-English articles were filtered out. The remaining 426 articles were reviewed based on titles and abstracts and 336 of them were excluded due to low relevancy. The remaining 90 articles were then reviewed in full text, and 54 of them were excluded as they do not mainly discuss Web-based medical appointment services or a specific automatic or Web-based tool helping patients to choose a provider or make a medical appointment. The remaining 36 highly relevant articles discussing 21 Web-based medical scheduling systems were used in this literature review.

Figure 2. Trial flow diagram of identifying eligible articles.

Results

Literature on this topic is very recent, with 16/36 articles published after 2010 and 35/36 published after 2000. The studies are highly heterogeneous in research design. More than one third (14/36) of the articles [3,13,18-29] discuss Web-based scheduling as standalone systems or components of portals, and report measurable or perceived (unquantified) improvements in some metrics after the implementation. Eight studies [1,2,5,7,30-33] conducted structured or semistructured interviews to sample patients' attitudes toward specific Web-based appointment systems, and one study [34] surveyed both patients and providers regarding the transformation to patient-centered access to care. Six articles [9,35-39] discuss the necessity and the potential of computerized or Web-based appointment services. Three studies [8,14,40] retrospectively analyzed Web-based appointment data and compared them with traditional appointments. Two studies [6,41] surveyed people's interest in using the Internet to schedule appointments (not tied to any specific Web-based appointment systems). One study [42] used a randomized controlled trial to assess the impact of a Web-based health management system. Another study [43] reported a Web-based provider recommendation system and validated it with a field experiment. These articles also vary in interventions and the granularity of information provided. Many studies were implemented in only a single clinic and had interventions that spanned from basic websites to detailed patient portals. Details provided about the specific components of each system and functionality vary from study to study and many offer only a vague description. Many studies also used multiple interventions simultaneously, such as a Web-based scheduling system with automated reminders and patient decision tools and patient portals. As a result, these studies cannot be directly compared.

[Multimedia Appendix 1](#) summarizes the characteristics of the 21 Web-based appointment systems discussed in the literature. Of these 21 Web-based scheduling systems, 1 is based in Australia, 1 in Canada, 1 in mainland China, 1 in Taiwan, 2 in the United Kingdom, and the remaining 15 in the United States.

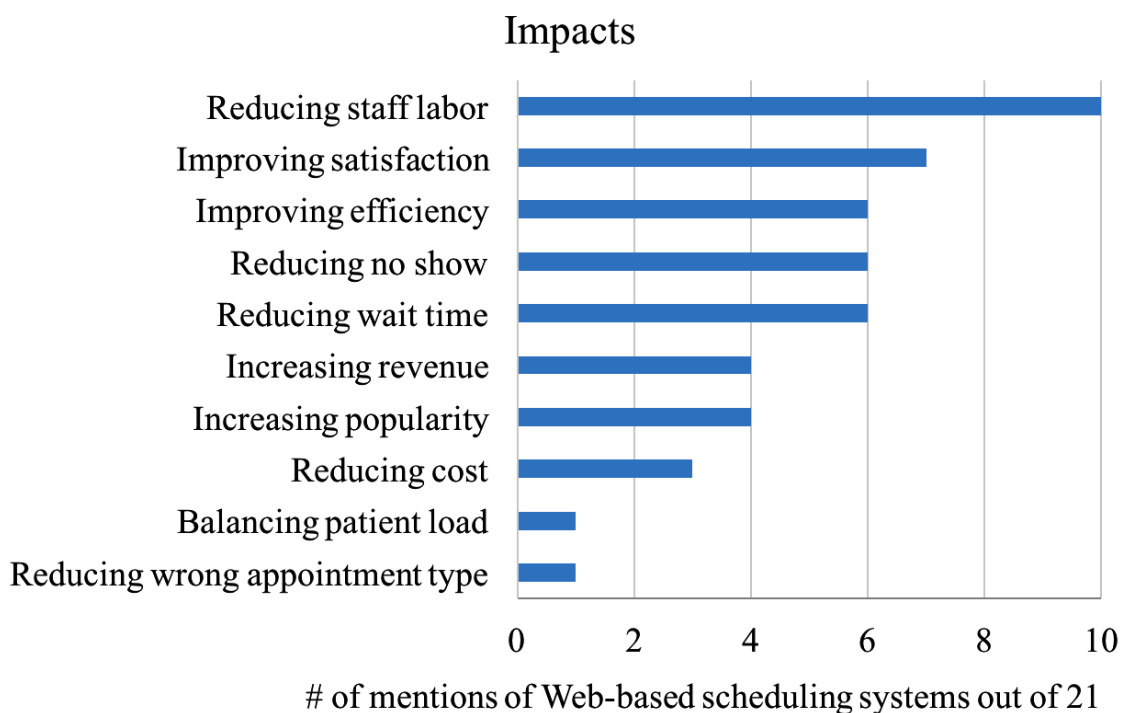
Many articles specifically measured reductions in no-show rate and waiting time as metrics to evaluate Web-based scheduling services.

Siddiqui et al [8] reported a no-show rate of 6.9% for dermatology appointments made with ZocDoc, significantly

lower than the no-show rates of appointments made by traditional appointment making means (17-31%). The UK national online electronic referral and booking service "Choose and Book" was reported to have a significantly better rate of attendance than traditional appointment methods (95% CI 4.3, 20.5%, $P < .01$) [40]. Walters et al [25] reported the Web-based communication tool "Patient Online" reduced no-shows by 42%. The Dartmouth-Hitchcock Medical Center in New Hampshire has reduced no-shows by 40% after implementing an asynchronous clinical messaging service that allows patients to request, review, reschedule, and cancel appointments [26]. The US Department of Defense's health care program Tricare achieved a no-show rate of 2% from Web-based scheduling compared with 8% from phone-based scheduling [20]. The Murry Hill Medical Group based in New York had a similar pattern in the no-show rate: less than 1% of Web-based appointments were missed compared with about 8% of phone-based appointments [21].

Cao et al [31] reported the Web-based appointment system (WAS) reduced the total average waiting time to 7 min from 98 min in a Chinese hospital because patients don't need to queue up for the appointments when they use WAS. In the United Kingdom, the Department of Health requires the maximum waiting time for sexual health service appointments to be 48 h. The introduction of eTriage increased the percentage of patients offered an appointment within 48 h from 48% to 100% [2].

Besides reductions in no-show rate and waiting time, many other improvements were also reported from the literature and they are summarized in [Figure 3](#). The horizontal axis indicates the number of mentions of Web-based scheduling systems for each impact after implementing the 21 Web-based scheduling systems. To limit the number of categories (on the vertical axis), some of the close metrics were merged into a broader category. For example, "optimizing the referral process" and "streamlining operations" were merged into "improving efficiency," as they both indicate improvements in the internal operations of the practices. [Figure 3](#) shows that the most cited (10/21) positive change is "reducing staff labor," closely followed by "improving satisfaction" (7/21), "improving efficiency" (6/21), "reducing no-show" (6/21), "reducing wait time" (6/21), "increasing revenue" (4/21), "increasing popularity" (4/21), "reducing cost" (3/21), "balancing patient load" (1/21), and "reducing wrong appointment type" (1/21).

Figure 3. Impacts after implementing the 21 Web-based scheduling systems.

Discussion

Principal Findings

The Web-based medical appointment reframes the way to communicate with providers' appointment management systems. Compared with traditional appointment methods, Web-based appointment scheduling has unique advantages and disadvantages. In this section, the key benefits and barriers to the adoption of Web-based appointment scheduling will be discussed.

Patient-Centeredness

Patient-centeredness is one of the six quality aims proposed by the Institute of Medicine to improve health care quality in the United States [34]. Web-based medical scheduling as a medical self-service offers a more patient-centered means to make appointments [6]. Most Web-based appointment systems are interfaced with a calendar-like list. Patients can browse and select the most convenient appointment time from the available time slots [21]. In contrast, patients are only given very limited options of available time slots in traditional appointment systems. Besides time slots, some of the Web-based systems allow patients to filter physicians by physicians' attributes such as education background, experience, gender, and reviews from other patients [8].

Another convenience from improved patient access is that patients can fill out registration forms [26], get prescreened and review practice policies online [23] before they show up and this can smooth workflow and reduce misunderstandings.

In the self-servicing Web-based appointments, patients' own descriptions of the reason for visit are often more detailed and illuminating [13]. Sometimes, patients might be uncomfortable

or unable to vocalize certain symptoms (eg, sexual health problems) to the scheduler over the phone or in person, and they may make an untrue statement [2,13]. They tend to be more candid when they schedule online by themselves [13,20].

Reduced No-Show Rates

No-show is a significant cause of wasted clinical resources [40]. The patient-centered design in Web-based appointments has the potential to decrease no-show rates [8,25]. The reasons for the reduction of no-shows after implementing Web-based scheduling have not been systematically studied in the literature, but it could be attributed to the improved access in Web-based scheduling that allows patients to easily verify, cancel, and reschedule their appointments [25]. A possible reason is that patients feel more responsible for their appointments when they make appointments by themselves [44].

Reduced Waiting Time

Waiting is an indicator of service quality and a source of dissatisfaction that affects health care outcomes and patient retention [45,46]. Long waiting time may make patients seek care from other providers and thus this can potentially cause a loss in revenue.

The most cited benefit of real-time scheduling is after-hour access [1,3,21]. Real-time scheduling requires minimal intervention of schedulers and thus can help reduce the waiting time caused by human factors. The available time slots are transparent to patients through the Web interface. Patients are free to claim available appointment slots anytime and anywhere [3,20,37].

The support of same-day or soon appointments by some real-time systems can help further shorten the time between when the appointment is requested and when the medical service

is fulfilled [3]. Although there is a concern that the ability to book in advance for chronic conditions might be diminished by same-day appointments due to the limited number of appointment slots [47], same-day appointments could produce positive outcomes as long as the provider can find a balance in his or her capacity. For providers, it is possible to reuse the time slots released due to late cancellations. These allotted time slots will be otherwise wasted if traditional appointment methods are used because of the longer turnaround time [8].

Barriers to Adoption

It is well known that medicine has lagged in the adoption of new technologies. Although Web-based appointment scheduling comes with many benefits, some providers and patients are reluctant to use it. By 2007, only about 3.2% of the population in 7 European countries (Denmark, Germany, Greece, Latvia, Norway, Poland, and Portugal) had used the Internet to make medical appointments [41]. Only about 15% of public hospitals and 18% of private hospitals in Italy allowed appointments to be made online in 2008-2009 [39]. According to a study conducted by Google and Compete (a research vendor) in 2012, only 21% of patients booked appointments via computer or mobile devices [48]. Only about 7% of primary care practices in Canada and 30% in the United States offered Web-based appointment services in 2012 [49]. As of 2014, 67% of general practitioner (GP) practices in Scotland have websites and only 10% of them support Web-based appointments [35].

There are many reasons for the slow adoption. First, the transition requires the practices to give up legacy systems they have relied on and change the fundamental workflow and administration already established [3,13,28,37]. A large investment would be required for the providers to move toward new centralized Web-based scheduling systems [28].

Second, real-time Web-based scheduling lacks flexibility in the medical setting because the automatic appointment systems are not intelligent enough to handle cases not predefined. Unlike the appointment scheduling in other industries such as airline ticket booking, which has strict rules, medical appointments are tailored based on the knowledge of physicians and patients, and thus can be rather flexible [13,28]. Physicians have their own preferences in appointment patterns, whereas the booking preferences for different patients can be rather distinct and can change over time [4]. The “Mabel factor” depicts a situation in which a scheduler knows how to balance the practice’s available resources and human factors such as physicians’ preferences and patients’ needs [3,13]. It is challenging for real-time Web-based scheduling systems to achieve the same level of flexibility. In reality, physicians have to give up their preferred scheduling patterns to accommodate the simplified real-time scheduling rules [3,13].

Third, safety is a concern. It is challenging to triage patients who made appointments through real-time Web-based appointment systems. Patients may misuse Web-based appointment systems for urgent conditions that need to be handled immediately by an emergency room or urgent care [13,20]. Because schedulers are no longer involved in the appointment process, the systems should be capable of triaging patients and stratifying their risks accurately. Some practices

just display static warning messages on their Web presence to stop patients from using their appointment systems for urgent conditions [13]. Some real-time systems still rely on human reviewers to screen for possible emergencies [3]. Very few real-time appointment systems reported in the literature can automatically identify emergency conditions [2].

Finally, many providers have a fear of losing control of their appointment systems, as they think patients may abuse the systems [20,23,44]. For example, patients may book appointment slots and end up with no-shows or late cancellations. As a result, valuable clinical time would be wasted. However, this issue can be addressed by enforcing predefined appointment rules, such as rules for cancellation and a penalty for no-shows [37]. Providers can also block out appointment slots and limit visit types to accommodate their schedules [20]. Blocking patients with no-show history and collecting copay up front when making an appointment can discourage no-shows [37]. Automatically generated email- or message-based reminders can also help reduce no-shows [37]. Some practices refuse to expose physicians’ open time slots, because they believe that patients might think the physicians do not work hard enough when they see many openings [21].

In addition to the four main barriers, studies found that the following common problems from the patient side considerably affect the adoption of Web-based scheduling: unawareness of the Web-based appointment service, low penetration and distrust of the Internet, low computer skills, and the preference for verbal communications [1,8,30,31].

Limitations

This review has a few limitations. First, the collection of literature has a long time span ranging from 1990 to 2016. With the rapid development of information technology, many systems, especially those implemented in 1990s and early 2000s, experienced significant changes after they were introduced and reported. Some of the original services have been discontinued and replaced with other services [24], whereas some practices have switched software service vendors [28].

Second, many studies lack statistical research designs and have used multiple interventions at once. Although there are many improved metrics reported in the literature, it is difficult to determine whether these improvements are solely resulted from the implementation of the Web-based appointment systems. In addition, as many reported Web-based appointment services are components of health care Web services or patient portals, it is possible that the positive changes could be attributed to other components of the system.

Third, several studies have discrepant and even contradicting results. This is because the studies are from various sources with differences in care type, patient population, study period, and study design. Therefore, it is hard to compare their results systematically.

Fourth, many studies failed to report the information about assessment methods used in their studies, making it hard to judge their findings.

Finally, this work only reviews Web-based scheduling systems reported in the academic literature and does not reflect all systems available in the market.

Conclusions

In this study, we sought evidence from the literature to discuss the benefits and challenges of implementing Web-based medical appointment systems. Compared with traditional appointment methods, Web-based appointment scheduling is more patient-centered and has many advantages due to improved access. After implementing Web-based appointment systems, many practices have shown positive changes such as reduced no-show rate, decreased staff labor, decreased waiting time, and improved patient satisfaction.

Although these changes suggest Web-based appointment systems could produce positive outcomes, this assertion should be further reinforced by more sophisticated study designs. As in some studies, the Web-based appointment services are

components of portals and it is hard to measure their impacts statistically. Some studies reported results without controlling for other factors. It is possible that the positive outcomes are produced by the other factors or by the combination of the Web-based appointment systems and the other factors.

Providers and patients both have reasons for the slow adoption of Web-based appointment scheduling. Cost, flexibility, safety, and integrity are major reasons discouraging providers from using Web-based scheduling. Patients' reluctance to adopt Web-based appointment scheduling is mainly influenced by their past experiences using computers and the Internet, as well as their communication preferences.

Overall, the literature suggests a growing trend for the adoption of Web-based appointment systems. The findings of this review suggest that there are benefits to a variety of patient outcomes from Web-based scheduling interventions with the need for further studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of the 21 Web-based scheduling systems.

[[PDF File \(Adobe PDF File\), 50KB - jmir_v19i4e134_app1.pdf](#)]

References

1. Zhang X, Yu P, Yan J. Patients' adoption of the e-appointment scheduling service: a case study in primary healthcare. *Stud Health Technol Inform* 2014;204:176-181. [Medline: [25087546](#)]
2. Jones R, Menon-Johansson A, Waters AM, Sullivan AK. eTriage - a novel, web-based triage and booking service: enabling timely access to sexual health clinics. *Int J STD AIDS* 2010;21(1):30-33. [doi: [10.1258/ijsa.2008.008466](#)] [Medline: [19884355](#)]
3. -. Online scheduling applications may improve customer satisfaction, but setup is not always easy. *Internet Healthc Strateg* 2004;6(5):1-5. [Medline: [15181764](#)]
4. Wang W, Gupta D. Adaptive appointment systems with patient preferences. *Manuf Serv Oper Manag* 2011;13(3):373-389. [doi: [10.1287/msom.1110.0332](#)]
5. Tang PC, Black W, Buchanan J, Young CY, Hooper D, Lane SR, et al. PAMFOnline: integrating EHealth with an electronic medical record system. In: *AMIA Annu Symp Proc. 2003 Presented at: AMIA Annual Symposium; November 8, 2003; Washington, DC p. 644-648 URL: <http://europepmc.org/abstract/MED/14728253>*
6. Denizard-Thompson NM, Feiereisel KB, Stevens SF, Miller DP, Wofford JL. The digital divide at an urban community health center: implications for quality improvement and health care access. *J Community Health* 2011;36(3):456-460. [doi: [10.1007/s10900-010-9327-5](#)] [Medline: [21086028](#)]
7. Zhang X, Yu P, Yan J, Hu H, Gourea N. Patients' perceptions of web self-service applications in primary healthcare. *Stud Health Technol Inform* 2012;178:242-249. [Medline: [22797048](#)]
8. Siddiqui Z, Rashid R. Cancellations and patient access to physicians: ZocDoc and the evolution of e-medicine. *Dermatol Online J* 2013;19(4):14. [Medline: [24021373](#)]
9. Clingan SA. Going online: the role of web-based initiatives in health information technology. *J Med Pract Manage* 2011;26(4):225-227. [Medline: [21506462](#)]
10. Healthit. What is a patient portal? URL: <https://www.healthit.gov/providers-professionals/faqs/what-patient-portal> [accessed 2016-09-26] [WebCite Cache ID 6kqTitGM1]
11. Healthit. Meaningful use definition and objectives URL: <https://www.healthit.gov/providers-professionals/meaningful-use-definition-objectives> [accessed 2016-10-03] [WebCite Cache ID 6kzyM99gX]
12. HIMSS. Using patient portals to achieve Meaningful Use Stage 2 URL: <http://www.himss.org/using-patient-portals-achieve-meaningful-use-ep-edition?ItemNumber=35966> [accessed 2016-10-03] [WebCite Cache ID 6kzz2tUpN]
13. -. Real-time scheduling faces operational challenges. *Internet Healthc Strateg* 2005;7(1):1-6. [Medline: [15712748](#)]

14. Abbott KC, Boocks CE, Sun Z, Boal TR, Poropatich RK. Walter Reed Army Medical Center's internet-based electronic health portal. *Mil Med* 2003;168(12):986-991. [Medline: [14719622](#)]
15. Cayirli T, Veral E. Outpatient scheduling in health care: a review of literature. *Prod Oper Manag* 2003;12(4):519-549. [doi: [10.1111/j.1937-5956.2003.tb00218.x](#)]
16. Gupta D, Denton B. Appointment scheduling in health care: challenges and opportunities. *IIE Transactions* 2008;40(9):800-819. [doi: [10.1080/07408170802165880](#)]
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](#)] [Medline: [19621072](#)]
18. Romano M. Not just a Web site. In response to consumers as well as competitors, hospitals are getting more sophisticated with their Internet presence. *Mod Healthc* 2003;33(21):22-26. [Medline: [12800289](#)]
19. Knight A, Lembke T. Appointment zen - shaping demand and matching capacity. *Aust Fam Physician* 2014;43(4):234-238 [FREE Full text] [Medline: [24701629](#)]
20. Lowes R. Let patients book their own appointments? *Med Econ* 2006;83(11):27-28. [Medline: [16827528](#)]
21. Friedman JP. Internet patient scheduling in real-life practice. *J Med Pract Manage* 2004;20(1):13-15. [Medline: [15500015](#)]
22. Federowicz M. Improving access to improve quality: evaluation of an organizational innovation. *Find Brief* 2008;11(5):1-4. [Medline: [18807646](#)]
23. Farr C. Dentistry takes the cybercure: scheduling, consultations, records move to the net. *Dent Today* 2000;19(5):106-113. [Medline: [12524769](#)]
24. Illman J. 'Expedia' system for appointments. *Health Serv J* 2013;123(6351):9. [Medline: [23879123](#)]
25. Walters BA, Danis K. Patient Online at Dartmouth-Hitchcock - interactive patient care web site. In: *AMIA Annu Symp Proc*. 2003 Presented at: AMIA Annual Symposium; December 8, 2003; Washington, DC p. 1044 URL: <http://europepmc.org/abstract/MED/14728547>
26. Lowes R. Phones driving you crazy? Try clinical messaging. *Med Econ* 2004;81(6):65, 69-72, 76. [Medline: [15077492](#)]
27. -. Appointments. Patient choice is the only option. *Health Serv J* 2009;119(6183):27. [Medline: [20131473](#)]
28. Baldwin FD. Book 'em. Enterprisewide scheduling presents challenges, but CIOs will find it worth the time and effort. *Healthc Inform* 2001;18(9):37-42. [Medline: [11569223](#)]
29. Judd J. The benefits of a practice-managed Web site. *MGMA Connex* 2002;2(4):25-26. [Medline: [11949523](#)]
30. Zhang X, Yu P, Yan J, Ton A M Spil I. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. *BMC Health Serv Res* 2015;15:71 [FREE Full text] [doi: [10.1186/s12913-015-0726-2](#)] [Medline: [25885110](#)]
31. Cao W, Wan Y, Tu H, Shang F, Liu D, Tan Z, et al. A web-based appointment system to reduce waiting for outpatients: a retrospective study. *BMC Health Serv Res* 2011;11:318 [FREE Full text] [doi: [10.1186/1472-6963-11-318](#)] [Medline: [22108389](#)]
32. Chen S, Liu S, Li S, Yen DC. Understanding the mediating effects of relationship quality on technology acceptance: an empirical study of e-appointment system. *J Med Syst* 2013;37(6):9981. [doi: [10.1007/s10916-013-9981-0](#)] [Medline: [24141491](#)]
33. Ross JD, Copas A, Stephenson J, Fellows L, Gilleran G. Optimizing information technology to improve sexual health-care delivery: public and patient preferences. *Int J STD AIDS* 2007;18(7):440-445. [doi: [10.1258/095646207781147210](#)] [Medline: [17623499](#)]
34. Ralston JD, Martin DP, Anderson ML, Fishman PA, Conrad DA, Larson EB, et al. Group health cooperative's transformation toward patient-centered access. *Med Care Res Rev* 2009;66(6):703-724. [doi: [10.1177/1077558709338486](#)] [Medline: [19549993](#)]
35. Christie B. Scottish GP practices need to beef up online presence, think tank says. *Br Med J* 2014;349:g5308. [Medline: [25163731](#)]
36. Manji I. Solving your scheduling woes. *J Can Dent Assoc* 1991;57(9):713-715. [Medline: [1913303](#)]
37. Craig A. Self-service appointments. *Adv Nurse Pract* 2007;15(10):24-25. [Medline: [19998876](#)]
38. McKenzie S. Patient communication: moving beyond 2 tin cans and a string. *Dent Today* 2005;24(11):148, 150. [Medline: [16358813](#)]
39. Maifredi G, Orizio G, Bressanelli M, Domenighini S, Gasparotti C, Perini E, et al. Italian hospitals on the web: a cross-sectional analysis of official websites. *BMC Med Inform Decis Mak* 2010;10:17 [FREE Full text] [doi: [10.1186/1472-6947-10-17](#)] [Medline: [20359366](#)]
40. Parmar V, Large A, Madden C, Das V. The online outpatient booking system 'Choose and Book' improves attendance rates at an audiology clinic: a comparative audit. *Inform Prim Care* 2009;17(3):183-186 [FREE Full text] [Medline: [20074431](#)]
41. Santana S, Lausen B, Bujnowska-Fedak M, Chronaki C, Kummervold PE, Rasmussen J, et al. Online communication between doctors and patients in Europe: status and perspectives. *J Med Internet Res* 2010;12(2):e20 [FREE Full text] [doi: [10.2196/jmir.1281](#)] [Medline: [20551011](#)]

42. Lau AY, Sintchenko V, Crimmins J, Magrabi F, Gallego B, Coiera E. Impact of a web-based personally controlled health management system on influenza vaccination and health services utilization rates: a randomized controlled trial. *J Am Med Inform Assoc* 2012;19(5):719-727 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2011-000433](#)] [Medline: [22582203](#)]
43. Chen T. Ubiquitous multicriteria clinic recommendation system. *J Med Syst* 2016;40(5):113. [doi: [10.1007/s10916-016-0469-6](#)] [Medline: [26984357](#)]
44. Riddell S. Technology: the race to get the NHS online. *Health Serv J* 2012;122(6319):32. [Medline: [23074883](#)]
45. Prakash B. Patient satisfaction. *J Cutan Aesthet Surg* 2010;3(3):151-155 [[FREE Full text](#)] [doi: [10.4103/0974-2077.74491](#)] [Medline: [21430827](#)]
46. Camacho F, Anderson R, Safrit A, Jones AS, Hoffmann P. The relationship between patients' perceived waiting time and office-based practice satisfaction. *N C Med J* 2006;67(6):409-413. [Medline: [17393701](#)]
47. Sampson F, Pickin M, O'Cathain A, Goodall S, Salisbury C. Impact of same-day appointments on patient satisfaction with general practice appointment systems. *Br J Gen Pract* 2008;58(554):641-643 [[FREE Full text](#)] [doi: [10.3399/bjgp08X330780](#)] [Medline: [18801282](#)]
48. Ssl.gstatic. The digital journey to wellness 2012 Google/Compete hospital study URL: https://ssl.gstatic.com/think/docs/the-digital-journey-to-wellness-hospital-selection_research-studies.pdf [accessed 2016-10-04] [[WebCite Cache ID 6kzzCkacW](#)]
49. Paré G, Trudel M, Forget P. Adoption, use, and impact of e-booking in private medical practices: mixed-methods evaluation of a two-year showcase project in Canada. *JMIR Med Inform* 2014;2(2):e24 [[FREE Full text](#)] [doi: [10.2196/medinform.3669](#)] [Medline: [25600414](#)]

Abbreviations

GP: general practitioner

MU: meaningful use

PRISMA: preferred reporting items for systematic reviews and meta-analyses

SaaS: software as a service

WAS: Web-based appointment system

Edited by G Eysenbach; submitted 07.10.16; peer-reviewed by D Miller, P Yu, X Zhang; comments to author 17.11.16; revised version received 31.12.16; accepted 04.03.17; published 26.04.17.

Please cite as:

Zhao P, Yoo I, Lavoie J, Lavoie BJ, Simoes E

Web-Based Medical Appointment Systems: A Systematic Review

J Med Internet Res 2017;19(4):e134

URL: <http://www.jmir.org/2017/4/e134/>

doi: [10.2196/jmir.6747](#)

PMID: [28446422](#)

©Peng Zhao, Illhoi Yoo, Jaie Lavoie, Beau James Lavoie, Eduardo Simoes. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Timely Digital Patient-Clinician Communication in Specialist Clinical Services for Young People: A Mixed-Methods Study (The LYNC Study)

Frances Griffiths^{1,2}, MBBS, PhD[‡]; Carol Bryce¹, PhD; Jonathan Cave³, PhD; Melina Dritsaki⁴, PhD; Joseph Fraser⁵, MA; Kathryn Hamilton⁶, PhD; Caroline Huxley¹, PhD; Agnieszka Ignatowicz¹, PhD; Sung Wook Kim¹, MSc; Peter K Kimani¹, PhD; Jason Madan¹, PhD; Anne-Marie Slowther¹, PhD; Mark Sujan¹, PhD; Jackie Sturt⁶, PhD

¹Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, United Kingdom

²Centre for Health Policy, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

³Department of Economics, University of Warwick, Coventry, United Kingdom

⁴Oxford Clinical Trials Research Unit, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom

⁵Patient and Public Involvement, Coventry, United Kingdom

⁶Florence Nightingale Faculty of Nursing and Midwifery, King's College London, London, United Kingdom

[‡]Written on behalf of LYNC study authorship group

Corresponding Author:

Frances Griffiths, MBBS, PhD

Division of Health Sciences

Warwick Medical School

University of Warwick

Gibbet Hill Road

Coventry, CV4 7AL

United Kingdom

Phone: 44 24765 ext 22534

Email: f.e.griffiths@warwick.ac.uk

Abstract

Background: Young people (aged 16-24 years) with long-term health conditions can disengage from health services, resulting in poor health outcomes, but clinicians in the UK National Health Service (NHS) are using digital communication to try to improve engagement. Evidence of effectiveness of this digital communication is equivocal. There are gaps in evidence as to how it might work, its cost, and ethical and safety issues.

Objective: Our objective was to understand how the use of digital communication between young people with long-term conditions and their NHS specialist clinicians changes engagement of the young people with their health care; and to identify costs and necessary safeguards.

Methods: We conducted mixed-methods case studies of 20 NHS specialist clinical teams from across England and Wales and their practice providing care for 13 different long-term physical or mental health conditions. We observed 79 clinical team members and interviewed 165 young people aged 16-24 years with a long-term health condition recruited via case study clinical teams, 173 clinical team members, and 16 information governance specialists from study NHS Trusts. We conducted a thematic analysis of how digital communication works, and analyzed ethics, safety and governance, and annual direct costs.

Results: Young people and their clinical teams variously used mobile phone calls, text messages, email, and voice over Internet protocol. Length of clinician use of digital communication varied from 1 to 13 years in 17 case studies, and was being considered in 3. Digital communication enables timely access for young people to the right clinician at the time when it can make a difference to how they manage their health condition. This is valued as an addition to traditional clinic appointments and can engage those otherwise disengaged, particularly at times of change for young people. It can enhance patient autonomy, empowerment and activation. It challenges the nature and boundaries of therapeutic relationships but can improve trust. The clinical teams studied had not themselves formally evaluated the impact of their intervention. Staff time is the main cost driver, but offsetting savings are likely elsewhere in the health service. Risks include increased dependence on clinicians, inadvertent disclosure of confidential

information, and communication failures, which are mostly mitigated by young people and clinicians using common-sense approaches.

Conclusions: As NHS policy prompts more widespread use of digital communication to improve the health care experience, our findings suggest that benefit is most likely, and harms are mitigated, when digital communication is used with patients who already have a relationship of trust with the clinical team, and where there is identifiable need for patients to have flexible access, such as when transitioning between services, treatments, or lived context. Clinical teams need a proactive approach to ethics, governance, and patient safety.

(*J Med Internet Res* 2017;19(4):e102) doi:[10.2196/jmir.7154](https://doi.org/10.2196/jmir.7154)

KEYWORDS

digital communication; long-term conditions; young people; digital health care; patient communication; NHS; National Health Service

Introduction

Young people living with long-term conditions are vulnerable to disengagement from health care, which endangers their current and future health [1-5]. Health service factors affecting young peoples' engagement with health care include poor patient-clinician communication, inflexible access to people and information, lack of person-centered health care, and the need for continuity and relationship development [2,6-9].

In the United Kingdom, 90% of young people aged 16-24 years own a smartphone [10]. Studies have reported requests from young people to be able to communicate via email, text message, and social media with their health care team [6,11]. There are reports of specialist clinical teams using digital channels for monitoring and information sharing [12,13]. In the United Kingdom, government policy and investment is driving the digitization of the National Health Service (NHS) [14,15]. With the rollout of NHSmail 2 [16], NHS clinicians now have access to secure email and other digital channels for communicating with patients on clinical matters.

Evidence for effectiveness on health outcome of the use of digital channels with patients on clinical matters is not strong. Prior to starting our project, we found 16 systematic reviews [17-32] and 1 clinical review [33] published from 2010 to 2012 on the effectiveness of digital communication between clinicians and patients with long-term conditions, where the long-term condition was relevant to young people (only 2 reviews focused on young people [25,32]). Evidence of an impact on clinical outcomes was equivocal, although no trials reported poorer outcomes in the intervention arm. The reporting of interventions was generally poor. The systematic reviews identified the following gaps in evidence: how digital communication might work [19,21,27,30,33,34], in particular examining the function of the communication rather than the communication channel [33]; what was important to patients and clinicians [19,20,22,23,25-27,32,33,35]; cost and resource use [17,19,21,23,25,27-29]; risks including privacy and data security [19,22,23,27,33]; the need to focus on widely used digital communication rather than being future focused [29]; and research to inform policy, practice, and implementation or rollout [21,22,28].

Given the poor quality of the evidence, and the gaps in the evidence about how digital communication might work, its

value to patients and clinicians, and its cost and risks, we had the following aims. First, we wanted to identify how the use of digital channels for communication between young people and their clinicians was addressing the health service factors influencing young people's engagement with health care, and the perceived impact and value of the digital communication. Second, we aimed to identify cost, ethical, and patient safety issues that need to be considered in the NHS policy-driven rollout of digital communication. To meet both these aims, we studied NHS clinicians and young people with long-term conditions requiring specialist care, who were already using, or considering using, digital channels for communication about clinical issues, where the communication was two-way (synchronous or asynchronous), and where both the clinician and the young person could be mobile.

Methods

This was an observational mixed-methods study of cases [36] undertaken in the UK NHS, where services are free at the point of delivery.

Case Study Sampling

We used multiple strategies to identify clinical teams. First, between December 2013 and February 2014, using Google (Google Inc, Mountain View, CA, USA), we searched the Internet for reports of the use of digital communication with patients in the NHS using the keywords "e-health," "telehealth," "telemedicine," "digital communication," "young people," and "young persons." We scrutinized the first 35 pages of each search for relevant reports. Further information was then sourced from individual NHS Trust websites, documents, and reports and by contacting key individuals. Second, we listed the project on the UK National Institute for Health Research (NIHR) portfolio inviting participation. Third, we contacted clinicians we knew personally or had encountered at applied health conferences and asked them to distribute information about the project to their networks. Fourth, clinical teams expressing interest in the study were asked to pass on the study information to potentially interested colleagues.

Study inclusion criteria were that (1) the clinical team was providing specialist care for young people (age 16-24 years) with long-term conditions (eg, sickle cell, liver disease, cystic fibrosis, cancer, or mental health issues), (2) the team had interest in the use of two-way digital communications with the

young people, and (3) the long-term condition had considerable cost implications for the NHS.

We sampled 20 teams purposively for diversity of clinical condition, use of digital communication with patients, size, and geographic location. Studying 20 teams ensured both diversity and anonymity of study teams. Anonymity was important as; at the time of undertaking the study, some study teams may have been in breach of information governance policies.

We obtained ethical approval (14/WM/0066) from National Research Ethics Service Committee West Midlands - The Black Country.

Data Collection

We collected data between November 2014 and March 2016. Prior to commencing fieldwork with each clinical team, we requested to see any in-house evaluations of their digital communication with patients that they had done. During recruitment of clinical teams, some teams mentioned that they were using digital channels without formal approval from their Trust. We therefore sought to interview the Trust information governance specialist before collecting data from any clinical teams.

To understand how digital clinical communication was used, including its perceived impact, and to identify issues related to ethics and patient safety, we observed and interviewed clinical team members at all study sites. We collected data within a 2-week data collection period during the team's day-to-day work, recruiting as many team members as were prepared to participate. To explore the cost of using digital communication with patients, we collected data on equipment and clinician time spent on the use of digital communication with patients and its cost. To gather these data, we developed a questionnaire based on early interview data and used this as part of clinical team interviews.

We also recruited for interview young people aged 16-24 years under the care of the clinical team and due to be seen by the team during the 2-week data collection period. The young people were sent study information prior to their scheduled encounter. The clinical team or the study researcher approached each young person at the scheduled encounter—usually before the appointment time—to take consent and confirm interview arrangements. Those not attending were further contacted to request participation. At the interview, we asked the young people about their use of digital communication with the clinical team, its impact on their day-to-day life and ability to manage their health condition, and ethical and safety issues. To explore the value young people placed on digital access to their clinical team, we asked them what they would be willing to pay for the service. Young people were each offered a £20 store voucher as a thank-you token. Interviews used any communication channel preferred by the participant, such as phone, in person, or email. We recruited for interview until we were confident we were not gaining any new data from the young people on their experience and views of the use of digital communication with clinicians at their clinic.

Data Management and Analysis

Observation notes were taken, and then typed up and expanded immediately after observation. Interviews were audio-recorded or notes taken, typed up, and expanded. We made reflective notes after each observation or interview. Recordings were transcribed and checked for accuracy. All identifiers were removed, and data were identified with a site and participant number. Independent coding was undertaken on 20% of all coding, and discrepancies were discussed. Quality checks were undertaken on data entry of survey data.

We coded all qualitative data for the major prespecified analysis themes related to our aims: (1) how digital communication with patients is used and its perceived impact, (2) the value of the communication to young people, (3) its ethical impact, and (4) patient safety and governance issues and their mitigation. Within these coded data, we undertook (1) further analysis identifying the mechanisms by which the digital communication had an immediate impact, and its context [37], (2) thematic analysis [38], (3) thematic analysis informed by theory [39], and (4) thematic analysis with an established safety framework [40].

Using staff questionnaire data, we calculated, for each respondent, the annual direct costs associated with digital communication with patients. We used NHS Agenda for Change pay scales 2014-2015 [41] for salaries and University of Warwick information technology service price lists for employer-provided equipment, annualized assuming a 3-year life span and a discount rate of 3.5% [42]. We estimated total costs at each site where over 50% of clinical team members responded to the survey. Where data permitted, we estimated cost per patient based on the size of each clinic's patient list.

Patient, Public, and Stakeholder Involvement

We explored early research ideas with an experienced Patient, Public, and Stakeholder Involvement group with which we had worked for over 10 years (Warwick Diabetes Research and Education User Group). Subsequently, to gain input from young people, 20 students (15-17 years of age) from 5 local schools collected opinions from their peers and reported this as a film [43]. Patient, Public, and Stakeholder Involvement coapplicant and coauthor JF drew on this to advise the project team about collecting data from the young people living with long-term conditions. He subsequently chaired the project management group. This group included 4 young adults and a parent of a young person living with a long-term condition, and representation from NHS Digital (UK Department of Health, Leeds, UK). They advised on recruitment, data collection procedures, analysis, and impact strategies.

Each clinical team is receiving a copy of the project report. Study results in the form of Quick Reference Guides [44] are being disseminated to patient support and advocacy groups, professional organizations, and all NHS Trusts, Health Boards, and Clinical Commissioning Groups.

Results

Study Sites and Participants

We identified 104 clinical teams (via Internet search, 15; NIHR portfolio, 7; networks and contacts, 58; contacts of already interested clinical teams, 24), of which 47 were eligible and interested in participating (see [Figure 1](#)). We initiated site setup at 25 sites and studied 20, covering 13 clinical specialties (see

[Table 1](#)). Clinic populations included children and adolescent services, transition services, young adult services, and adult services. There were 9 clinical teams in the South and East of England, 7 in the Midlands, 3 in the North of England, and 1 in Wales. At recruitment, clinical teams variously reported using with their young people the following communication channels: mobile phone calls, text messages, email, voice over Internet protocol, and personal health records. A total of 3 clinical teams used no digital communication with their young people.

Figure 1. Flowchart showing case site recruitment.

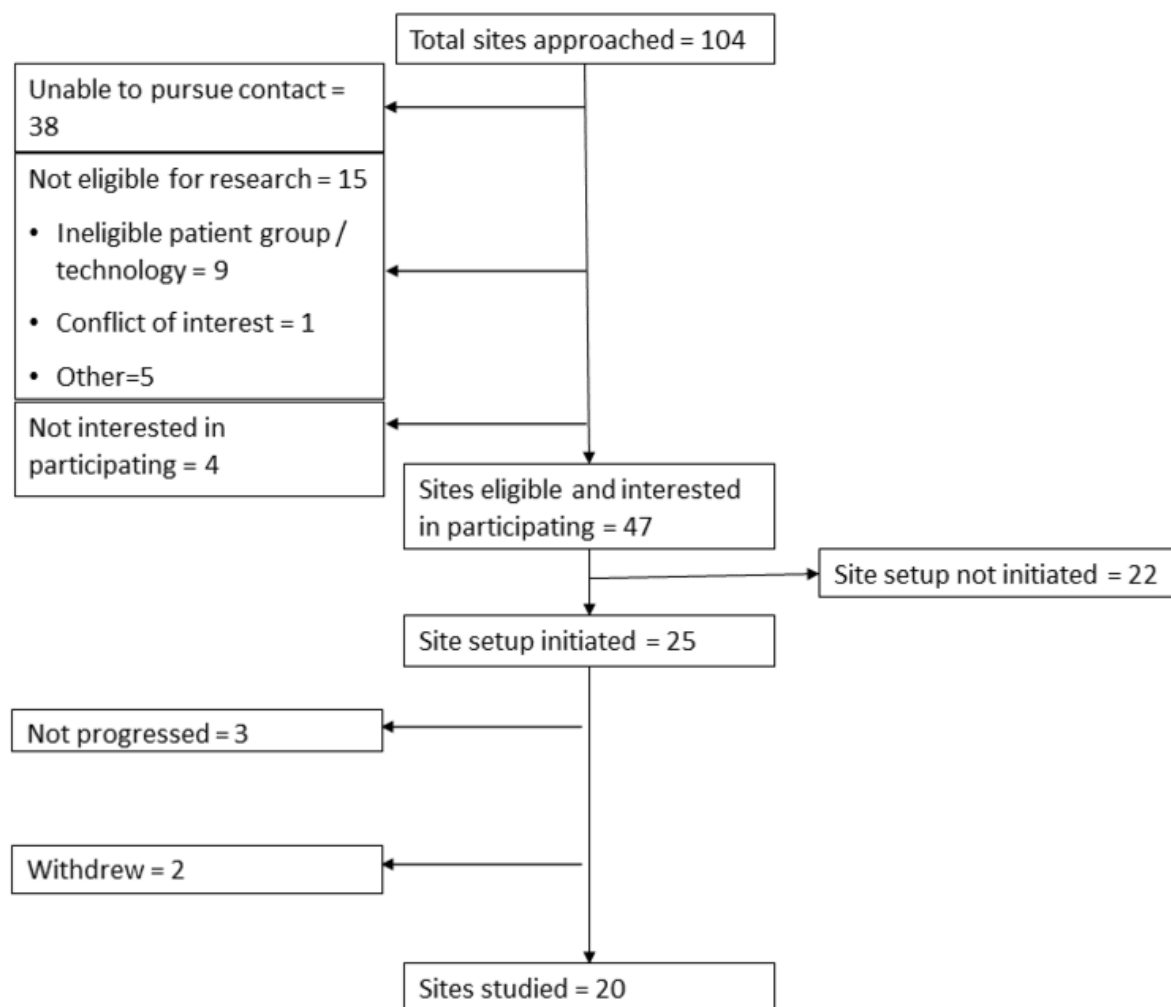


Table 1. Case study site health condition, clinic type, age group, digital communication used with patients, and data collected.

Site identifier	Clinic population ^a	Patient age range (years)	Digital communication channels used in clinic	No. of young people interviewed	No. of staff interviewed	No. of information governance specialists interviewed	No. of staff shadowed
Diabetes 1	Transition	12-19	Mobile phone, text message, email	12	8	0	5
Mental Health 1 (Early Intervention)	Age independent	>16	Mobile phone, text message, email	1	8	1	2
Cystic Fibrosis 1	Adult	>16	Email	2	5	1	1
Dermatology	Adult	>18	Email	7	4	0	2
Mental Health 2 (CAMHS ^b)	Child and adolescent	<18	None	4	11	1 ^c	4
Mental Health 3 (Out-reach team)	Child and adolescent	<18	Mobile phone, text message, VoIP ^d	5	11	1	3
Arthritis	Transition	16-25	None	16	8	1	1
Cystic Fibrosis 2	Adult	>16	Mobile phone, text message, VoIP	13	11	0	3
School nurse service	Young people	14-19	Text message, VoIP (pilot)	0	7	1	4
Kidney	Young adult	16-22	Email	7	7	0	3
Liver	Transition	12-25	Text message, email	15	12	2	7
Sickle Cell	Transition	12-24	Mobile phone, text message	10	13	2	9
Mental Health 4 (Early Intervention in Psychosis Team)	Youth	14-35	Mobile phone, text message, email	4	9	2	8
Cancer 1	Teenage and young adult	15-24	Mobile phone, text message, email	12	7	0	3
Diabetes 2	Transition	16-25	Mobile phone, VoIP	11	6	2	2
Inflammatory Bowel Disease 1	Adult	>16	Web portal, email	1	6	1	3
Inflammatory Bowel Disease 2	Adolescent	13-23	Email	13	7	0	4
HIV ^e	Adult	>18	None	9	12	0	4
Sexual Health	Adult and young people	>16	Testing kits ordered online	12	10	0	3
Cancer 2	Teenage and young adult	15-24	Mobile phone, text message, email	11	11	2	8

^aAs described by clinic staff.^bCAMHS: Child and Adolescent Mental Health Services.^cInformation governance specialist was the same person as for Mental Health 1.^dVoIP: voice over Internet protocol.^eHIV: human immunodeficiency virus.

We recruited for interview 165 young people. Interviews were undertaken by phone (n=82), face-to-face (n=41), email (n=35), Facebook (n=4), Skype (n=2), and text message (n=1). Speech-based interviews lasted 20-60 minutes, with the majority lasting approximately 30 minutes. Text message-based interviews took up to 2 weeks.

We recruited 16 information governance specialists and 173 clinical team members for interview. The clinicians included consultants, registrars, community nurses, advanced nurse practitioners, psychologists, dietitians, physiotherapists, occupational therapists, and pharmacists (7 interviews were with clinic administrators closely involved with patient care).

Interviews were undertaken face-to-face (n=158) or by phone (n=31) and lasted up to 2 hours, with the majority lasting approximately 45 minutes. We shadowed 79 clinical team members, usually for 1-2 hours, longer if appropriate (eg, when observing home visits with a clinician). Of the 173 clinical team members, 115 completed staff health economic questionnaires across 18 sites.

Timely Digital Contact Between Young People and Their Clinical Team

Young people and clinicians mostly used digital channels to be in contact at times when the contact could make a difference to how the young people managed their condition. This timely access was not scheduled, although clinicians often planned their contacts.

The channels of communication used for this timely access varied across clinical teams (see Table 1) and according to the reason for making contact. Mobile phones were used for urgent issues and when discussion was needed to resolve the clinical problem. Text messaging was used for keeping in touch, raising less-urgent concerns such as new symptoms or changing trends in home monitoring (eg, blood sugars), personal reminders about upcoming appointments, and reminders about therapy. Additional clinical team members used text messaging to make direct contact with young people where parents were involved in a young person's treatment, as this allowed issues to be raised that would not be raised in front of parents. Email was useful for sending complex information and summaries of discussion at a consultation, as the young people were then able to read and reread the information, and for sending test results where the results were routine or as expected and the individual was well known to the service. The young people emailed questions such as how to use a skin cream or fix equipment, concerns such as suitability of vaccinations for travel, photographs of their condition, such as a rash, and requests for supplies.

Although the clinical teams we studied were motivated to use timely digital communication with their young people to improve their health outcome, none of the clinical teams had evaluated the impact of its use on health outcome. However, our data revealed many mechanisms by which timely digital access improved health care and so had the potential to improve health outcome. Young people and clinicians reported that timely digital communication enhanced engagement, reduced patient anxiety, and improved trust between the young people and their clinicians. Young people felt they received personalized care and valued the continuity of care they received by being able to contact the clinicians who knew them when they needed to. The timely access prompted activation and better self-management by the young people:

I sort of just avoided doing anything really and just thought it might sort itself. But...I do need to accept the help that's out there for me...it's a lot better just being able to speak over email and then when you do need a test done you're only going into your doctors every four or five months, if that. [Young person 06, Diabetes 2]

I can email them anytime, I can get a response anytime and sort it out myself. [Young person 07, Inflammatory Bowel Disease 2]

Young people who were already engaged with their clinical team sometimes used email or text messaging to communicate about sensitive issues that they found difficult to raise face-to-face, knowing that the email or text message would prompt the clinician to raise the issue when they saw them. Some disengaged young people reengaged with their clinical team via email or text message when the young person had not responded to phone calls or regular mail.

Young people and clinicians reported examples of where timely digital communication had been used to diagnose minor issues, sometimes avoiding unnecessary clinic visits, or to treat symptoms before they became serious, thus avoiding the need for emergency care.

Timely digital access was considered, by both young people and their clinicians, to be a valuable addition to traditional clinic appointments, not a replacement. Face-to-face communication was considered important for establishing relationships and for conveying bad or potentially upsetting news. The use of digital channels for routine issues and exchange of information between appointments left more time in clinic appointments for complex issues, and so increased the value of the face-to-face consultations for both the young people and clinicians.

Digital communication was convenient for the clinicians and young people, as it avoided disruption to their other activities and sometimes avoided unnecessary consultations. The use of asynchronous communication such as email and text messaging allowed them to think about their questions or responses. Clinicians liked the opportunity to consult clinical records before responding, although where a clinician knew the patient well they did not always do this.

Value to the Young People and Cost to the NHS of Timely Digital Access

A total of 110 of the 165 young people answered the exploratory question on their willingness to pay for digital communication with their clinical team. The median willingness to pay was £5 per month (interquartile range £0-£16, maximum £120). A total of 27 young people reported being willing to pay £30 or more per month, 35 were willing to pay between £0 and £30, 30 would not be willing to pay extra, and 18 were unable or unwilling to answer this question. The reasons young people gave for wanting the service mostly related to resolving problems quickly, such as an issue with self-injection; enabling easier contact with a named clinician for continuity of care; and saving time travelling to a clinic to report progress and to hear or provide results.

Young people from one mental health site objected to the question on the basis that the service should be free, or because of a perception that payment would be discriminatory if applied only in their service. Those not willing to pay were not currently using digital channels for communicating with their clinical team, even where they were available, and thought that conventional channels could be used just as well.

Our exploratory data indicated that the mean time spent by staff per day using digital channels to communicate with young people was 76 minutes (median 45 minutes, interquartile range 0-120 minutes). The mean and median times were not typical levels of activity (see Figure 2): 33 of 115 staff (28.7%) reported using digital communication with patients “rarely” or “never,” and 25 staff (21.7%) reported using it for over 2 hours per day. Use varied by grade and profession. Medical consultants reported substantially lower use (mean time 28 minutes per day) than nurses (120 minutes per day) and physiotherapists (120

minutes per day), but consultants’ use was similar to that of dietitians (14 minutes per day) and psychologists (34 minutes per day). The major cost for providing digital communication between clinic appointments was staff time (see Table 2). Staff time was typically 90%-95% of total cost. For sites where the clinic was able to provide the size of their patient list so that cost per patient could be calculated, the reported figure was between £0 and £20 per month, with the exception of the 2 cystic fibrosis sites, where costs per patient were much higher (£73-£130).

Figure 2. Minutes per day reported by clinical team members (n=115) as spent using digital communication with patients.

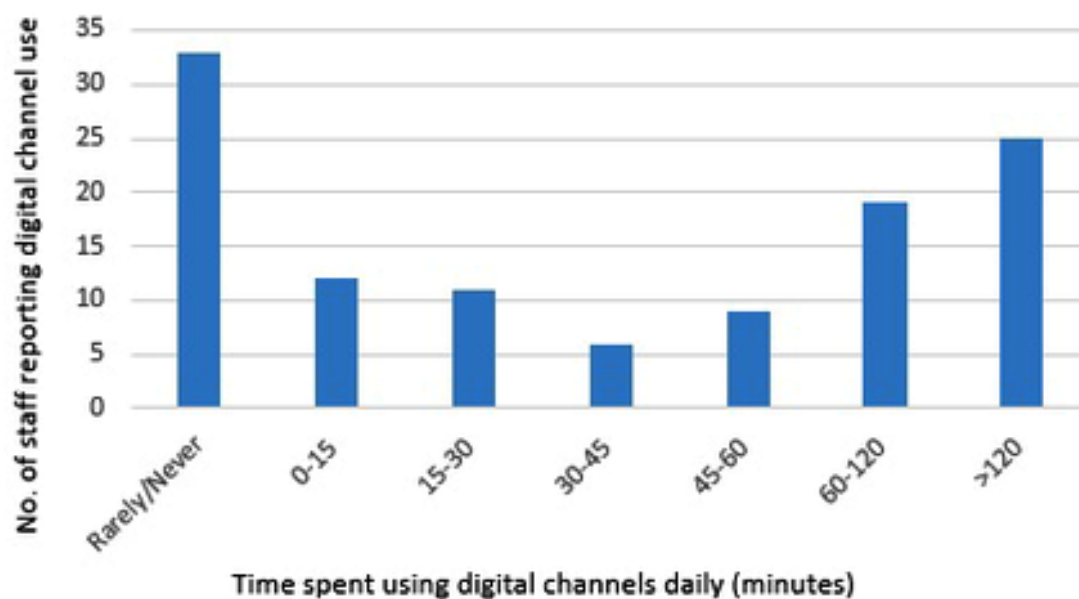


Table 2. Site-level costing analysis.

Site	No. of health economic questionnaires completed	Cost per month (£)			
		Clinical team cost	Equipment cost	Total cost	Average cost per patient
Cancer 1	5	2920	97	3017	N/A ^a
Mental Health 3 (Outreach team)	8	9230	330	9560	N/A
Arthritis	8	0	0	0	0
Kidney	6	135	26	161	16
Diabetes 1	6	2648	85	2733	4
Cystic Fibrosis 1	11	5323	383	5706	73
Sexual Health	10	3673	120	3793	N/A
HIV ^b	9	1055	51	1106	N/A
Cancer 2	11	6090	267	6357	N/A
Inflammatory Bowel Disease 1	7	3604	26	3630	3
Mental Health 2 (CAMHS ^c)	6	212	18	230	2
Liver	7	3806	71	3877	N/A
Inflammatory Bowel Disease 2	4	2672	63	2735	N/A
Cystic Fibrosis 2	3	1490	69	1559	130
Dermatology ^d	2	–	–	–	–
Mental Health 4 (Early Intervention in Psychosis Team) ^d	4	–	–	–	–
Sickle Cell ^d	6	–	–	–	–
Diabetes 2 ^d	2	–	–	–	–

^aN/A: not available.

^bHIV: human immunodeficiency virus.

^cCAMHS: Child and Adolescent Mental Health Services.

^dInsufficient data to calculate clinic costs.

Most staff reported that their workload was manageable. When asked what they would do without digital communication, they said they would spend time trying to contact young people by phone or arranging appointments for them. However, staff did not report that digital communication reduced their workload overall, and several reported an increase.

From the qualitative data we identified mechanisms by which NHS costs may be reduced through the use of digital communication with patients. These were as follows:

- Reducing costly complications of illness through early treatment
- Reducing the number of appointments young people had to attend
- Reducing “did not attend” rates
- Responding to queries, for example, for young people with sickle cell, to avoid a visit to an emergency department
- Improving response to therapy through provision of advice and support, so reducing future health care cost.

Managing Access Through Digital Channels

Although the ease of access that digital channels allowed was appreciated by both young people and clinicians, *both* were aware of the need to manage expectations. These were still being worked out by some clinical teams and their young people. Clarity about response times, working hours, and the channels of communication suitable for different purposes was considered important. Young people and clinicians wanted this information to be easily available through email and text messages, bounce-back messages, and voicemail and to be reinforced during consultations and communications. Clinical teams reported response times between a few minutes to a few days, depending on the health condition and channel of communication. Poor network coverage in some rural areas and the cost to young people of digital communication were identified as limiting digital access for some young people.

Ethical Impact of the Use of Timely Digital Communication

Digital communication has the potential to both enhance and undermine patient autonomy. Clinicians explained that it increased patient autonomy [45] by giving the young people

more control in both the management of their condition and the way in which they communicated with their clinical team. But they also noted that it may discourage some young people from taking responsibility for their own health by providing easy access to a decision maker. Young people placed more emphasis on the personalization of their care with digital communication than on increased empowerment. The ability to have more frequent contact with a specific clinician meant the clinician was more likely to know that particular young person, their circumstances, and what is important to them, so enabling the clinician to deliver person-centered care [46]:

Your relationship with the nurse is a lot easier...because they know you and they know your condition...[rather]than just another nurse that you come to see. They understand how yours is different to everybody else's. [Young person 04, Diabetes 1]

Communicating digitally reduced the power imbalance in the patient-clinician relationship, with clinicians fitting into the young person's world rather than the young person being expected to fit into the clinical world. However, there were consequences for clinician autonomy. Clinicians expressed concerns about blurring of the patient-clinician boundary:

[The patient]started sending me huge numbers of emails and chasing me a lot...I had to then think about what's a reasonable time frame for getting back to[the patient]. [Consultant 04, Mental Health 3]

Another concern was losing control over clinician information; one clinician reported how their patient had put the clinician's text messages out on social media.

The concept of a duty of care to an individual patient is enshrined in professional codes and common law [42,47]. The development of a more personalized relationship through digital communication created uncertainty for both patient and clinician about their understanding of the duty of care and its limits. Clinicians described their concerns about the patient's use of text messaging or email for communication about serious health

concerns outside of the clinic's normal working hours. They were unsure where the boundary was to the duty of care:

I was worried she [the patient] was going to do something dangerous like commit suicide or something, because she has mental health issues. And then felt awful the fact that I'd given her my email as a point of contact and then she'd reached out but it was two o'clock in the morning and of course I hadn't picked it up. [Clinical team member 01, Liver Disease]

However, across all sites, few participants were able to recall an instance where a patient had left an urgent communication that was not picked up in a timely manner.

There is an implied promise at the heart of the patient-clinician relationship that information disclosed to the clinician by the patient, or gained in the process of that patient's care, will not be disclosed to others without the patient's consent. Young people varied in their level of understanding of, and concern about, confidentiality and privacy. Clinicians were usually cautious about sending confidential data digitally, and many distinguished between a clinician sending information to the young person (risk of breach of confidentiality) and the young person sending data to the clinician (young person's choice and their responsibility).

Patient Safety

In addition to the inadvertent disclosure of sensitive information discussed above, our data revealed three other major categories of hazards from the use of timely digital communication between young people and their clinical team: communication failures; failure to record the content of the communication; and failure to consult the patient's notes prior to engaging in communication. Table 3 summarizes the causes, consequences, and current form of mitigation of these hazards. These hazards are common to all forms of clinical communication, but the ease and speed of use of digital channels magnifies the risks.

Table 3. Hazards, consequences, causes, and current form of mitigation identified by young people with long-term conditions and their clinicians using digital channels to communicate about clinical issues.

Hazard	Consequences	Causes	Current form of mitigation
Inadvertent disclosure of sensitive information	Negative effects on patient well-being; jeopardizing trust between clinician and patient	Hacking, interception of communication, loss or theft of hardware, poor usability of encrypted mail service, shared email accounts and computers, sending communication to wrong recipient, excessive distribution of communication	Limiting the use of digital communication; technical solutions; double-checking contact details; ensuring explicit or implicit patient consent
Communication failures	Failure or delay in providing relevant clinical information and advice; patients discouraged from seeking relevant advice; delays in escalation to emergency care; unnecessary escalation to emergency care; patient uncertain or anxious; clinician stressed or anxious	Not answering communication from unknown numbers, not being able to establish the patient's identity, delay in picking up or responding to messages, inability to access the Internet on mobile phones due to lack of signal or credit, poor usability of devices, difficulty expressing clearly information requests in text messages, patients downplaying seriousness of their condition in text messages, difficulty of checking correct understanding of communication content using asynchronous channels	Limiting the use of digital communication; clinician training in use of asynchronous digital channels with patients; planning for sufficient time to write and read digital communication carefully; using alternative means of emergency advice seeking; following up using a different communication channel
Failure to record content of digital communication	Other clinicians unaware of prior communication; unnecessary duplication of questions and advice given to patients; gaps in clinical record; lack of clarity for patients and clinicians about what was communicated	Digital communication not logged automatically; content of text and email messages not easily transferred to clinical notes; time consuming to record all digital communication; limited storage on communication device; lack of common understanding of how to document content of digital communication in clinical record	Treating every communication as equivalent to a face-to-face consultation; limiting the use of digital communication to forms readily integrated with patient's clinical record; restricting the use of digital communication
Failure to consult patient's notes prior to engaging in digital communication	Reliance on an incomplete understanding of patient's clinical history; duplicate or contradictory advice giving	Perceived familiarity with the patient because of frequent contact; acute problem requiring urgent response; nonclinical nature of many of the digital communications between young person and clinician	Familiarity with the patient; double-checking notes after the communication has taken place

Information governance specialists expressed a willingness to support their clinical teams in using digital communication to improve health care. They are also required to monitor adherence to NHS Trust policy. Of the 16 information governance specialists interviewed, 13 reported the existence of policies in their organizations that specifically covered text messaging, emailing, and the use of handheld mobile devices. They recognized that policies need to evolve as digital communications evolve, with a majority of those interviewed currently developing policies. A few Trusts did not permit digital communication with patients, and some information governance specialists in these Trusts were aware that it was nonetheless taking place. During their interviews, information governance specialists discussed the hazards of digital communication with patients in general terms and recognized the need for training clinical teams in its use. None of the clinical teams we studied reported that they had undertaken a formal patient ethical or safety appraisal of their service. As Table 3 describes, young people and clinicians were often left to mitigate the risks by relying on common-sense strategies (eg, escalation by other means for emergencies) and by restricting the use of digital communication (eg, restricting it to nonurgent matters). A trusting relationship between the young people and their clinical team was important for mitigating both patient safety and ethical risks.

Discussion

Principal Findings

The provision of timely digital communication between young people with long-term conditions and their clinicians is addressing the health system factors that in the past have led to these young people disengaging from health services. Digital channels enable contact between young people and their clinical teams when this contact can make a difference to how the patient manages their condition. This digital service improves the patient's experience of and engagement with care and prompts greater levels of self-management. Offering both digital and face-to-face contact is important to young people and clinicians. It also has the potential to reduce health care inequalities by engaging young people who are otherwise hard to reach. Young people value the enhanced access. Providing this access increases staff workload. The cost of providing this access is mostly attributable to staff costs. This cost is not immediately apparent to patients in the NHS, where services are free at the point of delivery. There is potential for offsetting savings from reduced adverse events and enhanced long-term outcomes, but these will not generally accrue to the service facing increased initial costs. As NHS policy prompts further rollout of digital access between patients and clinicians, there are ethical, governance, and patient safety issues to be considered by the patients, clinical teams, and their service organizations. These

issues are currently mitigated by patients and clinicians working together in relationships of trust.

Strengths and Limitations

Our study findings are likely to apply to adult populations, as they do not relate specifically to the age or clinical condition of the young people, particularly as smartphone ownership among older people is rapidly increasing [10], giving them easier access to text messaging and email. We included a relatively large number of case study sites for mixed-methods data collection covering a wide range of clinical conditions. Clinical teams were using widely used digital channels. The study captured the perspectives of many young people living with long-term conditions and those of a wide range of clinical team members. The generalizability of our study findings is limited by the study design, as with any empirical study of practice-initiated behavior. We may have detected only what is most obvious and may have missed more subtle issues. Our participants may have reported particularly positive or, perhaps to a lesser extent, negative experiences. We were unable to recruit young people who were not engaging with their health care provider. Interviews did not elicit explicit ethical reflection. Clinicians found it difficult to estimate their workload during the interviews, and we did not attempt to collect data about digital communication activity via their digital communication system. Some clinics were unable to provide the size of their patient list. We did not attempt to collect cost data for patients as we had no comparator group. We were able to estimate the direct costs associated with the delivery of providing digital communication with patients, but we did not have comparator data to estimate the costs incrementally. While we identified qualitatively how this communication could lead to NHS savings, we did not have accurate-enough incremental outcome data to quantify the savings or cost-effectiveness. The study was undertaken in the NHS, where costs of care are not made clear to individual patients.

Comparison With Prior Work

The clinical teams we studied did not need convincing of the benefits of implementing timely digital clinical communication [48]. There is evidence that motivation and enthusiasm make a difference when implementing digital clinical interventions [49]. However, despite the importance of evaluating their digital access service for justifying its further development [50], none of the clinical teams had done so. Systematic reviews of intervention studies, usually focused on specific disease areas and published since we were preparing for this study, have

mostly found some benefit from the use of digital channels for communication between patient and clinical teams, but some found no benefit. One review of text messaging for diabetes found no clear impact on glycemic control and self-management [51], whereas other reviews have found that telehealth improves glycemic control [52,53]. Systematic reviews on mental health found mostly positive findings [54,55]. A systematic review of telehealth to support family caregivers of people with chronic disease concluded that telehealth can positively affect care [56].

Conclusions

Our study findings suggest how the introduction of timely access for patients to their clinical team using digital channels could be safely and ethically achieved, to improve the experience of health care and enhance self-management. First, implement the service initially with patients where there is an existing relationship of trust between patient and clinical team. Examples include patients with long-term conditions and women in the antenatal period. Second, focus on delivery to the population of patients where the service is responding to an identified need—for example, patients in transition between services (such as the young people we studied); patients in transition between treatments (eg, a person with diabetes starting insulin), or before or after treatment (eg, liver transplant); and patients in transition in their life (eg, starting university)—and monitor impact on staff workload. Third, prior to introducing the timely digital access, clinical teams need to work out how they will proactively manage safety [57] (eg, inadvertent information disclosure) and ethical issues (eg, role boundaries), and use their information governance specialists as a resource (eg, provision of training). Improvements in the technological infrastructure (eg, NHSmail 2 [16]) have solved, or will solve, some safety and ethical concerns, but others need to be addressed within the clinical team, often with simple measures such as a timetable of clinical team members' availability within email signatures (see LYNC study Quick Reference Guides designed to support clinical team discussions on these issues) [44].

Introducing timely access for patients to their clinical team using digital channels will require trained leadership [15] and patient involvement [50].

An experimental research design is needed to evaluate the impact of timely digital access to clinical teams on health outcome and health care provision. The research will need to take account of the nature of the existing relationship between patient and clinician.

Acknowledgments

This project was funded by the NIHR, Health Services and Delivery Research Programme (project number 12/209/51).

Department of Health disclaimer: the views and opinions expressed herein are those of the authors and do not necessarily reflect those of the Health Services and Delivery Research Programme, NIHR, NHS, or the Department of Health.

Authors' Contributions

FG and JS led the study from design through to writing up study reports. FG drafted this paper. JC, JF, JM, AMS, and MS contributed to study design, and supervised data collection and analysis. CB, MD, KH, CH, and PKK undertook data collection and analysis. AI and SWK contributed to data analysis. All authors contributed to writing the paper.

LYNC study authorship group: Xavier Armoiry, Helen Atherton, Abigail Buckle, Rachel Court, Patrick Elder, Eleni Karasouli, Mike May, Paul A Sutcliffe, Isabelle Svahnstrom, Frances Taggart, Ayesha Uddin, and Alice Verran, Warwick Medical School, University of Warwick; Thandiwe Rebecca Dliwayo, Vera Forjaz, and Richard Goodwin, Florence Nightingale Faculty of Nursing and Midwifery, King's College London; Harjit Matharu and Sailesh Sankaranarayanan, University Hospitals Coventry and Warwickshire NHS Trust; Luhanga Musumadi, Guy's and St Thomas' NHS Foundation Trust; Moli Paul, Coventry and Warwickshire Partnership Trust; and Gyanu Raut, King's College Hospital NHS Foundation Trust, UK

Conflicts of Interest

FG and AMS received grants from the UK National Institute of Health Research during the conduct of the study. JF received personal fees from NHS England, outside the submitted work.

References

1. Watson AR. Non-compliance and transfer from paediatric to adult transplant unit. *Pediatr Nephrol* 2000 Jun;14(6):469-472. [Medline: [10872185](#)]
2. Heery E, Sheehan AM, While AE, Coyne I. Experiences and outcomes of transition from pediatric to adult health care services for young people with congenital heart disease: a systematic review. *Congenit Heart Dis* 2015;10(5):413-427. [doi: [10.1111/chd.12251](#)] [Medline: [25659600](#)]
3. Witherspoon D, Drotar D. Correlates of adherence to prophylactic penicillin therapy in children with sickle cell disease. *Child Health Care* 2006 Dec;35(4):281-296. [doi: [10.1207/s15326888chc3504_1](#)]
4. Singh SP. Transition of care from child to adult mental health services: the great divide. *Curr Opin Psychiatry* 2009 Jul;22(4):386-390. [doi: [10.1097/YCO.0b013e32832c9221](#)] [Medline: [19417667](#)]
5. National Paediatric Diabetes Audit Project Board, Royal College of Paediatrics and Child Health. National paediatric diabetes audit report 2010-11. London, UK: Royal College of Paediatrics and Child Health; 2012 Sep. URL: http://www.rcpch.ac.uk/system/files/protected/page/NPDA%20Annual%20Report_25%2009%2012%20for%20web_0.pdf [accessed 2017-03-24] [WebCite Cache ID 6oz3XBB1G]
6. Dovey-Pearce G, Hurrell R, May C, Walker C, Doherty Y. Young adults' (16-25 years) suggestions for providing developmentally appropriate diabetes services: a qualitative study. *Health Soc Care Community* 2005 Sep;13(5):409-419. [doi: [10.1111/j.1365-2524.2005.00577.x](#)] [Medline: [16048529](#)]
7. Kennedy A, Sloman F, Douglass JA, Sawyer SM. Young people with chronic illness: the approach to transition. *Intern Med J* 2007 Aug;37(8):555-560. [doi: [10.1111/j.1445-5994.2007.01440.x](#)] [Medline: [17640188](#)]
8. Kreyenbuhl J, Nossel IR, Dixon LB. Disengagement from mental health treatment among individuals with schizophrenia and strategies for facilitating connections to care: a review of the literature. *Schizophr Bull* 2009 Jul;35(4):696-703 [FREE Full text] [doi: [10.1093/schbul/sbp046](#)] [Medline: [19491314](#)]
9. Bell LE, Sawyer SM. Transition of care to adult services for pediatric solid-organ transplant recipients. *Pediatr Clin North Am* 2010 Apr;57(2):593-610. [doi: [10.1016/j.pcl.2010.01.007](#)] [Medline: [20371054](#)]
10. Ofcom. Communications market report 2016. London, UK: Ofcom; 2016 Aug 4. URL: https://www.ofcom.org.uk/_data/assets/pdf_file/0024/26826/cmr_uk_2016.pdf [accessed 2017-03-24] [WebCite Cache ID 6oz3lyQWU]
11. Van Wallegghem N, Macdonald CA, Dean HJ. Evaluation of a systems navigator model for transition from pediatric to adult care for young adults with type 1 diabetes. *Diabetes Care* 2008 Aug;31(8):1529-1530 [FREE Full text] [doi: [10.2337/dc07-2247](#)] [Medline: [18458141](#)]
12. Simon J, Budge K, Foster J, Bell S, Christensen J, Price J, et al. Impact of True Colours mood monitoring on mental health service utilisation. 2011 Presented at: 9th International Conference on Bipolar Disorder; June 9-11, 2011; Pittsburgh, PA.
13. Bartlett C, Simpson K, Turner AN. Patient access to complex chronic disease records on the Internet. *BMC Med Inform Decis Mak* 2012 Aug 06;12:87 [FREE Full text] [doi: [10.1186/1472-6947-12-87](#)] [Medline: [22867441](#)]
14. National Information Board. National Information Board's workstream roadmaps. Leeds, UK: National Information Board; 2015 Oct 01. URL: <https://www.gov.uk/government/publications/national-information-boards-workstream-roadmaps> [accessed 2017-03-24] [WebCite Cache ID 6pCXS54NB]
15. Wachter R. Making IT work: harnessing the power of health information technology to improve care in England. London, UK: Department of Health; 2016 Sep 07. URL: <https://www.gov.uk/government/publications/using-information-technology-to-improve-the-nhs/making-it-work-harnessing-the-power-of-health-information-technology-to-improve-care-in-england> [accessed 2017-03-15] [WebCite Cache ID 6oz5LKTUo]
16. NHS Digital. NHSmail. 2016. URL: <https://digital.nhs.uk/nhsmail> [accessed 2016-12-13] [WebCite Cache ID 6mixYKfG8]

17. Ye J, Rust G, Fry-Johnson Y, Strothers H. E-mail in patient-provider communication: a systematic review. *Patient Educ Couns* 2010 Aug;80(2):266-273 [[FREE Full text](#)] [doi: [10.1016/j.pec.2009.09.038](https://doi.org/10.1016/j.pec.2009.09.038)] [Medline: [19914022](https://pubmed.ncbi.nlm.nih.gov/19914022/)]
18. Blackburn S, Brownsell S, Hawley MS. A systematic review of digital interactive television systems and their applications in the health and social care fields. *J Telemed Telecare* 2011;17(4):168-176. [doi: [10.1258/jtt.2010.100610](https://doi.org/10.1258/jtt.2010.100610)] [Medline: [21398387](https://pubmed.ncbi.nlm.nih.gov/21398387/)]
19. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. *Cochrane Database Syst Rev* 2012;12:CD007459. [doi: [10.1002/14651858.CD007459.pub2](https://doi.org/10.1002/14651858.CD007459.pub2)] [Medline: [23235644](https://pubmed.ncbi.nlm.nih.gov/23235644/)]
20. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. *Telemed J E Health* 2012 Apr;18(3):175-184. [doi: [10.1089/tmj.2011.0119](https://doi.org/10.1089/tmj.2011.0119)] [Medline: [22356525](https://pubmed.ncbi.nlm.nih.gov/22356525/)]
21. Johansen MA, Henriksen E, Horsch A, Schuster T, Berntsen GK. Electronic symptom reporting between patient and provider for improved health care service quality: a systematic review of randomized controlled trials. part 1: state of the art. *J Med Internet Res* 2012 Oct 03;14(5):e118 [[FREE Full text](#)] [doi: [10.2196/jmir.2214](https://doi.org/10.2196/jmir.2214)] [Medline: [23032300](https://pubmed.ncbi.nlm.nih.gov/23032300/)]
22. Meyer B, Atherton H, Sawmynaden P, Car J. Email for communicating results of diagnostic medical investigations to patients. *Cochrane Database Syst Rev* 2012 Aug 15(8):CD007980. [doi: [10.1002/14651858.CD007980.pub2](https://doi.org/10.1002/14651858.CD007980.pub2)] [Medline: [22895970](https://pubmed.ncbi.nlm.nih.gov/22895970/)]
23. Atherton H, Sawmynaden P, Sheikh A, Majeed A, Car J. Email for clinical communication between patients/caregivers and healthcare professionals. *Cochrane Database Syst Rev* 2012;11:CD007978. [doi: [10.1002/14651858.CD007978.pub2](https://doi.org/10.1002/14651858.CD007978.pub2)] [Medline: [23152249](https://pubmed.ncbi.nlm.nih.gov/23152249/)]
24. Osborn CY, Mayberry LS, Mulvaney SA, Hess R. Patient web portals to improve diabetes outcomes: a systematic review. *Curr Diab Rep* 2010 Dec;10(6):422-435 [[FREE Full text](#)] [doi: [10.1007/s11892-010-0151-1](https://doi.org/10.1007/s11892-010-0151-1)] [Medline: [20890688](https://pubmed.ncbi.nlm.nih.gov/20890688/)]
25. Sutcliffe P, Martin S, Sturt J, Powell J, Griffiths F, Adams A, et al. Systematic review of communication technologies to promote access and engagement of young people with diabetes into healthcare. *BMC Endocr Disord* 2011;11:1 [[FREE Full text](#)] [doi: [10.1186/1472-6823-11-1](https://doi.org/10.1186/1472-6823-11-1)] [Medline: [21210964](https://pubmed.ncbi.nlm.nih.gov/21210964/)]
26. Verhoeven F, Tanja-Dijkstra K, Nijland N, Eysenbach G, van Gemert-Pijnen GL. Asynchronous and synchronous teleconsultation for diabetes care: a systematic literature review. *J Diabetes Sci Technol* 2010 May;4(3):666-684 [[FREE Full text](#)] [Medline: [20513335](https://pubmed.ncbi.nlm.nih.gov/20513335/)]
27. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for communicating results of medical investigations. *Cochrane Database Syst Rev* 2012;6:CD007456. [doi: [10.1002/14651858.CD007456.pub2](https://doi.org/10.1002/14651858.CD007456.pub2)] [Medline: [22696369](https://pubmed.ncbi.nlm.nih.gov/22696369/)]
28. Van Gaalen JL, Hashimoto S, Sont JK. Telemanagement in asthma: an innovative and effective approach. *Curr Opin Allergy Clin Immunol* 2012 Jun;12(3):235-240. [doi: [10.1097/ACI.0b013e3283533700](https://doi.org/10.1097/ACI.0b013e3283533700)] [Medline: [22475997](https://pubmed.ncbi.nlm.nih.gov/22475997/)]
29. 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](https://doi.org/10.1002/14651858.CD007198.pub2)] [Medline: [23076932](https://pubmed.ncbi.nlm.nih.gov/23076932/)]
30. Gentles SJ, Lokker C, McKibbin KA. Health information technology to facilitate communication involving health care providers, caregivers, and pediatric patients: a scoping review. *J Med Internet Res* 2010;12(2):e22 [[FREE Full text](#)] [doi: [10.2196/jmir.1390](https://doi.org/10.2196/jmir.1390)] [Medline: [20562092](https://pubmed.ncbi.nlm.nih.gov/20562092/)]
31. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. *Pain* 2011 Aug;152(8):1740-1750. [doi: [10.1016/j.pain.2011.02.012](https://doi.org/10.1016/j.pain.2011.02.012)] [Medline: [21565446](https://pubmed.ncbi.nlm.nih.gov/21565446/)]
32. Martin S, Sutcliffe P, Griffiths F, Sturt J, Powell J, Adams A, et al. Effectiveness and impact of networked communication interventions in young people with mental health conditions: a systematic review. *Patient Educ Couns* 2011 Nov;85(2):e108-e119. [doi: [10.1016/j.pec.2010.11.014](https://doi.org/10.1016/j.pec.2010.11.014)] [Medline: [21239133](https://pubmed.ncbi.nlm.nih.gov/21239133/)]
33. McLean S, Protti D, Sheikh A. Telehealthcare for long term conditions. *BMJ* 2011;342:d120. [Medline: [21292710](https://pubmed.ncbi.nlm.nih.gov/21292710/)]
34. Johansen MA, Berntsen GK, Schuster T, Henriksen E, Horsch A. Electronic symptom reporting between patient and provider for improved health care service quality: a systematic review of randomized controlled trials. part 2: methodological quality and effects. *J Med Internet Res* 2012 Oct 03;14(5):e126 [[FREE Full text](#)] [doi: [10.2196/jmir.2216](https://doi.org/10.2196/jmir.2216)] [Medline: [23032363](https://pubmed.ncbi.nlm.nih.gov/23032363/)]
35. Gagnon M, Légaré F, Labrecque M, Frémont P, Pluye P, Gagnon J, et al. Interventions for promoting information and communication technologies adoption in healthcare professionals. *Cochrane Database Syst Rev* 2009 Jan 21(1):CD006093 [[FREE Full text](#)] [doi: [10.1002/14651858.CD006093.pub2](https://doi.org/10.1002/14651858.CD006093.pub2)] [Medline: [19160265](https://pubmed.ncbi.nlm.nih.gov/19160265/)]
36. Yin R. *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications; 2013.
37. Pawson R, Tilley N. *Realistic Evaluation*. London, UK: Sage; 1997.
38. Boyatzis R. *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: Sage Publications; 1998.
39. Ives J, Draper H. Appropriate methodologies for empirical bioethics: it's all relative. *Bioethics* 2009 May;23(4):249-258. [doi: [10.1111/j.1467-8519.2009.01715.x](https://doi.org/10.1111/j.1467-8519.2009.01715.x)] [Medline: [19338525](https://pubmed.ncbi.nlm.nih.gov/19338525/)]
40. International Electrotechnical Commission. IEC 60812: Analysis Techniques for System Reliability-Procedure for Failure Mode and Effects Analysis (FMEA). Geneva, Switzerland: International Electrotechnical Commission; 2006:1-93.

41. Royal College of Nursing. Agenda for Change pay scales for 2014-2015. London, UK: Royal College of Nursing; 2014. URL: https://www2.rcn.org.uk/_data/assets/pdf_file/0007/572812/nurse-pay-scales-nhs-agenda-for-change-england-2014-2015.pdf [accessed 2017-03-24] [WebCite Cache ID 6pCYuUqFm]
42. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013 Apr 4. URL: <https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781> [accessed 2017-03-15] [WebCite Cache ID 6oz6EtQEK]
43. LYNCS Project. LYNCS Young Researchers Symposium: the films. 2011 Jul 7. URL: <https://www.youtube.com/watch?v=enngoM2BoFU> [accessed 2017-03-15] [WebCite Cache ID 6oz6PbHtW]
44. Warwick Medical School, University of Warwick. LYNCS study quick reference guides. 2016 Oct 5. URL: <http://www2.warwick.ac.uk/fac/med/research/hscience/sssh/research/lynscs/outputs> [accessed 2017-03-24] [WebCite Cache ID 6mixxr5OR]
45. Slowther A. The concept of autonomy and its interpretation in health care. *Clin Ethics* 2007;2(4):173-175.
46. Lee Y, Lin JL. Do patient autonomy preferences matter? Linking patient-centered care to patient-physician relationships and health outcomes. *Soc Sci Med* 2010 Nov;71(10):1811-1818. [doi: [10.1016/j.socscimed.2010.08.008](https://doi.org/10.1016/j.socscimed.2010.08.008)] [Medline: [20933316](https://pubmed.ncbi.nlm.nih.gov/20933316/)]
47. General Medical Council. Duties of a doctor: the duties of a doctor registered with the General Medical Council. 2017. URL: http://www.gmc-uk.org/guidance/good_medical_practice/duties_of_a_doctor.asp [accessed 2017-03-15] [WebCite Cache ID 6oz6oHjGo]
48. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. *J Am Med Inform Assoc* 2016 Jan;23(1):212-220. [doi: [10.1093/jamia/ocv052](https://doi.org/10.1093/jamia/ocv052)] [Medline: [26078410](https://pubmed.ncbi.nlm.nih.gov/26078410/)]
49. O' Cathain A, Drabble SJ, Foster A, Horspool K, Edwards L, Thomas C, et al. Being human: a qualitative interview study exploring why a telehealth intervention for management of chronic conditions had a modest effect. *J Med Internet Res* 2016 Jun 30;18(6):e163 [FREE Full text] [doi: [10.2196/jmir.5879](https://doi.org/10.2196/jmir.5879)] [Medline: [27363434](https://pubmed.ncbi.nlm.nih.gov/27363434/)]
50. Greenhalgh T, Procter R, Wherton J, Sugarhood P, Hinder S, Rouncefield M. What is quality in assisted living technology? The ARCHIE framework for effective telehealth and telecare services. *BMC Med* 2015;13:91 [FREE Full text] [doi: [10.1186/s12916-015-0279-6](https://doi.org/10.1186/s12916-015-0279-6)] [Medline: [25902803](https://pubmed.ncbi.nlm.nih.gov/25902803/)]
51. Herbert L, Owen V, Pascarella L, Streisand R. Text message interventions for children and adolescents with type 1 diabetes: a systematic review. *Diabetes Technol Ther* 2013 May;15(5):362-370. [doi: [10.1089/dia.2012.0291](https://doi.org/10.1089/dia.2012.0291)] [Medline: [23550554](https://pubmed.ncbi.nlm.nih.gov/23550554/)]
52. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2013;3:CD008776. [doi: [10.1002/14651858.CD008776.pub2](https://doi.org/10.1002/14651858.CD008776.pub2)] [Medline: [23543567](https://pubmed.ncbi.nlm.nih.gov/23543567/)]
53. Guljas R, Ahmed A, Chang K, Whitlock A. Impact of telemedicine in managing type 1 diabetes among school-age children and adolescents: an integrative review. *J Pediatr Nurs* 2014;29(3):198-204. [doi: [10.1016/j.pedn.2013.10.013](https://doi.org/10.1016/j.pedn.2013.10.013)] [Medline: [24269308](https://pubmed.ncbi.nlm.nih.gov/24269308/)]
54. Ye X, Bapuji SB, Winters SE, Struthers A, Raynard M, Metge C, et al. Effectiveness of internet-based interventions for children, youth, and young adults with anxiety and/or depression: a systematic review and meta-analysis. *BMC Health Serv Res* 2014 Jul 18;14:313 [FREE Full text] [doi: [10.1186/1472-6963-14-313](https://doi.org/10.1186/1472-6963-14-313)] [Medline: [25037951](https://pubmed.ncbi.nlm.nih.gov/25037951/)]
55. Aardoom JJ, Dingemans AE, Spinhoven P, Van Furth EF. Treating eating disorders over the internet: a systematic review and future research directions. *Int J Eat Disord* 2013 Sep;46(6):539-552. [doi: [10.1002/eat.22135](https://doi.org/10.1002/eat.22135)] [Medline: [23674367](https://pubmed.ncbi.nlm.nih.gov/23674367/)]
56. Chi N, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. *J Telemed Telecare* 2015 Jan;21(1):37-44 [FREE Full text] [doi: [10.1177/1357633X14562734](https://doi.org/10.1177/1357633X14562734)] [Medline: [25475220](https://pubmed.ncbi.nlm.nih.gov/25475220/)]
57. Sujan M, Spurgeon P, Cooke M, Weale A, Debenham P, Cross S. The development of safety cases for healthcare services: practical experiences, opportunities and challenges. *Reliability Eng Syst Safety* 2015;140:200-207.

Abbreviations

NHS: National Health Service

NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 21.12.16; peer-reviewed by J Oldenburg, A Farmer, S Mclean, S Hinder, MP Gagnon; comments to author 13.01.17; revised version received 24.01.17; accepted 24.01.17; published 10.04.17.

Please cite as:

Griffiths F, Bryce C, Cave J, Dritsaki M, Fraser J, Hamilton K, Huxley C, Ignatowicz A, Kim SW, Kimani PK, Madan J, Slowther AM, Sujan M, Sturt J

Timely Digital Patient-Clinician Communication in Specialist Clinical Services for Young People: A Mixed-Methods Study (The LYNC Study)

J Med Internet Res 2017;19(4):e102

URL: <http://www.jmir.org/2017/4/e102/>

doi:[10.2196/jmir.7154](https://doi.org/10.2196/jmir.7154)

PMID:[28396301](https://pubmed.ncbi.nlm.nih.gov/28396301/)

©Frances Griffiths, Carol Bryce, Jonathan Cave, Melina Dritsaki, Joseph Fraser, Kathryn Hamilton, Caroline Huxley, Agnieszka Ignatowicz, Sung Wook Kim, Peter K Kimani, Jason Madan, Anne-Marie Slowther, Mark Sujan, Jackie Sturt. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 10.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Is Provider Secure Messaging Associated With Patient Messaging Behavior? Evidence From the US Army

Vickee Wolcott^{1*}, PhD; Ritu Agarwal^{2*}, PhD; D. Alan Nelson^{3*}, PhD

¹Army-Baylor University Graduate Program in Health and Business Administration, JBSA Ft Sam Houston, TX, United States

²Center for Health Information and Decision Systems, Robert H Smith School of Business, University of Maryland, College Park, MD, United States

³School of Medicine, Stanford University, Stanford, CA, United States

* all authors contributed equally

Corresponding Author:

Vickee Wolcott, PhD

Army-Baylor University Graduate Program in Health and Business Administration

3599 Winfield Scott Rd

Suite 1224

JBSA Ft Sam Houston, TX, 78234

United States

Phone: 1 210 701 0415

Fax: 1 210 466 4728

Email: vickee_wolcott@baylor.edu

Abstract

Background: Secure messaging with health care providers offers the promise of improved patient-provider relationships, potentially facilitating outcome improvements. But, will patients use messaging technology in the manner envisioned by policy-makers if their providers do not actively use it?

Objective: We hypothesized that the level and type of secure messaging usage by providers might be associated with messaging initiation by their patients.

Methods: The study employed a dataset of health care and secure messaging records of more than 81,000 US Army soldiers and nearly 3000 clinicians with access to a patient portal system. We used a negative binomial regression model on over 25 million observations to determine the adjusted association between provider-initiated and provider-response messaging and subsequent messaging by their patients in this population over a 4-year period.

Results: Prior provider-initiated and response messaging levels were associated with new patient messaging when controlling for the patient's health care utilization and diagnoses, with the strongest association for high provider-response messaging level. Patients whose providers were highly responsive to the messages of other patients initiated 334% more secure messages ($P < .001$) than patients with providers who did not personally respond to other patients' messages.

Conclusions: Our results indicate that provider messaging usage levels and types thereof predict their patients' subsequent communication behavior. The findings suggest the need for more study into the factors associated with provider messaging to fully understand the mechanisms of this relationship.

(*J Med Internet Res* 2017;19(4):e103) doi:[10.2196/jmir.6804](https://doi.org/10.2196/jmir.6804)

KEYWORDS

patient portal; physician-patient relations; health communication

Introduction

Background

Better patient-provider communication is important because their relationship is at the center of health care service delivery [1]. A supportive patient-provider relationship has been shown to be associated with many positive outcomes, including

increased patient compliance [2], decreased pain [3], and shortened recovery periods [4]. Opportunities to interact with one another are critical to the development of ideal patient-provider relationships [5]. In "Crossing the Quality Chasm," the Institute of Medicine (2001) [6] recommended the use of phone and email communication between appointments

as a visit extender to support a continuous patient-provider relationship [5,7].

Ongoing and expanded communication stands in contrast to the historical, episodic patient-provider relationship that mainly comprises infrequent office visits. Secure messaging could facilitate the development of deeper relationships by increasing interaction time, making patients more comfortable about asking questions and discussing embarrassing issues [8], and allowing physicians to provide better advice and education [9]. However, such benefits are likely to be realized only if patients and providers are both committed users of the technology.

Secure messaging is often provided as part of a patient portal. Unfortunately, studies show that portal use among patients is low, with only 10-32% of patient portal adopters actually using the portal [10,11]. The Centers for Medicare & Medicaid Services recently proposed a change to the Health Information Technology for Economic and Clinical Health Act because providers appear to struggle to engage their patients via electronic means. This recommendation would reduce the current requirement that 5% of patients use secure messaging to simply requiring the presence of the feature [12].

Reducing the communication recommendation may be the wrong approach, given that relatively little evidence has been developed in this area. Researchers have paid surprisingly limited attention to provider and patient usage of secure messaging, and the associated factors are potentially complex. Patients perceive the overall health service quality they receive and develop an associated level of trust and comfort with their clinicians [13], potentially driving behavior such as secure messaging. Patients may lose interest in such resources if providers do not encourage the use of electronic tools or lead by example by becoming active users of the tools themselves [14]. Furthermore, the way and degree to which providers generally engage in messaging may represent a marker of the level of approachability that is perceived by their patients during care. Provider receptiveness to communication, an otherwise difficult-to-assess factor, might be indicated by the willingness of patients to initiate secure messages with their clinicians.

Aim of This Study

Our associated hypothesis was that providers' overall messaging behavior might serve as an indicator of their accessibility for or interest in communication, which patients directly or indirectly perceive. Therefore, patients' use of secure messaging might be related to the extent to which their providers generally use it. Our resulting specific aim was to answer the question: What is the relationship between providers' past secure messaging types and levels and the initiation of messaging by patients?

We defined provider messaging levels as the extent to which a patient's primary care provider exchanges messages with his or her other patients, compared with the messaging rates of other providers in the same population. We further distinguished between messages the provider initiated on their own (defined as "provider-initiated messages") and messages the provider sent in response to patient messages (defined as

"provider-response messages") due to the potentially different causes for these messaging events.

We theorized that provider-initiated messages might be mostly representative of routine operational matters, rather than a personal philosophy or stance on communication that patients could perceive and act upon. For example, such messages could occur due to specific clinical needs including notifications of laboratory or imaging results or reminders. In contrast, choosing to personally respond to patient messages (rather than delegating this function to the supporting clinical team, or even ignoring messages) might represent a stronger commitment to communication. This receptive mindset could be perceived by other patients, creating increased comfort among them for message initiation. Because we had no data on the specific content of messages, controlling for the distinct message types provided an initial method for isolating any difference in impact between them.

Methods

Data

We employed messaging data from the Army Medicine Secure Messaging Service (AMSMS) used by the US Army Medical Department (AMEDD), introduced in January 2011. Patients in this health system use AMSMS to securely message their primary care and medical teams to request medical advice, appointments, lab results, referrals, and prescription renewals; record medical information; and access educational materials. Providers initiate messages in AMSMS to send care reminders, appointment reminders, and direct patient messages. AMSMS was rolled out in a consistent manner across Army hospitals and clinics, with a team visiting each location to conduct training and provide system access.

We theorized that AMEDD would be an excellent setting for this study because it is a large, integrated organization providing health care to the Army's 3.95 million service members, retirees, and family members. We were able to utilize a rich, extensive dataset with over 25 million observations constituting eligible months for health care within the Army's medical system. The system comprises 8 medical centers, 27 community hospitals, and 180 primary care clinics [15], with common policies and procedures and similar patient populations across medical facilities.

The primary dataset therefore consisted of de-identified administrative, medical, and training data from official military information systems documenting AMEDD care. This repository was established at the University of Maryland Center for Health Information and Decision Systems (CHIDS) as the Military Medical Informatics Data Set (MMIDS). MMIDS contains data on over 820,000 active duty soldiers in total, capturing military service and associated events during January 2011 through December 2014. The data are arranged into a longitudinal record of observed person-months of military service during this time, across which values for the selected variables were free to vary with time. Data elements in the dataset include, among other variables, age, deployment history, time-in-service, rank, race, marital status, body mass index, self-reported health measures,

medical diagnoses, medical appointment data, prescription medications, physical fitness test scores, and tobacco use.

We obtained AMSMS usage logs from its implementation in January 2011 through November 2014, and linked the patients in the MMIDS who were portal users with their specific messaging actions. The AMSMS data included 727,951 secure messages for 439,368 patient users, of which 81,645 were active duty soldiers, involving 2983 provider and staff users. We studied active duty soldiers only because MMIDS solely consists of information on these soldiers.

This study was reviewed and determined to be exempt by the University of Maryland Institutional Review Board and underwent secondary review at the Defense Health Agency's Human Research Protection Office. All statistical analysis was conducted using Stata 13 software (StataCorp).

Variables

Our dependent variable represented the number of messages, if any, sent by each patient registered to use the AMSMS in each observed person-month. Values for this parameter varied with time within each such month across the longitudinal dataset, where applicable. Our requirement was that qualifying messages were initiated by the patient and did not represent a reply to providers. To control for each primary care provider's overall AMSMS message rate, we calculated the number of messages which a patient's provider initiated and responded to other patients in each month (excluding the focal patient). We then divided this value by the number of patients enrolled in

AMSMS for that provider. The quotients were then categorized by tertiles into low, medium, and high messaging when compared with the messaging values of all providers in the sample. We note that these measures were exogenous to the patient, and therefore expected to be uncorrelated with individual outcomes. Our approach appeared adequate to control for provider workload because the providers in our sample worked full-time and had approximately equal patient empanelments in accordance with Army policy.

We expected that patients might have been more likely to send secure messages following health care visits and in response to ongoing medical issues. The patient-specific number of health care visits could have been associated with problem severity and chronicity, which in turn could have been associated with an individual increased need for messaging. Therefore, we included health care utilization measures and medical conditions within the previous 3 months as independent variables to eliminate variance explained by these factors.

The International Classification of Disease System, 9th Revision, Clinical Modification (ICD-9) remained in use by the Military Health System at the time of the studied events. We included medical conditions in the following categories defined by ICD-9 codes because we observed these to be the five most prevalent condition types among the active duty Army soldiers in our dataset: musculoskeletal issues, mental health diagnoses, hypertension, sleep apnea, and dyslipidemia. Additionally, we controlled for calendar month and location. [Table 1](#) describes each of the variables.

Table 1. Description of variables.

Variables	Variable name	Description
Independent variables		
Recently diagnosed patient medical conditions		
	mentaldx	Whether or not the soldier had a diagnosis of anxiety disorder, adjustment disorder, personality disorder, depression, or post-traumatic stress disorder within the previous 3 months
	mskdx	Whether or not the soldier had a diagnosis of musculoskeletal issue (eg, back injury, joint pain) within the previous 3 months
	sleepapndx	Whether or not the soldier had a diagnosis of sleep apnea within the previous 3 months
	hypertensiondx	Whether or not the soldier had a diagnosis of hypertension within the previous 3 months
	dyslipidemiadx	Whether or not the soldier had a diagnosis of dyslipidemia within the previous 3 months
Health care utilization measures		
	primcaretot	Number of monthly primary care visits
	ervisit	Number of monthly emergency room visits
	speccaretot	Number of monthly specialty care visits
Secure messaging factors		
	provinitiatecat	The number of messages the focal patient's provider initiated to other patients in a month, representing the providers' messaging level, categorized by tertiles. Message types included are care reminders, appointment reminders, and patient communication. Three categories: low, medium, and high
	provresponsecat	The number of messages in which the focal patient's provider responded to other patients in a month, representing the providers' messaging level and categorized by tertiles. Message types included responses to appointment requests, billing questions, lab or test results, doctor notes, referral requests, and prescription refills. Three categories: low, medium, and high
Dependent variable		
	patientmsg	The number of messages initiated by each patient in each observed person-month (excludes replies to provider messages)
Other factors		
	installation	Patient's site of military service, one of 32 possible locations
	month	Monthly dummies for time controls

Analysis

Because the outcome measure was a count variable and was overdispersed, we utilized a negative binomial regression model [16]. We included patient-level fixed effects to control for patient-level heterogeneity that could impact portal usage. We note that the fixed effect is able to account for differences in patient characteristics such as demographics as well as technology acceptance factors idiosyncratic to a patient, such as perceived usefulness of the technology, perceived incompatibility with needs, and so on [17]. Patient-level fixed effects allow for a separate intercept for each patient, controlling for unobserved differences among individuals [16]. The provider messaging categories and health care utilization measures were lagged to the previous month to ensure they occurred before the patient sending the message. As data were reported at the

monthly level, it was not possible to distinguish order of events within a month. The regression model was as follows:

$$\log(\text{patientmsg}_{it}) = \beta_{0i} + \beta_1 \text{provinitiatecat}_{it-1} + \beta_2 \text{provresponsecat}_{it-1} + \beta_3 \text{primcaretot}_{it-1} + \beta_4 \text{ervisit}_{it-1} + \beta_5 \text{speccaretot}_{it-1} + \beta_6 [\text{Patient Medical Conditions}]_{it-1} + \beta_7 \text{installation} + \beta_8 \text{month} + \varepsilon_{it}$$

Results

Descriptive Statistics

Tables 2 and 3 provide descriptive statistics of the 81,645 patients who adopted the portal between January 2011 and November 2014. Each month, 7% of patients initiated a secure message. Health care providers initiated on average 0.007 (SD 0.06) messages per patient per month, and responded on average to 0.09 (SD 0.19) messages per patient per month.

Table 2. Patient messaging and health care utilization characteristics.

Description	Mean (SD ^a)
Number of patient-initiated messages per month per patient	0.07 (0.38)
Patient primary care visits	0.48 (0.87)
Patient specialty care visits	0.23 (1.39)
Patient emergency room visits	0.01 (0.14)

^aSD: standard deviation.

Table 3. Characteristics of patients who adopted the portal (N=81,645).

Characteristics	Portal adopter ^a , n (%)
Gender	
Male	64,206 (78.64)
Female	17,439 (21.36)
Age category (in years)	
18-22	13,496 (16.53)
23-27	17,652 (21.62)
28-35	23,187 (28.40)
36+	27,310 (33.45)
Education level	
High school equivalency	4139 (5.07)
High school diploma	39,018 (47.79)
Some college	13,733 (16.82)
Bachelor's degree	14,957 (18.32)
Graduate	9798 (12.00)
Marital status	
Never married	21,367 (26.17)
Married	54,253 (66.45)
Divorced	6025 (7.38)
Had a dyslipidemia diagnosis within the previous 3 months (in 3 months before adoption)	2792 (3.42)
Had a hypertension diagnosis within the previous 3 months	727 (0.89)
Had a mental health diagnosis within the previous 3 months	1755 (2.15)
Had a musculoskeletal diagnosis within the previous 3 months	6572 (8.05)
Had a sleep apnea diagnosis within the previous 3 months	1886 (2.31)

^aAdopters were all patients who signed up for the portal. Not all of them were actual users.

Main Results

As displayed in [Table 4](#) and [Figure 1](#), we found that patients receiving care from high response- and high initiation-messaging providers were substantially more likely to initiate a secure message than patients with nonmessaging providers, that is, those providers who did not send a message in the previous month. Patients with high initiation-messaging providers were 60% more likely to send a secure message than patients with noninitiation-messaging providers. Strikingly, patients with high response-messaging providers sent 334% more messages than those with nonresponse-messaging providers. Patient

message initiation among low response-messaging providers was 254% higher than among nonresponse-messaging providers. Among medium response-messaging providers, patients demonstrated increased messaging by 167% when compared with nonresponse-messaging providers.

As might be expected, health care utilization and medical conditions also impacted patient messaging. For every additional primary care visit during the month prior, patients sent 14% more messages in a given observed month. Specialty care and emergency room visits in the month prior were not associated with the number of messages a patient sent. Having a musculoskeletal or dyslipidemia diagnosis in the previous 3

months were each associated with statistically significant increases in the number of patient messages by 14% and 13%, respectively. But, mental health, hypertension, and sleep apnea were not associated with patient messaging habits. For comparison, we conducted regression analysis without including the five medical conditions. Results were largely unchanged, with very modest differences. The incidence rate ratios (IRRs)

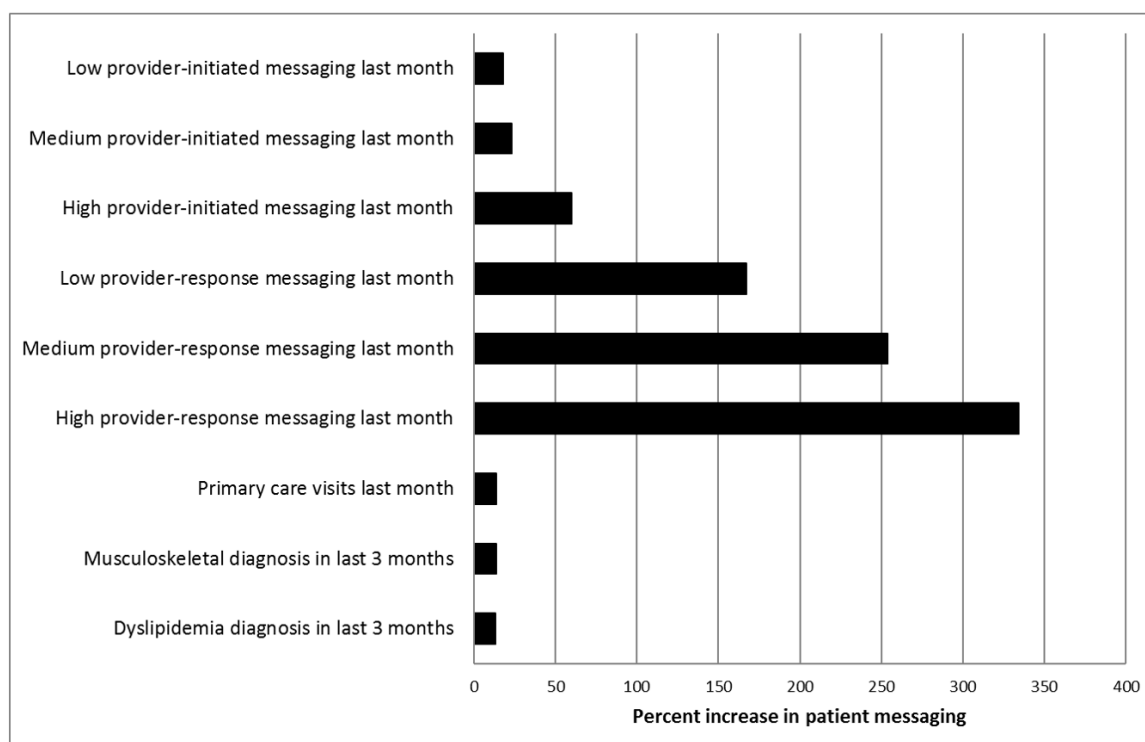
for high provider-initiated messaging and medium and high provider-response messaging increased by 0.01, and the specialty care visits last month became statistically significant when excluding medical conditions, but only provided an IRR of 1.01. It appeared that when only controlling for specialty care visits, the model failed to adequately address the more specific information provided by selected medical conditions.

Table 4. Regression results. The model included time controls, location controls, and patient fixed effects.

Variables	Patient-initiated messages		
	IRR ^a	95% CI	<i>P</i> value
Monthly provider-initiated messaging			
No provider-initiated messaging last month	1.00	Referent	
Low provider-initiated messaging last month	1.18	1.12-1.24	< .001 ^b
Medium provider-initiated messaging last month	1.23	1.17-1.30	< .001
High provider-initiated messaging last month	1.60	1.51-1.70	< .001
Monthly provider-response messaging			
No provider-response messaging last month	1.00	Referent	
Low provider-response messaging last month	2.67	2.53-2.82	< .001
Medium provider-response messaging last month	3.54	3.37-3.73	< .001
High provider-response messaging last month	4.34	4.13-4.55	< .001
Primary care visits last month	1.14	1.12-1.15	< .001
Emergency room visits last month	1.02	0.97-1.07	.49
Specialty care visits last month	1.00	0.99-1.01	.21
Musculoskeletal diagnosis within the previous 3 months	1.14	1.10-1.19	< .001
Mental health diagnosis within the previous 3 months	1.01	0.95-1.07	.76
Hypertension diagnosis within the previous 3 months	1.00	0.92-1.07	.92
Sleep apnea diagnosis within the previous 3 months	1.01	0.96-1.07	.64
Dyslipidemia diagnosis within the previous 3 months	1.13	1.08-1.19	< .001

^aIRR: incidence rate ratio.

^b*P* value in italics indicate statistical significance ($P < .05$).

Figure 1. Graph of relative associations of patients' provider messaging and health care factors with patient messaging.

Discussion

Principal Findings

In this study, we demonstrated that among US Army soldiers, increased provider-initiated and provider-response messaging were associated with statistically-significant increases in the adjusted probability of patient-initiated secure messaging. We also demonstrated that provider-response messaging had a much larger impact on patient messaging than provider-initiated messaging. Given our ability to control for health care utilization and medical conditions, the study offers strong evidence that provider messaging usage is a critical, overlooked factor associated with their patients' behavior.

We can suggest no direct mechanism by which provider messaging with a given patient would impact the decision of another patient to initiate a message. We therefore theorize that patients' willingness to initiate messages may stem from their appreciation of provider traits that, in turn, are associated with the provider's propensity to robustly engage in different types of secure messaging. These traits might be more evident among clinicians who choose to personally respond to patient messages and do so at high rates. Provider-response messaging may be more influential than provider-initiated messaging because provider-response messaging is likely a more personal type of messaging, tailored to each specific patient's needs.

Patients may appreciate the general communicative nature of providers who take the time to respond to messages personally, rather than having a staff member respond. Alternatively, it is possible that patients may initiate messages due to frustration

with their ability to communicate with providers because their providers use secure messaging to avoid face-to-face and telephone encounters. Provider messaging use therefore requires substantial further study in order to better understand how providers differ when stratified by secure messaging usage levels and types.

Our findings carry implications for policy addressing the wider diffusion and uptake of critical patient-centered health information technologies. Health information technologies have been heralded as one possible solution to addressing the high cost and often low quality of health services delivery. However, as noted, authorities may reduce the pressure on providers to engage in patient portal and secure messaging use. Perhaps a more useful approach would be to require that providers demonstrate a minimum level of engagement with secure messaging and to sponsor studies that examine the factors associated with all use levels and types.

Additionally, our findings revealed that medical problem types were related to patient message initiation in the studied population. Health care organizations may therefore expect increased secure messaging from patients following primary care visits and recent diagnoses of certain conditions, such as musculoskeletal issues and dyslipidemia. Organizations that are at early stages of implementing secure messaging resources might need to ensure that health care teams and patients in care settings in which these problems predominate are notified of the potential for high messaging rates. These settings might also provide the greatest opportunities for the study of secure messaging and the emergence of associated best practices.

Limitation

A limitation of this study is that the findings may not be broadly generalizable because we studied a younger, more male population than that seen in the general public, and this population is preselected for health as a requirement of military service.

Future Work

Further study in other groups will be required to assess the external validity of our findings. However, we note that the patient portal software, including secure messaging capability used by the military, is the same software used in many civilian health care settings. The ability to verify or refute our findings in large civilian medical systems should therefore be feasible in future research employing data from this or similar systems.

Conclusions

This was the first study to use a large, robust dataset to empirically investigate provider messaging behavior and its potential relationship with the willingness of patients to send secure messages. New data will be needed to address the

potential, unobservable factors that explain our main finding of a provider-patient usage association. Candidate data might include new surveys assessing patient perceptions, supporting a study comparing those cared for by clinicians with varying secure messaging use levels and types. In preparation for such research, we will leverage the datasets employed for this study that include large reserves of additional information on patient trajectories and provider behavior. We expect these data to provide new insights with our ongoing research to better understand the impact of patient and provider utilization of technologies and the factors associated with these critical contributors to health.

The project also revealed relatively high outpatient health care utilization (Table 2) for a generally young population of individuals who were preselected for health in order to serve. As the military is an environment with universal, free health care, this finding suggests the potential to study care utilization behavior in such an environment. We will assess this study concept for feasibility as part of our ongoing review of the substantial data resources available to the research team.

Acknowledgments

We thank Dr Guodong (Gordon) Gao, University of Maryland, College Park, for his insightful ideas and suggestions for improvement. This research was supported by the University of Maryland-Leidos seed grant program. The grant program sponsors had no role in the review or approval of the manuscript. The views expressed in this paper are those of the authors and do not necessarily reflect the views or official policies of the US Government, Department of Defense, Defense Health Agency, Department of the Army, or the US Army Medical Department.

Conflicts of Interest

None declared.

References

1. Alexander JA, Hearld LR, Mittler JN, Harvey J. Patient-physician role relationships and patient activation among individuals with chronic illness. *Health Serv Res* 2012 Jun;47(3 Pt 1):1201-1223 [FREE Full text] [doi: [10.1111/j.1475-6773.2011.01354.x](https://doi.org/10.1111/j.1475-6773.2011.01354.x)] [Medline: [22098418](https://pubmed.ncbi.nlm.nih.gov/22098418/)]
2. Francis V, Korsch BM, Morris MJ. Gaps in doctor-patient communication: patients' response to medical advice. *N Engl J Med* 1969 Mar 6;280(10):535-540. [doi: [10.1056/NEJM196903062801004](https://doi.org/10.1056/NEJM196903062801004)] [Medline: [5764453](https://pubmed.ncbi.nlm.nih.gov/5764453/)]
3. Gryll SL, Katahn M. Situational factors contributing to the placebo effect. *Psychopharmacology (Berl)* 1978 May 31;57(3):253-261. [Medline: [97705](https://pubmed.ncbi.nlm.nih.gov/97705/)]
4. Olsson B, Olsson B, Tibblin G. Effect of patients' expectations on recovery from acute tonsillitis. *Fam Pract* 1989 Sep;6(3):188-192. [Medline: [2792618](https://pubmed.ncbi.nlm.nih.gov/2792618/)]
5. Baur C. Limiting factors on the transformative powers of e-mail in patient-physician relationships: a critical analysis. *Health Commun* 2000;12(3):239-259. [doi: [10.1207/S15327027HC1203_02](https://doi.org/10.1207/S15327027HC1203_02)] [Medline: [10938915](https://pubmed.ncbi.nlm.nih.gov/10938915/)]
6. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press; 2001.
7. Matusitz J, Breen G. Telemedicine: its effects on health communication. *Health Commun* 2007;21(1):73-83. [doi: [10.1080/10410230701283439](https://doi.org/10.1080/10410230701283439)] [Medline: [17461754](https://pubmed.ncbi.nlm.nih.gov/17461754/)]
8. Haun JN, Patel NR, Lind JD, Antinori N. Large-scale survey findings inform patients' experiences in using secure messaging to engage in patient-provider communication and self-care management: a quantitative assessment. *J Med Internet Res* 2015;17(12):e282 [FREE Full text] [doi: [10.2196/jmir.5152](https://doi.org/10.2196/jmir.5152)] [Medline: [26690761](https://pubmed.ncbi.nlm.nih.gov/26690761/)]
9. Houston TK, Sands DZ, Jenckes MW, Ford DE. Experiences of patients who were early adopters of electronic communication with their physician: satisfaction, benefits, and concerns. *Am J Manag Care* 2004 Sep;10(9):601-608 [FREE Full text] [Medline: [15515992](https://pubmed.ncbi.nlm.nih.gov/15515992/)]
10. 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.cjcd.2013.10.005](https://doi.org/10.1016/j.cjcd.2013.10.005)] [Medline: [24485208](https://pubmed.ncbi.nlm.nih.gov/24485208/)]

11. Weppner WG, Ralston JD, Koepsell TD, Grothaus LC, Reid RJ, Jordan L, et al. Use of a shared medical record with secure messaging by older patients with diabetes. *Diabetes Care* 2010 Nov;33(11):2314-2319 [[FREE Full text](#)] [doi: [10.2337/dc10-1124](#)] [Medline: [20739686](#)]
12. Centers for Medicare and Medicaid Services. Federalregister. 2014. Medicare and Medicaid Programs: Electronic Health Record Incentive Program-Modifications to Meaningful Use in 2015 Through 2017; Proposed Rule URL: <https://www.federalregister.gov/documents/2015/10/16/2015-25595/medicare-and-medicicaid-programs-electronic-health-record-incentive-program-stage-3-and-modifications> [accessed 2017-01-16] [WebCite Cache ID 615VPad9z]
13. Chang C, Chen S, Lan Y. Service quality, trust, and patient satisfaction in interpersonal-based medical service encounters. *BMC Health Serv Res* 2013 Jan 16;13:22 [[FREE Full text](#)] [doi: [10.1186/1472-6963-13-22](#)] [Medline: [23320786](#)]
14. Wells S, Rozenblum R, Park A, Dunn M, Bates DW. Organizational strategies for promoting patient and provider uptake of personal health records. *J Am Med Inform Assoc* 2015 Jan;22(1):213-222 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2014-003055](#)] [Medline: [25326601](#)]
15. Hudak RP, Julian R, Kugler J, Dorrance K, Lynch S, Dinneen M, et al. The patient-centered medical home: a case study in transforming the military health system. *Mil Med* 2013 Feb;178(2):146-152. [Medline: [23495459](#)]
16. Trivedi PK, Cameron AC. Microeconometrics: methods and applications. Cambridge: Cambridge University Press; 2005.
17. Or CK, Karsh B. A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc* 2009 Aug;16(4):550-560 [[FREE Full text](#)] [doi: [10.1197/jamia.M2888](#)] [Medline: [19390112](#)]

Abbreviations

AMEDD: US Army Medical Department
AMSMS: Army Medicine Secure Messaging Service
CHIDS: Center for Health Information and Decision Systems
IRR: incidence rate ratio
MMIDS: Military Medical Informatics Data Set

Edited by G Eysenbach; submitted 13.10.16; peer-reviewed by K Nazi, S Simon, S Shimada; comments to author 22.12.16; revised version received 01.02.17; accepted 04.03.17; published 06.04.17.

Please cite as:

Wolcott V, Agarwal R, Nelson DA
Is Provider Secure Messaging Associated With Patient Messaging Behavior? Evidence From the US Army
J Med Internet Res 2017;19(4):e103
URL: <http://www.jmir.org/2017/4/e103/>
doi: [10.2196/jmir.6804](#)
PMID: [28385681](#)

©Vickey Wolcott, Ritu Agarwal, D. Alan Nelson. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 06.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Connecting Families to Their Health Record and Care Team: The Use, Utility, and Impact of a Client/Family Health Portal at a Children's Rehabilitation Hospital

Gillian King^{1,2}, PhD; Joanne Maxwell^{2,3}, MSc; Amir Karmali³; Simon Hagens⁴; Madhu Pinto¹, MASP; Laura Williams^{3,5}, MSW; Keith Adamson^{3,5}, PhD

¹Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada

²Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada

³Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada

⁴Canada Health Infoway, Toronto, ON, Canada

⁵Faculty of Social Work, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Gillian King, PhD

Bloorview Research Institute

Holland Bloorview Kids Rehabilitation Hospital

150 Kilgour Rd

Toronto, ON, M4G 1R8

Canada

Phone: 1 416 425 6220 ext 3323

Fax: 1 416 425 1634

Email: gking27@uwo.ca

Abstract

Background: Health care portals have the potential to provide consumers with timely, transparent access to health care information and engage them in the care process.

Objective: The objective was to examine the use, utility, and impact on engagement in care and caregiver-provider communication of a client/family portal providing access to electronic health records (EHRs) and secure, 2-way e-messaging with care providers.

Methods: We conducted a prospective, mixed-methods study involving collection of caregivers' portal usage information over a 14-month period (from portal introduction in January 2015 to the end of the study period in March 2016), a Web-based survey for caregivers administered after a minimum of 2 months' exposure to the portal and repeated 2 months later, and focus groups or individual interviews held with caregivers and service providers at the same points in time. The survey assessed caregivers' perceptions of the utility of and satisfaction with the EHR and e-messaging, and the portal's impact on client engagement and perceptions of caregiver-provider communication. A total of 18 caregivers (parents) completed surveys and 6 also took part in focus groups or interviews. In addition, 5 service providers from different disciplines took part in focus groups or interviews.

Results: Although usage patterns varied, the typical pattern was a steady level of use (2.5 times a month over an average of 9 months), which is higher than typically reported use. The portal pages most frequently accessed were the home page, health record main page, appointment main page, and reports main page. The Web-based survey captured caregivers' perceptions of usefulness of and satisfaction with the EHR and portal messaging, as well as the portal's impact on their engagement in care and perceptions of caregiver-provider communication. The surveys indicated a moderate degree of utility of and satisfaction with the portal features, and a low but emerging impact on engagement in care and caregiver-provider communication (survey scales measuring these outcomes displayed excellent internal consistency, with Cronbach alpha ranging from .89 to .95). Qualitative themes from focus groups and interviews supported and extended the survey findings. Caregivers and service providers saw appreciable information benefits and provided recommendations to increase portal use and utility. Caregivers focused on the scope of organizational adoption of the portal system and indicated their hopes for the future of the portal, whereas service providers were concerned about how to best manage their investment of time and effort in preparing client-friendly reports and messaging clients via the portal.

Conclusions: Overall, the findings show the promise of the portal and the need for ongoing evaluation to show the portal's ultimate potential in enhancing engagement in care and communication with care providers.

(*J Med Internet Res* 2017;19(4):e97) doi:[10.2196/jmir.6811](https://doi.org/10.2196/jmir.6811)

KEYWORDS

disability; engagement; health information; caregivers; children with disabilities; patient portals; electronic health records

Introduction

Health care portals have the potential to improve consumers' access to information, engagement in care, and health outcomes [1]. Few children's hospitals in North America offer online portals to families [2] and, to our knowledge, there are no published research studies of portal use, utility, or impact in pediatric rehabilitation hospitals. Through electronic access to health records and e-messaging (secure 2-way messaging), clients and families may feel a greater personal connection to care and partnership in the care process.

In pediatric rehabilitation, the chronic nature of children's difficulties and the need for parental support have resulted in widespread adoption of family-centered care, which stresses the importance of engaging in partnerships with care providers and exchanging information [3]. Engagement in care refers to a sense of meaningful involvement, true collaboration, or authentic participation [4], and the active involvement and full investment of children and parents in therapy is presumed to be essential to achieving positive client outcomes [5,6]. Similarly, good client-provider communication is considered essential in the delivery of children's rehabilitation services [7]. Effective communication is strongly associated with client satisfaction [8,9], and studies have indicated that communication plays an important role in the ongoing client-practitioner relationship, and especially in engaging clients [10].

The Promise of Health Care Portals

Consumers are asking for the ability to view their appointment schedules online, have electronic access to documentation and other resources, and have their questions answered over the Internet. E-scheduling, e-visits, and e-messaging are seen as important portal functions [11], as they allow information exchange and can potentially enhance consumers' ability to manage their own health and health care [1]. By having access to a portal that does more than just provide information and data, consumers can become more active participants in their own health care [12]. Thus, portals have the potential to provide more personalized, collaborative, and effective care. There is, however, limited (but growing) research information on the use of electronic communication in health care, including electronic health records (EHRs) [1,13].

Research on EHRs has typically focused on adult medical care and management of chronic diseases such as rheumatoid arthritis [14] and diabetes [15] rather than children's chronic conditions [16]. We searched for articles on portals in pediatric rehabilitation hospitals (using "pediatric" or "paediatric" with "online" and "portal," with searches limited to 2015 and more recent). We found articles on portals for children with mental health conditions [17], with asthma [18], and those receiving

acute care in a pediatric hospital [19]. We found no articles on portals in pediatric rehabilitation hospitals. A recent evaluation of a portal at a Canadian children's acute care hospital indicated that, at 3 months after EHR implementation, service providers perceived a negative impact on their efficiency, productivity, and the quality of care they provided, whereas patients and families did not perceive a negative impact on care [20].

Even for adult chronic disease management, there is sparse evidence to support the ability of personal health records and EHRs to improve efficiencies, reduce costs, and improve outcomes [15,21]. Most existing research on health care portals has examined the frequency of use of their features, satisfaction with the portal, or efficiencies in health care utilization (eg, number of office and emergency department visits, phone calls) [22-24].

Systematic reviews on the effects of portals with EHR access [25-27] indicate that few studies have investigated outcomes concerning patient engagement or empowerment, with studies typically providing nonsignificant findings. Kruse et al [27] found only 27 studies of relevance to impact—those measuring meaningful outcomes such as patient participation in medical decisions, patient-provider communication, or satisfaction. For example, a study by van der Vaart and colleagues [14] found that approximately 44% of patients with rheumatoid arthritis reported feeling more involved in treatment as a result of access to their EHRs; however, significant pre-post differences in empowerment were not found. Thus, research is needed to substantiate "meaningful use" impacts [1,15].

Summary of Research Gaps

To our knowledge, there has been no research on client portals for pediatric rehabilitation service users. As well, the literature has focused on portal use, satisfaction, and service utilization outcomes, and there is a recognized need to examine meaningful outcomes such as client engagement in care and communication with providers. These 2 outcomes are of fundamental importance in pediatric rehabilitation, as well as in medical services for people with chronic conditions [27]. Furthermore, although some studies have tracked portal use over time (eg, [28]), studies have not examined portal impact prospectively. Examining portal use, utility, and impact over time provides an opportunity to see how users access portal features and may also indicate emergent short-term impacts [29], such as enhanced engagement in care. Since studies typically evaluate portal features in isolation [27], there is benefit in examining client perceptions of the introduction of a full portal over a period of time. Lastly, the literature indicates the value of mixed-methods studies, which are rare in portal research [27].

The Client Portal at Holland Bloorview Kids Rehabilitation Hospital

Holland Bloorview Kids Rehabilitation Hospital (Holland Bloorview) in Toronto, Ontario, is Canada's largest children's rehabilitation hospital. The hospital is a provincial resource for children with cerebral palsy, acquired brain injury, muscular dystrophy, amputation, epilepsy, spina bifida, arthritis, cleft lip and palate, autism, and other physical and developmental disabilities. Holland Bloorview's vision is to create a world of possibility for kids with disability by embracing client- and family-centered care, and participating in applied research and education.

In early 2015, Holland Bloorview launched a consumer portal with the ultimate goal of helping clients and their caregivers (family members, most typically parents) take an active role in managing their own or their children's care. The consumer health portal (called connect2care) was developed in partnership with clients and families [2]. Connect2care provides clients and families with electronic access to their medical records, online appointment cancelling and booking features, transparent and timely access to clinical documentation, and e-messaging to connect with their care providers.

Beginning in January 2015, enrollment to the portal was initiated, first for clients on the inpatient units, and then also for ambulatory and community programs. In this first phase, the functions available included the ability to view the client's schedule and visit history, view and print clinical notes, and update demographic details. In May 2015, improvements were made to the viewing of laboratory and microbiology test results, and new processes were established to increase the number of clinical notes that would flow to the portal (allowing providers from additional health disciplines to share clinical notes). Over the summer of 2015, training of more than 100 health care providers was completed to support this enhanced sharing of clinical notes. In the summer and fall of 2015, messaging functionality was rolled out, allowing secure e-messaging between portal users and their providers. By December 2015, the training of providers for e-messaging was completed.

Enrollment and usage targets for 2015 were established prior to launching the portal, with a target enrollment of 721 users (clients and caregivers), and a target use (unique logins) of 1440. The targets were met well ahead of schedule and, by the end of 2015, there were 869 enrolled users and more than 4800 uses. The adoption rate was approximately 12.41% (869 out of about 7000 unique clients). As well, over 200 staff were live with e-messaging at the end of 2015, including physicians,

occupational therapists, physical therapists, speech-language pathologists, social workers, psychologists, therapeutic recreation staff, orthotics and prosthetics staff, ambulatory care nurses, child life workers, and nurse practitioners.

Study Purpose and Objectives

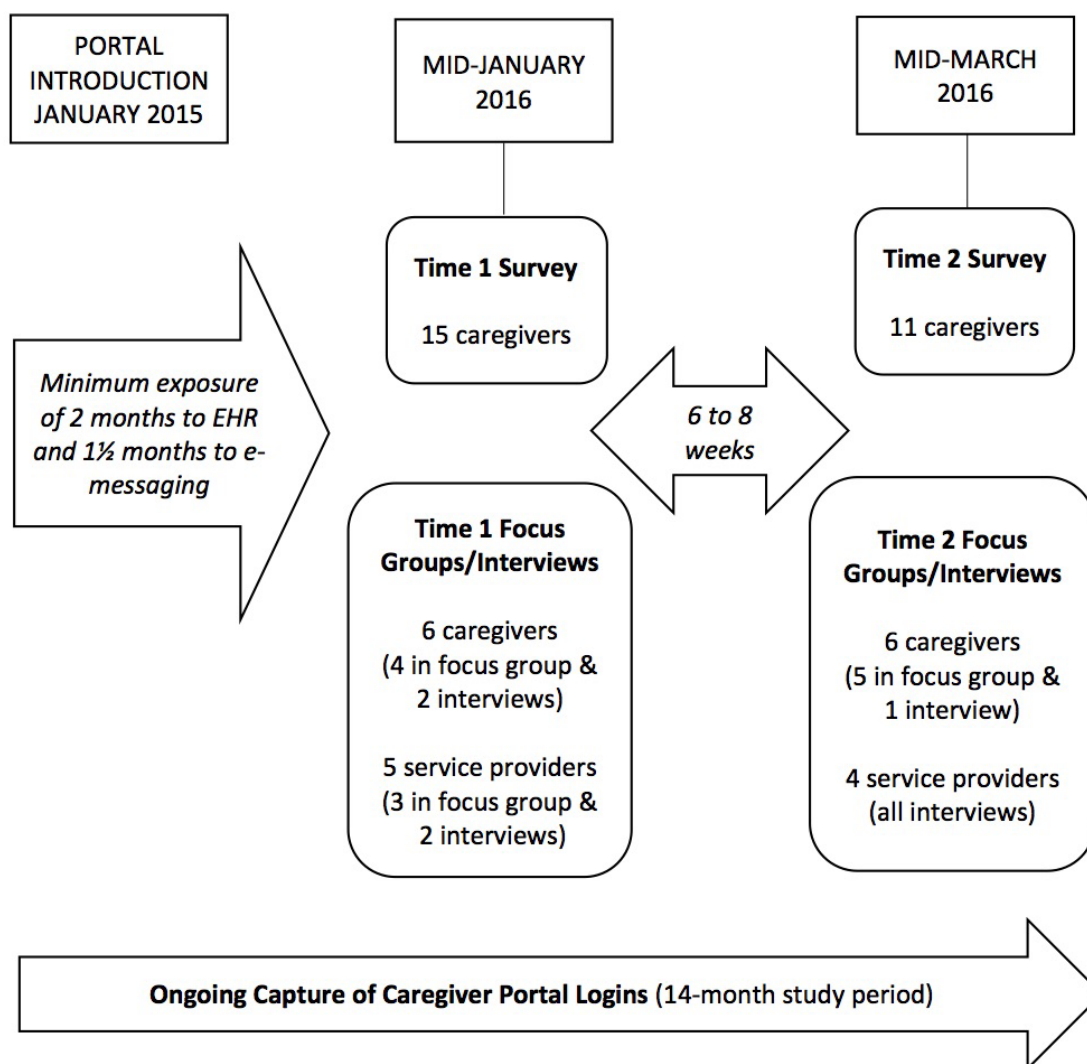
The purpose of our study was to examine the use, utility, and impact of the connect2care portal from the beginning of portal introduction until the end of data collection 14 months later. To meet this aim, we conducted a prospective, mixed-methods study collecting quantitative survey data and qualitative data from focus groups and interviews with caregivers and service providers at 2 points in time. We adopted a concurrent triangulation approach, in which quantitative and qualitative data were collected at the same time [30] and integrated at the level of interpretation [31]. Our research objectives were (1) to determine caregivers' portal use over the study period, (2) to examine levels of perceived usefulness of and satisfaction with the portal's EHR and e-messaging, as well as the portal's impact on engagement in care and caregiver-provider communication, and (3) to ascertain caregivers' and service providers' perceptions of the portal, its utility, and how it could be enhanced.

Methods

Study Overview and Design

The study was conducted by a team with diverse organizational roles, including clinical directors, a project manager, researchers, and a family-centered care specialist who is a parent of a client. The team brought different perspectives and backgrounds (social work, occupational therapy, and psychology) to the design of the study and interpretation of the findings.

As Figure 1 shows, we used 3 methods of data collection: capture of portal login information, a survey (assessing utility and satisfaction, and impact on client engagement and perceptions of caregiver-provider communication), and focus groups or interviews held with caregivers and service providers. To ensure the opportunity for a base of experience prior to assessment, participants had a minimum of 2 months' exposure to the EHR and 1.5 months to e-messaging before time 1 assessment. There was an interval of 6 to 8 weeks between time 1 and time 2 measurement points, allowing us to examine the effects of additional exposure to the portal. Caregivers could opt into 1 of 2 arms of the study: survey only; or survey plus focus group or interview. Service providers participated only in focus groups or interviews.

Figure 1. Data collection procedures: logins, surveys, and focus groups or interviews. EHR: electronic health record.

Participant Recruitment

Caregivers

We obtained ethics approval from Holland Bloorview's Research Ethics Board. Members of Holland Bloorview's Family Advisory Committee introduced caregivers to the study when they were registering for the portal. Interested caregivers provided their contact information to learn more about the study from a research assistant. Other recruitment methods were (1) study flyers, (2) emails sent to connect2care users, (3) an advertisement posted on the Participate in Research page on the hospital's website, and (4) messages posted on the hospital's Facebook page for parents.

We administered study forms for caregivers (invitation, background form, and surveys) online using Research Electronic Data Capture (REDCap), a secure, Web-based app designed to support data capture for research studies [32]; all forms and materials were in English, as this is the language used in connect2care. Other than the ability to read English, there were no inclusion or exclusion criteria. Caregivers (of clients of any age or diagnosis) accessed a link on the portal's home page that provided access to REDCap. The REDCap link contained a

downloadable PDF version of the information letter and a consent form. At this time, caregivers indicated their interest in 1 of 2 study options (survey only or survey plus focus group), and provided their contact information so that study information and consent forms could be sent to them by email. As part of providing consent, caregivers gave permission for their portal usage information to be included in the research.

Service Providers

We recruited service providers using an announcement on the hospital's internal home page, as well as emails with the study information flyer attached. Interested service providers were emailed the background form, which they returned to the research assistant.

Background Forms

The background form for caregivers captured demographic information (sex, education), and information about their child or adolescent with a disability (primary diagnosis). It also captured Internet use and ratings of Internet skills [14], since lack of Internet experience is the primary barrier to portal use [27]. The background form for service providers captured discipline, education, and years in practice.

Portal Survey

Since existing surveys did not capture the outcomes of interest, we developed our own. Satisfaction and utility items were informed by surveys developed to evaluate patient access to EHRs (eg, [14,33,34]). Engagement with care items were informed by the Pediatric Rehabilitation Intervention Measure of Engagement for Parents (GK, unpublished measure, 2015), and caregiver-provider communication items were based on constructs from existing surveys [33,35].

We refined the survey items, piloted them with 6 caregivers to ensure the items were acceptable and easily understood, and then reviewed them for health literacy. Caregivers indicated that the surveys took less than 10 minutes to complete. Based on caregivers' feedback, we clarified the wording (eg, "complete" meant not missing important information), added definitions of key terms to the instruction section (eg, clinical provider team, care, and involved or engaged), and instructed respondents to click on "not applicable" or "to a small extent or not at all" when they had not had relevant opportunities to do what items specified.

The final survey (Portal Survey on Satisfaction and Impact on Care [36]) contained 38 items in 5 sections: (1) utility/satisfaction: client health record, (2) utility/satisfaction: portal messaging, (3) involvement in the care process, (4) impact of portal messaging with service providers (eg, improvements to communication, ability to express concerns and get clarifications, trust or rapport with clinical provider team), and (5) portal overall (ie, satisfaction and utility overall, future intention to use, impact on care).

Focus Groups and Interviews

Focus groups and interviews were conducted by an interviewer with more than 10 years of interview experience. Individual interviews were conducted (either by phone or in person) when individuals could not attend scheduled focus groups. Focus groups and interviews were audio-recorded and transcribed by a trained transcriptionist, with identifying information removed. Focus groups lasted an average of 48 minutes (range 35-60 minutes), and individual interviews lasted an average of 19 minutes (range 7-34 minutes). Participants were asked to discuss (1) general thoughts about the portal (including satisfaction and the extent to which they accessed the portal, and for what purposes), (2) how the portal could be enhanced or improved, (3) the most helpful or useful parts of the portal, (4) difficulties or issues in using or accessing the portal, (5) needs and expectations regarding the portal, and (6) whether the portal made things more efficient.

Data Analysis

Portal Usage by Caregivers

Decision Support and Health Information Management at the hospital provided the study research assistant with an Excel 97-2003 (Microsoft Corporation) workbook of data covering the 14-month study period. We used this information to create a dataset containing the portal pages accessed by date for each person enrolled in the research. We then used this information to calculate total number of visits and days of use per person, usage across months per person, and the number of times each page was viewed over all participants.

Survey Analysis

We analyzed time 1 and time 2 survey data descriptively, given the small sample size. Aggregate scores were calculated for survey scales (the 5 survey sections), and Cronbach alpha was calculated to determine the scales' internal consistency reliabilities.

Qualitative Analysis

We analyzed transcripts using a content analysis approach, which involves coding statements based on key concepts, combining these coded concepts into themes, and then refining the identified themes [37,38]. Team members read all caregiver transcripts. To ensure confidentiality, service provider transcripts were analyzed by 3 team members (not clinical or project directors). Through an iterative process, the lead author then summarized the transcript data into themes, using tables to group and compare related ideas. The credibility and accuracy of data analysis were ensured by using multiple procedures, including maintaining an audit trail [39]. Trustworthiness was enhanced through peer examination and discussion of findings in team meetings, investigator triangulation (various disciplines, perspectives, and roles), and mutual confirmation of the data [40,41].

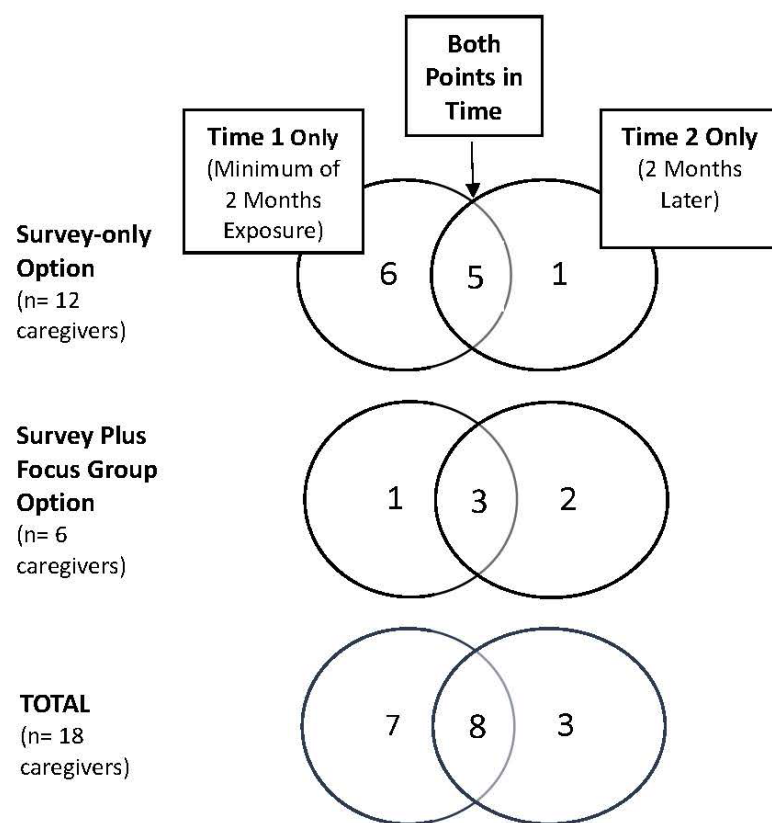
Results

Description of Participants

Caregivers

A total of 18 caregivers took part, some in the survey-only option and some in the survey plus focus group or interview. As Figure 2 shows, some individuals participated at both time points. There were 15 completed surveys at time 1 and 11 at time 2.

Table 1 presents the participants' characteristics for the entire sample. As well, almost all participants (n=17, 94%) had used the Internet for more than 5 years, and all reported using it daily or several days a week. They rated their Internet skills as either good to very good (n=12, 67%) or average to reasonable (n=6, 33%).

Figure 2. Caregiver participants by study options and 2 points in time.**Table 1.** Caregiver participant characteristics (n=18).

Characteristic	n	%
Sex		
Female	15	83
Male	3	17
Age range (years)		
20-34	4	22
35-49	8	44
50-64	6	33
Highest education level		
High school certificate or diploma	1	6
College or other nonuniversity certificate or diploma	3	17
University certificate or diploma (eg, bachelor's)	9	50
Postgraduate certificate or diploma (eg, master's, PhD)	5	28
Child's current primary health or development problem		
Acquired brain injury	4	22
Cerebral palsy	4	22
Autism	3	17
Spina bifida or hydrocephalus	2	11
Other (eg, global developmental delay, muscular dystrophy, central nervous system vasculitis or stroke)	5	28

Service Provider Participants

Our target was to recruit 8 to 10 service providers for the focus groups, as is typical in focus group research [42]. A total of 5 service providers took part in focus group or interviews: 2 physical therapists, 1 occupational therapist, 1 speech-language pathologist, and 1 nurse; 4 of the 5 participating service providers took part at both time points. They had been in practice 13.5 years on average; 2 had a bachelor's and 3 had a master's degree.

Portal Usage by Caregivers

Table 2 presents portal usage information over the study period, from the introduction of the portal (January 2015) to the end of data collection (March 2016). As Figure 1 shows, we collected login information from first use of the portal, whereas we held surveys and focus groups or interviews only after participants

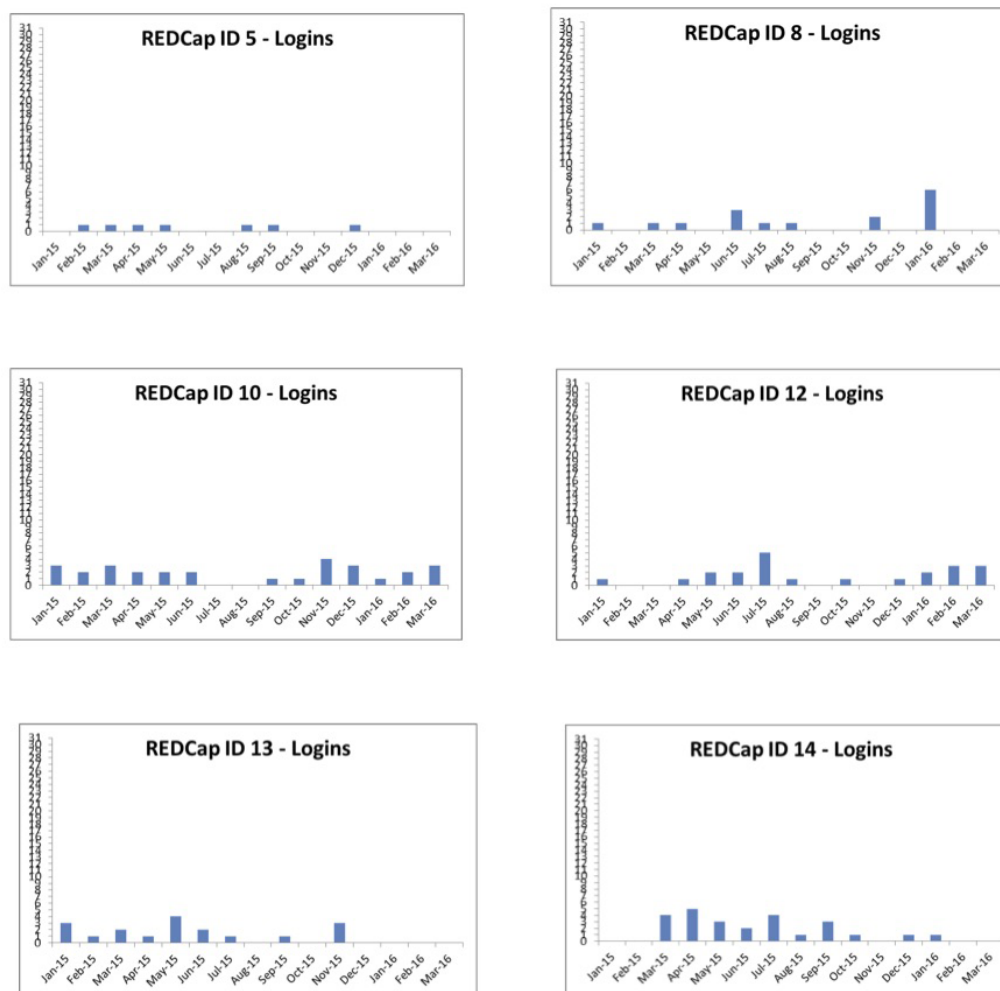
had a minimum of 2 months' exposure to the portal (to ensure a base of experience). Users joined connect2care at various points over this time period, with an average exposure of 253 days (approximately 9 months). Overall, users logged on to the portal an average of 22.2 times on 19.2 days (2.5 times a month). The most common user access pattern was a combination of home page, health record main page, appointment main page, and reports main page. Thus, users were most interested in their child's health record, appointments, and reports, as has been reported by others [20].

Login graphs for participants indicated differences in patterns of usage, with some users having high initial use that tapered off, whereas others had no use past their initial portal logon. The typical pattern, though, was a steady level of use (2 or 3 times a month). We chose the graphs in Figure 3 to show typical login patterns over the study period.

Table 2. Connect2care portal usage by caregivers over the 14-month study period.

Variable	Mean	Range
Number of days with exposure to connect2care (last session date minus first session date)	253	1-433
Number of times logged in to connect2care	22.2	1-87
Number of days logged in to connect2care	19.2	1-69

Figure 3. Typical login patterns (total number of logins per month) over the study period for 6 study participants. ID: identification; REDCap: Research Electronic Data Capture.



Survey Results at Time 1 and Time 2

Survey Responses

Table 3 shows mean scores for the survey scales, along with the scales' internal consistencies. Cronbach alpha values ranged from .89 to .95, indicating excellent internal consistency reliability [43,44]. There was a moderate extent of agreement with respect to utility of and satisfaction with the client health record and portal messaging. Scores for involvement in the care process and for the portal overall indicated agreement between "neutral" (score of 0) and "agree" (score of 1). Mean scores for portal messaging with service providers were neutral (−0.06) at time 1 and slightly higher (0.25) at time 2.

Mean Survey Item Responses

Table 4 shows a fair degree of overlap in items with the highest and lowest means at time 1 and time 2. The items with the

highest means (greatest endorsement) concerned the accuracy and convenience of the client health record (section A), the usefulness, timeliness, and adequacy of portal messaging (B), a willingness to invest effort in the intervention process (C), and convenience in accessing care or services (D). There was also evidence of a strong intention to use the portal in the future (E). Items with the lowest means (least endorsement) concerned actual use of the client health record (although this was moderate; section A) and portal messaging (to a fairly small extent; B). There was little evidence that the portal led to feelings of greater involvement in the care process (C), improved ability to express concerns to providers or enhanced relationships with providers (D), or reduced number of in-person visits (E).

Table 3. Survey scale scores at time 1 and time 2.

Survey section	No. of items	Cronbach alpha (times 1 and 2 combined)	Time 1 (n=15)		Time 2 (n=11)	
			Mean ^a	(SD)	Mean ^a	(SD)
Utility/satisfaction: client health record ^b	6	.89	3.68	0.76	3.14	1.01
Utility/satisfaction: portal messaging ^b	6	.90	2.62	1.08	3.22	1.11
Involvement in the care process ^c	10	.91	0.69	0.77	0.77	0.52
Impact of portal messaging with service providers ^c	7	.95	−0.06	0.67	0.25	0.99
Portal overall ^c	9	.90	0.61	0.41	0.40	0.97

^aSome data are missing (a score was calculated only if a respondent had data for ≥50% of the items in the scale).

^bScale choices were to a great extent=5, to a fairly great extent=4, to a moderate extent=3, to a fairly small extent=2, to a small extent or not at all=1.

^cScale choices were strongly agree=2, agree=1, neutral=0, disagree=−1, strongly disagree=−2.

Table 4. Survey items with the highest and lowest means at time 1 and time 2.

Survey section	Time 1 (n=15) ^a		Time 2 (n=11) ^a	
	Item(s) with the highest mean	Item(s) with the lowest mean	Item(s) with the highest mean	Item(s) with the lowest mean
A. Utility/satisfaction: client health record ^b <i>To what extent did you...</i>	Feel the client health record was accurate (there were no mistakes)? (mean 4.07)	Access your/your child's health record? (mean 2.93)	Find it easy to locate relevant information in the client health record? (mean 3.82)	Access your/your child's health record? (mean 2.45)
B. Utility/satisfaction: portal messaging ^b <i>To what extent did you...</i>	Feel this feature was useful to you? (mean 3.60)	Communicate with a member of the provider team using portal messaging? (mean 1.70)	Feel responses to your portal messages were timely? (mean 3.83) Feel responses to your portal messages were adequate? (mean 3.83)	Communicate with a member of the provider team using portal messaging? (mean 2.00)
C. Involvement in the care process ^c <i>Based on my portal experience...</i>	I am willing to invest effort in the intervention process (mean 1.20)	I believe the intervention process (ie, treatment plan) will be effective (due to greater involvement in the care process) (mean 0.40)	I am willing to invest effort in the intervention process (mean 1.27)	I believe the intervention process (ie, treatment plan) will be effective (due to greater involvement in the care process) (mean 0.18)
D. Impact of portal messaging with service providers ^c <i>Because of the portal...</i>	I feel that accessing care or services is more convenient for me (mean 0.67)	My ability to express concerns and/or provide comments to providers has improved (mean −0.36) I have built a more open and trusting relationship with health care providers (mean −0.36)	I feel that accessing care or services is more convenient for me (mean 0.64)	My communication with my provider has improved (mean 0.10) I have built a more open and trusting relationship with health care providers (mean 0.10) My sense of trust/rapport with the clinical provider team has increased (mean 0.10)
E. Portal overall ^c	I intend to use the portal in the future (mean 1.47)	The portal reduced the number of in-person visits I made (mean −0.27)	I intend to use the portal in the future (mean 1.00)	The portal reduced the number of in-person visits I made (mean −0.36)

^aSome data are missing (a score was calculated only if a respondent had data for ≥50% of the items in the scale).

^bScale choices were to a great extent=5, to a fairly great extent=4, to a moderate extent=3, to a fairly small extent=2, to a small extent or not at all=1.

^cScale choices were strongly agree=2, agree=1, neutral=0, disagree=−1, strongly disagree=−2.

Transcript Themes

Caregiver Themes

Three themes arose from our analysis of the caregiver transcripts, with the nature of these changing slightly at time 2

as caregivers acquired more experience with the portal. These themes were information benefits, recommendations to increase use and utility, and scope of adoption and future vision.

With respect to information benefits, at time 1, caregivers indicated the usefulness of the portal in providing easy access

to their child's medical history, reports, and appointments, and how this saved them time. At time 2, the benefits of access to the information itself received somewhat greater emphasis—participants expressed an appreciation for having more detailed information and knowing the technical language, as they felt they could then communicate on a more level playing field with providers (eg, “[if] we can use accurate language, language that they understand, we get a lot better dynamic where they're going to listen to us”).

There were many recommendations to increase use and utility, including being able to message all clinicians, receiving speedier notification of messages waiting for them in the portal (caregivers had to go into connect2care to get their emails, and notifications about awaiting emails took time to arrive) and better information about how to do things on the portal (eg, send a message), and having clarification of the scope of confidentiality and portal access after discharge. The recommendations also included uniform implementation of portal features (some caregivers did not have access to features that others did) and greater comprehensiveness in what was available (eg, medication information, care plans, and particularly reports from all clinicians). These comments may reflect differences in whether a caregiver's service providers were set up on e-messaging.

The theme of scope of adoption and future vision concerned caregivers' ultimate hopes for the portal. They were aware that the portal was a “work in progress” and “in its infancy,” and shared their visions of the portal in the future. They commented on the utility of cross-organization EHRs, connection to the adult health care system, provision of information about available programs, and personalization on a broad scale (eg, ability to upload one's own information to the portal and receive targeted information). They also commented on the need for organizationwide adoption, where portal use was embedded in routine practice (eg, “it's early days...but if just everybody could get on board...the opportunities for saving everybody's time and aggravation in messaging are huge”).

Service Provider Themes

Four themes emerged from our analysis of the service provider transcripts. There were no differences in these themes at time 1 and time 2. First, the theme utility for families indicated that service providers saw the utility of the portal in setting up appointments and providing secure messaging. They were less sure about the impact of the portal on client engagement but felt that the portal provided a positive, inviting message to families about being engaged (eg, “messages of we want you to be engaged, we're trying to reach out to you, we're trying to have multiple ways of engaging with you”).

Second, service providers identified technical shortcomings in several areas, including lack of notification of emails (requiring repeated logons to check whether emails from caregivers had been received), and lack of ability to post vacation messages and upload attachments. Problems in formatting occurred when reports from the clinical system were uploaded to connect2care, and reports had to be changed in several places to ensure consistency in what clients saw.

The third theme dealt with uncertainties in portal use, related to lack of knowledge, comfort, or confidence in using the portal, in addition to some of the portal's technical shortcomings. Service providers expressed uncertainty in knowing which families accessed the portal, whether reports were being used, whether a message was waiting or had been received, and what documents or materials families could access. They were also uncertain about their role in informing families about the portal and how to access it (eg, “it doesn't seem like a big deal, but when I'm seeing 5, sometimes 6 people in a day, I don't have that extra time to do that”).

The fourth service provider theme concerned use, effort, and investment in the portal. Due to low levels of perceived use by families, and the time, effort, and care required to see if messages were waiting and to produce user-friendly reports meeting professional standards, service providers expressed concerns about whether it was currently worth investing a great amount of time in the portal (eg, “it's another thing to log in to,” “doesn't make sense for me to go in every day, or multiple times a day, and check if in 6 months I've only ever got 2 messages”).

Discussion

This study contributes to the growing literature on portals, as there has been no previous research, to our knowledge, on the use, utility, or impact of client portals in pediatric rehabilitation. Furthermore, studies of health care portals have largely focused on use and utility, rather than impact-related outcomes such as engagement in care or client-provider communication [27]. Compared with new portal annual adoption rates of 5% to 10% and access rates of 0.4 to 0.6 uses per month per user [45], the connect2care portal had an adoption rate of 12.4% and 2.5 logins a month per user over 9 months of exposure. There may be many reasons for the higher adoption and access rates, including the involvement of families in the development of the portal [2], effort put into informing families about the portal, and differences between acute care and pediatric rehabilitation contexts.

There was a moderate degree of perceived usefulness of and satisfaction with the EHR and e-messaging features, and evidence that the portal was perceived to provide useful access to the clinical record. Utility and satisfaction scores did not change much over time; however, there was only a 6- to 8-week time period between the 2 survey administrations, thus reducing the amount of new portal experience that was possible.

With respect to impact, there was some evidence that portal access facilitated caregivers' perceptions of engagement in care, but this evidence was not strong (between neutral and agree, on average). As a point of comparison, van der Vaart et al [14] found that 44% of patients with rheumatoid arthritis felt more involved in their treatment as a result of portal access. Service providers in our study also indicated seeing little evidence of increased engagement. However, several caregivers indicated that reading available reports increased their understanding of the related technical language, and felt this enhanced their ability to engage in conversations about care with service providers.

We speculate that it takes time for portal use to have an effect on engagement in care (in this study, we collected time 1 engagement data after an average of approximately 9 months of portal exposure). At this early stage of portal use, caregivers appeared predominantly interested in the information they could access via the portal, although at the second measurement point 2 months later, there was some indication of greater appreciation of the actual content of the reports, which may reflect the fact that more and more clinical documentation was being added through the study period. With respect to impact on caregiver-provider communication, there was some evidence at time 2 of improved communication due to the introduction of e-messaging. We speculate that e-messaging had not been introduced long enough or widely enough to affect client engagement. Alternatively, caregivers may be already quite engaged in their child's care, so that the portal may not make such a big difference in engagement compared with nonpediatric portals.

The qualitative themes were informative, as they allowed comparison of caregivers' and service providers' perspectives. There was a common emphasis on the utility of the information provided via the portal. As well, both groups questioned the extent of the impact on engagement in care at this early stage of portal rollout (corroborating survey data), and both groups recommended increasing portal use and utility by addressing technical shortcomings. These recommendations were related to the portal not meeting expectations for technology and not being as user friendly as desired.

The major difference between caregivers and providers was that caregivers focused on the scope of adoption of the portal system in the organization and expressed their hopes for the future of the portal with respect to their family's life journey (the scope of adoption and future vision theme), whereas service providers were concerned about how to best manage their investment of time and effort (the use, effort, and investment in the portal theme). Both groups expressed a desire for the other group to use the portal to a greater extent: caregivers wanted to see organizationwide adoption, whereas some providers in our small sample questioned whether their investment in the portal was justified given low levels of perceived use. As well as full organizationwide adoption of the portal system, caregivers' hopes for the future included greater personalization and comprehensiveness of the provided information. As recognized by caregivers, portals are in their infancy and the maturity of portals does not appear to be where it needs to be to improve quality of care and involve the patient in care decisions [27].

Planned future enhancements to the connect2care portal will address some of the hopes of caregivers, including the ability to import reports from other service providers, link to records from other hospitals, and receive materials and resources specifically targeted to their care needs [2].

Study Strengths and Limitations

Strengths of this investigation include prospective data collection of login information starting at the portal's introduction, the breadth of information collected using multiple methods, inclusion of both caregiver and provider perspectives, and focused examination of caregiver engagement and

caregiver-provider communication using scales with excellent internal consistency. Study limitations include its descriptive nature, the relatively short time period between measurement points (6-8 weeks), and the relatively small number of respondents (18 caregivers and 5 providers). We likely did not have data saturation and therefore the robustness of the qualitative themes is uncertain [38,46], although triangulation with the survey data should be noted.

The study participants are likely representative of those who register for portal access but *not* of the Holland Bloorview client population overall, since people who register for portals are likely to be early adopters who embrace technology more readily. Research suggests that those who use patient portals are generally more highly educated, younger, more affluent, and have fewer medical problems than nonenrollees [28,47,48]. As with any innovation, expectations and technical difficulties play a role in adoption. Difficulties in getting portal technology to run as desired can affect portal use and may bias reports of utility, satisfaction, and impact [33,49], as may be the case in our study.

Research and Organizational Implications

Future research directions include continued use of the survey to see whether there is evidence of increasing impact of the portal on client outcomes. Other directions include examination of expectations that are held about portal use and utility, and examination of the perceptions of young adult clients. For other researchers, this study has indicated the utility of analyzing login statistics and using a portal survey with demonstrated internal consistency. The study has also indicated the utility of an integrated knowledge translation approach, where clinicians, family members, and researchers come together to address an important applied question [50].

Organizationally, the findings indicate specific areas for improvements to the portal and its processes, many of which are being or have been addressed as the portal continues to be developed, including the establishment of processes to improve efficiency related to portal activities (such as reminder emails when a new message is received) and mechanisms to support the enrollment of all interested clients and caregivers. The findings also suggest the need for ongoing education about portal intents and for transparency in communication about implementation issues. The literature indicates that education is needed to manage expectations and enhance the extent of portal adoption [27]. Education and resources are needed to support providers in feeling confident that their clinical documentation is family-friendly and in using e-messaging effectively. Similarly, education is needed to enable caregivers to make the best possible use of the portal features.

Although caregivers in our sample saw the portal's value, this was not yet the case for the service providers. This suggests that organizations may need to focus some of their efforts on ensuring that service providers see this value, as they are the ones who will share clinical notes and communicate via e-messaging—they too are partners in care. In directing attention to engaging clients or families, the engagement of service providers can be given less attention. As technology becomes more prominent in health care settings, it will be up to

organizations to support uptake and demonstrate potential benefits for both parties. An important future direction at Holland Bloorview is therefore to share family stories with providers so that they can better understand the portal's positive impact and feel that their time related to sharing information and communicating via the portal is worth the effort.

Lastly, the findings endorse the often-made statement that changing organizational culture takes time. Portal adoption is a process—not a one-time event—and requires a feedback loop (as provided by this study), allowing an organization to improve portal adoption through attention to the needs of the people who use it.

Acknowledgments

Canada Health Infoway Inc. funded the demonstration project supporting the development and evaluation of Holland Bloorview Kids Rehabilitation Hospital's client/family EHR portal. Canada Health Infoway Inc. also covered the publication costs for this article. Canada Health Infoway Inc. is an independent, not-for-profit organization funded by the federal government that works to accelerate the development, adoption, and effective use of digital health across Canada. The funders had no role in the data collection or decision to publish the manuscript. Simon Hagens (affiliated with Canada Health Infoway) was a member of the research team and involved in study design and interpretation. Gillian King holds the Canada Research Chair in Optimal Care for Children with Disabilities, funded by the Canadian Institutes of Health Research.

We thank Siamak Tenzif for assistance with obtaining information about usage patterns and frequency of login information, Becky Quinlan for assistance with registering people for the portal, and Alanna Rudzik for conducting the focus groups and interviews. We would also like to thank those who participated in the study and members of the Family Advisory Committee who provided input on the survey.

Conflicts of Interest

Simon Hagens is affiliated with Canada Health Infoway.

References

1. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). *J Am Med Inform Assoc* 2008;15(6):729-736 [FREE Full text] [doi: [10.1197/jamia.M2547](https://doi.org/10.1197/jamia.M2547)] [Medline: [18756002](https://pubmed.ncbi.nlm.nih.gov/18756002/)]
2. Maxwell J, Williams L, Adamson K, Karmali A, Quinlan B. Supporting client and family engagement in care through the planning and implementation of an online consumer health portal. *J Patient Experience* 2016 Aug 11;3(2):48-51. [doi: [10.1177/2374373516652256](https://doi.org/10.1177/2374373516652256)]
3. Rosenbaum P, King S, Law M, King G, Evans J. Family-centred service: a conceptual framework and research review. *Phys Occup Ther Pediatr* 2009 Jul 29;18(1):1-20. [doi: [10.1080/J006v18n01_01](https://doi.org/10.1080/J006v18n01_01)]
4. King G, Currie M, Petersen P. Child and parent engagement in the mental health intervention process: a motivational framework. *Child Adolesc Ment Health* 2012 Dec 14;19(1):2-8. [doi: [10.1111/camh.12015](https://doi.org/10.1111/camh.12015)]
5. Davidson J, Wiens S, Anderson K. Creating a provincial family council to engage youth and families in child & youth mental health systems. *J Can Acad Child Adolesc Psychiatry* 2010 Aug;19(3):169-175 [FREE Full text] [Medline: [20842271](https://pubmed.ncbi.nlm.nih.gov/20842271/)]
6. Tetley A, Jinks M, Huband N, Howells K. A systematic review of measures of therapeutic engagement in psychosocial and psychological treatment. *J Clin Psychol* 2011 Sep;67(9):927-941. [doi: [10.1002/jclp.20811](https://doi.org/10.1002/jclp.20811)] [Medline: [21633956](https://pubmed.ncbi.nlm.nih.gov/21633956/)]
7. King G, Cathers T, King S, Rosenbaum P. Major elements of parents' satisfaction and dissatisfaction with pediatric rehabilitation services. *Child Health Care* 2010 Jun 07;30(2):111-134. [doi: [10.1207/S15326888CHC3002_3](https://doi.org/10.1207/S15326888CHC3002_3)]
8. Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, Participants in the American Academy on Physician and Patient's Conference on Education and Evaluation of Competence in Communication and Interpersonal Skills. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med* 2004 Jun;79(6):495-507. [Medline: [15165967](https://pubmed.ncbi.nlm.nih.gov/15165967/)]
9. Mercer LM, Tanabe P, Pang PS, Gisondi MA, Courtney DM, Engel KG, et al. Patient perspectives on communication with the medical team: pilot study using the Communication Assessment Tool-Team (CAT-T). *Patient Educ Couns* 2008 Nov;73(2):220-223. [doi: [10.1016/j.pec.2008.07.003](https://doi.org/10.1016/j.pec.2008.07.003)] [Medline: [18703306](https://pubmed.ncbi.nlm.nih.gov/18703306/)]
10. King G, Desmarais C, Lindsay S, Piérart G, Tétreault S. The roles of effective communication and client engagement in delivering culturally sensitive care to immigrant parents of children with disabilities. *Disabil Rehabil* 2015;37(15):1372-1381. [doi: [10.3109/09638288.2014.972580](https://doi.org/10.3109/09638288.2014.972580)] [Medline: [25323397](https://pubmed.ncbi.nlm.nih.gov/25323397/)]
11. Stylus Consulting. Nova Scotia Personal Health Record demonstration project: benefits evaluation report. Toronto, ON: Canada Health Infoway; 2014. URL: <https://www.infoway-inforoute.ca/en/component/edocman/1995-nova-scotia-personal-health-record-demonstration-project-benefits-evaluation-report/view-document> [accessed 2017-03-27] [WebCite Cache ID 6pHHM3wv7]
12. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *J Am Med Inform Assoc* 2006;13(2):121-126 [FREE Full text] [doi: [10.1197/jamia.M2025](https://doi.org/10.1197/jamia.M2025)] [Medline: [16357345](https://pubmed.ncbi.nlm.nih.gov/16357345/)]

13. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. *J Am Med Inform Assoc* 2004;11(6):505-513 [[FREE Full text](#)] [doi: [10.1197/jamia.M1593](https://doi.org/10.1197/jamia.M1593)] [Medline: [15299001](#)]
14. van der Vaart R, Drossaert CH, Taal E, Drossaers-Bakker KW, Vonkeman HE, van de Laar MA. Impact of patient-accessible electronic medical records in rheumatology: use, satisfaction and effects on empowerment among patients. *BMC Musculoskelet Disord* 2014;15:102 [[FREE Full text](#)] [doi: [10.1186/1471-2474-15-102](https://doi.org/10.1186/1471-2474-15-102)] [Medline: [24673997](#)]
15. Ahmad FS, Tsang T. Diabetes prevention, health information technology, and meaningful use: challenges and opportunities. *Am J Prev Med* 2013 Apr;44(4 Suppl 4):S357-S363. [doi: [10.1016/j.amepre.2012.12.020](https://doi.org/10.1016/j.amepre.2012.12.020)] [Medline: [23498299](#)]
16. Patterson ES, Zhang J, Abbott P, Gibbons MC, Lowry SZ, Quinn MT, et al. Enhancing electronic health record usability in pediatric patient care: a scenario-based approach. *Jt Comm J Qual Patient Saf* 2013 Mar;39(3):129-135. [Medline: [23516763](#)]
17. Power TJ, Michel J, Mayne S, Miller J, Blum N, Grundmeier R. Coordinating systems of care using health information technology: development of the ADHD Care Assistant. *Adv School Ment Health Promotion* 2016;9(3-4):201-218.
18. Fiks AG, DuRivage N, Mayne SL, Finch S, Ross ME, Giacomini K, et al. Adoption of a portal for the primary care management of pediatric asthma: a mixed-methods implementation study. *J Med Internet Res* 2016 Jun 29;18(6):e172 [[FREE Full text](#)] [doi: [10.2196/jmir.5610](https://doi.org/10.2196/jmir.5610)] [Medline: [27357835](#)]
19. Kelly MM, Hoonakker PL, Dean SM. Using an inpatient portal to engage families in pediatric hospital care. *J Am Med Inform Assoc* 2017 Jan;24(1):153-161. [doi: [10.1093/jamia/ocw070](https://doi.org/10.1093/jamia/ocw070)] [Medline: [27301746](#)]
20. Children's Hospital of Eastern Ontario. Benefits evaluation report: Canada Health Infoway deliverable AMBHS-17B. Toronto, ON: Canada Health Infoway; 2016 Feb 29. URL: <https://www.infoway-inforoute.ca/en/component/edocman/3032-children-s-hospital-for-eastern-ontario-epiccare-ambulatory-program-benefits-evaluation-report/view-document> [accessed 2017-03-27] [[WebCite Cache ID 6pHHqxWge](#)]
21. Tenforde M, Jain A, Hickner J. The value of personal health records for chronic disease management: what do we know? *Fam Med* 2011 May;43(5):351-354 [[FREE Full text](#)] [Medline: [21557106](#)]
22. Chen C, Garrido T, Chock D, Okawa G, Liang L. The Kaiser Permanente Electronic Health Record: transforming and streamlining modalities of care. *Health Aff (Millwood)* 2009 Apr;28(2):323-333 [[FREE Full text](#)] [doi: [10.1377/hlthaff.28.2.323](https://doi.org/10.1377/hlthaff.28.2.323)] [Medline: [19275987](#)]
23. Palen TE, Ross C, Powers JD, Xu S. Association of online patient access to clinicians and medical records with use of clinical services. *JAMA* 2012 Nov 21;308(19):2012-2019. [doi: [10.1001/jama.2012.14126](https://doi.org/10.1001/jama.2012.14126)] [Medline: [23168824](#)]
24. Emont S. Measuring the impact of patient portals: what the literature tells us. Oakland, CA: California HealthCare Foundation; 2011 May. URL: <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/PDF%20M/PDF%20MeasuringImpactPatientPortals.pdf> [accessed 2017-03-27] [[WebCite Cache ID 6pHI5UCK0](#)]
25. Urowitz S, Smith K, Alkazaz N, Apatu E, Quartey NK, Wiljer D. Patient accessible electronic health records for the chronically ill: a review of the literature. *J Hosp Admin* 2012 Nov 01;1(2):64-72. [doi: [10.5430/jha.v1n2p64](https://doi.org/10.5430/jha.v1n2p64)]
26. 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](https://doi.org/10.2196/jmir.2238)] [Medline: [23183044](#)]
27. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. *J Med Internet Res* 2015;17(2):e44 [[FREE Full text](#)] [doi: [10.2196/jmir.3171](https://doi.org/10.2196/jmir.3171)] [Medline: [25669240](#)]
28. 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](https://doi.org/10.1197/jamia.M1833)] [Medline: [16221943](#)]
29. Currie M, King G, Rosenbaum P, Law M, Kertoy M, Specht J. A model of impacts of research partnerships in health and social services. *Eval Program Plann* 2005 Nov;28(4):400-412. [doi: [10.1016/j.evalproplan.2005.07.004](https://doi.org/10.1016/j.evalproplan.2005.07.004)]
30. Hanson WE, Creswell JW, Clark VLP, Petska KS, Creswell JD. Mixed methods research designs in counseling psychology. *J Couns Psychol* 2005;52(2):224-235. [doi: [10.1037/0022-0167.52.2.224](https://doi.org/10.1037/0022-0167.52.2.224)]
31. Hall B, Howard K. A synergistic approach: conducting mixed methods research with typological and systematic design considerations. *J Mixed Methods Res* 2008 Jul 01;2(3):248-269. [doi: [10.1177/1558689808314622](https://doi.org/10.1177/1558689808314622)]
32. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](#)]
33. Nazi KM, Hogan TP, McInnes DK, Woods SS, Graham G. Evaluating patient access to electronic health records: results from a survey of veterans. *Med Care* 2013 Mar;51(3 Suppl 1):S52-S56. [doi: [10.1097/MLR.0b013e31827808db](https://doi.org/10.1097/MLR.0b013e31827808db)] [Medline: [23407012](#)]
34. System and use survey: e-visits. Toronto, ON: Canada Health Infoway; 2015 Jun 11. URL: <https://www.infoway-inforoute.ca/en/resource-centre/advanced-search?q=System+and+Use+Survey++e> [accessed 2017-03-25] [[WebCite Cache ID 6pEO8o8IF](#)]
35. Kristjansson E, Tugwell PS, Wilson AJ, Brooks PM, Driedger SM, Gallois C, et al. Development of the effective musculoskeletal consumer scale. *J Rheumatol* 2007 Jun;34(6):1392-1400. [Medline: [17552066](#)]
36. Holland Bloorview Kids Rehabilitation Hospital. Portal Survey on Satisfaction and Impact on Care. flintbox. Chicago, IL: Wellspring Worldwide, LLC; 2017 Jan 04. URL: <https://flintbox.com/public/project/30920> [accessed 2017-03-21] [[WebCite Cache ID 6p8Qi5l8x](#)]

37. Fiese BH, Bickham NL. Qualitative inquiry: an overview for pediatric psychology. *J Pediatr Psychol* 1998;23(2):79-86. [Medline: [9585634](#)]
38. Lincoln Y, Guba E. *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications; 1985.
39. Creswell J. *Qualitative Inquiry and Research Design*. 2nd edition. Thousand Oaks, CA: Sage Publications; 2007.
40. Krefting L. Rigor in qualitative research: the assessment of trustworthiness. *Am J Occup Ther* 1991 Mar;45(3):214-222. [Medline: [2031523](#)]
41. 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](#)]
42. Krueger R, editor. *Focus Groups: A practical Guide for Applied Research*. 3rd edition. Thousand Oaks, CA: Sage Publications; 2000.
43. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford, UK: Oxford University Press; 1989:79-96.
44. Tabachnik B, Fidell L. *Using Multivariate Statistics*. New York, NY: Harper Collins; 1989.
45. Gheorghiu B, Hagens S. Use and maturity of electronic patient portals. *Stud Health Technol Inform* 2017;234:136-141. [Medline: [28186030](#)]
46. Strauss A, Corbin J. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 2nd edition. Thousand Oaks, CA: Sage Publications; 1998.
47. Leveille SG, Huang A, Tsai SB, Allen M, Weingart SN, Iezzoni LI. Health coaching via an internet portal for primary care patients with chronic conditions: a randomized controlled trial. *Med Care* 2009 Jan;47(1):41-47. [doi: [10.1097/MLR.0b013e3181844dd0](#)] [Medline: [19106729](#)]
48. Or CK, Karsh B. A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc* 2009 Aug;16(4):550-560 [FREE Full text] [doi: [10.1197/jamia.M2888](#)] [Medline: [19390112](#)]
49. Krause MS. Behavioral indexes of motivation for treatment. *J Couns Psychol* 1967;14(5):426-435. [doi: [10.1037/h0024937](#)]
50. King G, Thomson N, Rothstein M, Kingsnorth S, Parker K. Integrating research, clinical care, and education in academic health science centers. *J Health Organ Manag* 2016 Oct 10;30(7):1140-1160. [doi: [10.1108/JHOM-11-2015-0177](#)] [Medline: [27700478](#)]

Abbreviations

EHR: electronic health record

REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 14.10.16; peer-reviewed by S Lindsay, MC Trudel, K Nazi; comments to author 28.01.17; revised version received 22.02.17; accepted 15.03.17; published 06.04.17.

Please cite as:

King G, Maxwell J, Karmali A, Hagens S, Pinto M, Williams L, Adamson K
Connecting Families to Their Health Record and Care Team: The Use, Utility, and Impact of a Client/Family Health Portal at a Children's Rehabilitation Hospital
J Med Internet Res 2017;19(4):e97
URL: <http://www.jmir.org/2017/4/e97/>
doi: [10.2196/jmir.6811](#)
PMID: [28385680](#)

©Gillian King, Joanne Maxwell, Amir Karmali, Simon Hagens, Madhu Pinto, Laura Williams, Keith Adamson. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 06.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Original Paper

Young People's Preferences for an Asthma Self-Management App Highlight Psychological Needs: A Participatory Study

Dorian Peters¹, BHA, MMDes; Sharon Davis², BPharm, PhD; Rafael Alejandro Calvo¹, Lic, PhD; Susan M Sawyer^{3,4,5}, MBBS, MD; Lorraine Smith^{6*}, BA, PhD; Juliet M Foster^{2*}, BSc, PhD

¹School of Electrical and Information Engineering, Faculty of Engineering, University of Sydney, Sydney, Australia

²Clinical Management Group, Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia

³Royal Children's Hospital Centre for Adolescent Health, Melbourne, Australia

⁴Department of Paediatrics, The University of Melbourne, Melbourne, Australia

⁵Murdoch Childrens Research Institute, Melbourne, Australia

⁶Faculty of Pharmacy, University of Sydney, Sydney, Australia

*these authors contributed equally

Corresponding Author:

Dorian Peters, BHA, MMDes

School of Electrical and Information Engineering

Faculty of Engineering

University of Sydney

Bldg J03, Electrical Engineering

Maze Crescent

Sydney, 2006

Australia

Phone: 61 0289015604

Fax: 61 93513847

Email: dorian.peters@sydney.edu.au

Abstract

Background: Although the prevalence of mental illness among young people with asthma is known to be twice the rate of the wider population, none of the asthma apps reported have acknowledged or attempted to include psychological support features. This is perhaps because user involvement in the development of asthma apps has been scarce. User involvement, facilitated by participatory design methods, can begin to address these issues while contributing insights to our understanding of the psychological experience associated with asthma and how technology might improve quality of life.

Objective: The goal of this participatory user research study was to explore the experience, needs, and ideas of young people with asthma while allowing them to define requirements for an asthma app that would be engaging and effective at improving their well-being.

Methods: Young people aged 15-24 years with doctor-diagnosed asthma were invited to participate in a participatory workshop and to complete a workbook designed to elicit their thoughts and ideas about living with asthma, technology use, and the design of an app. Participants generated a number of artifacts (including collages, concept maps, and paper prototypes) designed to reify their ideas, tacit knowledge, and experience.

Results: A total of 20 participants (mean age 17.8 years; 60%, 12/20 female) representing a range from inadequately to well-controlled asthma completed a workbook and 13 of these also took part in a workshop (four workshops were held in total), resulting in 102 participant-generated artifacts. Theoretical thematic analysis resulted in a set of personal needs, feature ideas, and app characteristics considered relevant by young people for an asthma support app. The data revealed that psychological factors such as anxiety, and impediments to autonomy, competence, and relatedness (as consistent with self-determination theory [SDT]), were considered major influences on quality of life by young people with asthma. Furthermore, the incorporation of features pertaining to psychological experience was particularly valued by participants.

Conclusions: In addition to practical features for asthma management, an app for young people with asthma should include support for the mental health factors associated with lived experience (ie, anxiety, lack of autonomy, and social disconnectedness). We show how support for these factors can be translated into design features of an app for asthma. In addition to informing the

development of asthma-support technologies for young people, these findings could have implications for technologies designed to support people with chronic illness more generally.

(*J Med Internet Res* 2017;19(4):e113) doi:[10.2196/jmir.6994](https://doi.org/10.2196/jmir.6994)

KEYWORDS

asthma; mobile applications; quality of life; mental health; adolescents; chronic disease; mhealth; participatory design

Introduction

Asthma Prevalence and Psychological Comorbidities

Asthma is considered the 14th most important disorder in the world in terms of the extent and duration of disability, affecting over 300 million people [1]. Furthermore, people with asthma are significantly more likely to experience mental illness, for example, being six times more likely to develop anxiety disorder [2].

Young people (aged 12-24 years) with asthma are also significantly more likely to suffer psychological dysfunction [3], including greater levels of depression, anxiety, risk-taking behavior, and social isolation as compared with their peers without asthma and also as compared with people with asthma of other age groups [1,3]. Adolescence is a complex period including issues of puberty, personal identity, sexuality, and vocational decision-making and having a chronic illness adds to the psychosocial burden of adolescence, for example, through the requirement to take on more responsibility for asthma self-management and dealing with gradual transition from pediatric to adult care [4].

On the basis of the severity of asthma and its psychological comorbidities, and in light of survey data showing a desire among young people for mobile-based asthma support, Blanchard et al have called on researchers to “develop tailored Web-based and mobile apps to support the wellbeing of young people with asthma, targeting asthma control, mental health, and general wellbeing side-by-side.” They also report that 83% of young people with asthma surveyed reported using a mobile phone “every day or almost every day.” Nevertheless, no asthma app in the existing research literature has attempted to address any of the psychological comorbidities associated with the illness [5-7].

Participatory Design as a Method for User-Involvement

In this paper, we describe results from the participatory design phase of an app intended to improve the well-being of young people with asthma. We refer to young people aged between 10 and 24 years, which include adolescents aged between 10 and 19 years [8].

A participatory design approach was selected to overcome the typical lack of user-involvement reported in the development of existing health and asthma apps [9,10]. The need for end-user involvement from the outset is now acknowledged as critical to a successful health app development process [11-13].

Participatory design as a methodology, and general user-centered orientation, provides methods for the direct involvement of end-users in the codesign of technologies. It is distinct from other forms of user-centered design in that it positions designers

as facilitators and views users as active cocreators of the solution [14]. Participatory design has been used as a method for involving patients or end-users in the design of health technologies and services in order to develop tools that are more likely to be engaging, relevant, and effective [15,16]. The approach has also become increasingly popular when involving adolescents in technology design [17-19].

Self-Determination Theory as a Model for Well-Being and Technology Engagement

Self-determination theory (SDT) is a psychological model of motivation that has been successfully applied to understand and predict psychological well-being [20,21] specifically with respect to the satisfaction of three basic psychological needs: *autonomy* (to experience volition; act in accord with one's interests and values), *competence* (to feel effective), and *relatedness* (to interact with, feel connected to, and care for others) [22]. Health intervention research utilizing SDT has shown that the more support for autonomy an individual has for their health, the more likely they are to achieve improved health outcomes, such as symptom control and quality of life [23]. The SDT model has also been shown to explain and predict engagement with technologies [24,25]. Given the usefulness of SDT to both supporting health outcomes and engagement in technology design, it serves an ideal foundation to a “positive computing” [26] approach (a topic of this special issue) to designing health apps that support well-being as well as health outcomes.

Ryan and Deci [27] also provide characteristics of environments (ie, designable features) that can be considered supportive or hindering to autonomy, competence, and relatedness (ie, goal choice, positive feedback, informational rewards), thus providing some initial practical guidance to designers. As such, SDT provides a uniquely useful framework for both the qualitative analysis of psychological experience and for translating insights from this analysis into actionable design specifications for technologies.

Autonomy is of special importance to the health community. It has not only been implicated as critical to mental health [27,28] but also is now considered critical to doctor-patient relationships and to health organizations. Taking autonomy-supportive approaches to patient interaction (eg, enabling patients to make informed decisions, active listening, taking a patient's personal circumstances into account) is now considered a guiding principal of biomedical ethics [27,29]. Given its importance for doctor-patient relationships and to health organizations, it stands to reason that health *technologies* should also be designed in autonomy supportive ways. This paper contributes to a growing body of literature [30] on how human autonomy can be supported through technology design, in this case through the

design of an app which aims to support the psychological well-being of young people with asthma.

In this paper, we aim to help address the gap in psychological support for asthma by (1) presenting qualitative results of the psychological experience of asthma as reported by young people, and (2) providing examples of how an app for asthma can be designed to support psychological factors.

Methods

Recruitment

Using convenience sampling, young people aged 15-24 years were invited to participate in the study via communications sent to social media, university websites, community pharmacies, sporting clubs, high school nurses, asthma educators, and organizations providing services for young people located in the metropolitan areas of Sydney and other parts of Australia. Eligibility criteria for the study included age (15-24 years), residency in Australia, and doctor-diagnosed asthma (self-reported). Level of asthma control was determined via an Asthma Control Questionnaire (ACQ) given after initial recruitment. Participants were offered a Aus \$50 store voucher as compensation for their time. This research was approved by the Research Ethics Committee at the University of Sydney.

Participatory Design Workshops and Workbooks

Activity-based workshops (2.5 h each) were carried out with young people with asthma. During each workshop, a user experience specialist-guided participants through a series of activities designed to elicit (1) emotions, perceptions, and needs related to being a young person with asthma and (2) design ideas and preferences for an app to support young people with asthma in a relevant and engaging way.

Activities were selected in order to fulfill a series of user research questions defined by the research team at the start of the project. Activities were drawn from the literature in participatory design [31,32] and included collaborative collage, individual concept mapping, group ideation, and paper prototyping. Data collected included participant-generated artifacts from each of the activities (ie, collages, concept maps, and screen designs), field notes, and audio recordings with transcriptions of each workshop.

Workbooks were also received from young people with asthma. The workbook was originally intended as a sensitization tool [14] to be completed by participants before the workshop. However, as the difficulty in recruiting participants to workshops became apparent, the workbook was also repurposed as an alternative method of participation for individuals who could not attend workshops and allowed the researchers to collect data from a broader sample by including those outside the state.

Workbook data included information on commonly used apps, and perceptions and experiences to do with health.

Issues to do with mental health and psychological factors were not prompted in either the workshops or workbooks but emerged independently.

Questionnaires

Each participant completed a questionnaire that included questions on demographic characteristics and the 6-item Asthma Control Questionnaire (ACQ), a standard validated scale to measure the range of asthma control among participants (scored 0 [best] and 6 [worst]; ≤ 0.75 : well-controlled asthma; ≥ 1.50 : inadequately controlled asthma [33]).

User Feedback on a Prototype

Consistent with participatory design methods, user evaluation was conducted with nine of the original workshop participants on a high-fidelity prototype after the initial design phase. These user testing sessions were audiorecorded for later analysis.

The app's content was also reviewed for accuracy of medical content by a team of clinicians.

Crowdsourcing

In order to allow young people with asthma to name the app, an invitation was posted to the Young People with Asthma Facebook community inviting submissions for an asthma app name. The research team then selected from the options contributed (based on distinctiveness) and tested acceptability of the name during the user testing phase.

Data Analysis

Using theoretical thematic analysis [34] and consistent with methods for the analysis of generative participatory data [14], the text and imagery from the artifacts and workbooks were analyzed by a member of the research team and categorized into themes as detailed below.

Although the user research was exploratory and open in nature, it was conducted toward a specific outcome (the development of an effective app) and therefore analysis was guided by the research questions in support of that outcome. The theoretical thematic analysis focused on those categories of data that would inform development, namely: (1) categories of features and characteristics of an ideal app, and (2) categories of lived experience that could be addressed via technology. For the former, categories were straightforward feature types, for example, a category "social features" included "chat," "forum," and "story sharing," whereas a category "Feedback" included "encouraging messages," "evidence of achievement," and "evidence of progress."

For the more diversely articulated and abstract descriptors of lived experience, data was categorized using SDT as a framework in order to link elements of experience as reported by participants to evidence-based well-being constructs, thus facilitating the translation of findings to relevant interaction designs (as mentioned earlier, SDT literature provides some links from well-being constructs to technology design features). A similar application of SDT to content analysis has been used in the context of smoking [35]. For the collage activity involving imagery, participants were asked to describe the images they chose and these descriptive words were included in analysis.

Results

Demographics

A total of 48 young people contacted the research team expressing an interest in participation; of which 28 chose not to participate due to various factors such as geographical location, school, or work schedules and 20 young people (mean

age 17.8 years [range 15-24]; 12 female) with a range of asthma control (ACQ: 1.5 [0.2-4.5] and 12 prescribed a preventer inhaler) completed a workbook and 13 of these also took part in a workshop (workshops consisted of between 2 and 7 participants) between May 2015 and November 2015. Of the 20, 7 only completed a workbook, whereas 13 completed both a workshop and workbook. More demographic detail is given in [Table 1](#).

Table 1. Participant demographics.

Participants	N=20
Age, years, mean (range)	17.8 (15-24)
Female, n (%)	12 (60)
Occupational status, n (%)	
High school student	12 (60)
University student	6 (30)
Working	1 (5)
Socioeconomic indexes (median; range) ^a	5; 1-5
Asthma Control Questionnaire, mean (range) ^b	1.5 (0.2-4.5)

^aSocial disadvantage at home address: "Disadvantaged" SEIFA (Socioeconomic Indexes for Areas) Quintile ≤3, "Advantaged" SEIFA Quintile: 4-5 [36].

^bScored 0 (best) and 6 (worst); ≤0.75: well-controlled asthma; ≥1.50: inadequately controlled asthma [33].

A total of 12 submissions were received for the crowdsourced app naming campaign. The name "Kiss My Asthma" was selected from among the 12 by the research team, in consultation with Asthma Australia (national asthma charity and research funders) and based on criteria, such as distinctiveness from other app names and appropriateness to the goals of the app.

Qualitative analysis of 102 participant-generated artifacts (ie, collages, concept maps, and screen designs) yielded a set of user needs, goals, obstacles, and supports with respect to the lived experience of being young with asthma, as well as a list of preferred features and characteristics for an asthma app from the perspective of young people. Results showed clear evidence of a need to support psychological well-being for those with asthma, along with some specific ideas for how to do so within the context of an app, as detailed below.

The Experience of Being Young With Asthma

Workshop participants collaboratively created collages to represent "what it's like to be a young person with asthma." [Figure 1](#) shows combined imagery from workshop participant collages and includes images depicting asthma as a "battle;" asthma as a source of stress, anxiety, and social exclusion; the need to put on a healthy façade; and asthma as a barrier to living like a "normal teenager." The resulting collages and discussion overwhelmingly focused on the psychological experience of living with asthma (something unprompted and unanticipated by the researchers) and revealed key themes consistent with the basic psychological needs described by SDT. Examples of collage imagery depicting emotional experience as well as illustrative quotations are included below.

Figure 1. Combined imagery from workshop participant collages.

Autonomy: Lack of Control, Restrictions, and Uncertainty

Participants expressed feelings of lack of autonomy in the form of frustration and anxiety over lack of control, the inability to take part in activities, and uncertainty...

You never know what will happen, cause you never know when you'll have an asthma attack. [female, 18 years]

When it does happen, we become subject to the asthma attack and can't really control our actions or defend ourselves. [male, 16 years]

If you forget your (rescue) meds, you're panicking the whole time. [female, 18 years]

You're in a bit of a bubble with what you can do, eat and touch. [female, 19 years]

We want to be a bit more Zen, a bit more carefree (because having asthma is) pretty scary [female, 22 years]

Competence: Not Keeping Up, Struggling, Feeling Unprepared

Participants also expressed feelings of lack of competence with respect to physical competence, keeping up with peers, managing their health, and feeling unprepared for asthma emergencies...

We still do it, it's just a bit harder, it's not as easy to keep up. [female, 18 years]

I always struggled to keep up with others in training. [female, 18 years]

Asthma's holding us back. [male, 15 years]

Asthma is the guy that's punching you. [male, 17 years]

It's like you against asthma and it's always there. I have the words "health battle" because I feel that really captures it. [female, 19]

Relatedness: Being Left Out, Social Pretense, Isolation

Participants expressed lack of relatedness as feelings of isolation, distress at being left out, having to put on a healthy façade, and being different to peers.

You can't act your age because of asthma. [female, 18 years]

It's not as easy to keep up(so you) feel left out. [female, 18 years]

How we feel trying to navigate around growing up, cause it's different—even going to school each day—you have to worry about a lot more things than other kids, even through the day.[female, 22 years]

You pretend you're fine even when you're not. [female, 18 years]

Two different perspectives on people with asthma: the facade that we put on saying we're healthy and fine, and then the more chaotic side. [male, 16 years]

Life Goals, Obstacles, and Supports

The participants reported a number of life goals via the individual concept-mapping activity. This contributed insights into the intrinsic motivations young people have to manage their asthma. Participants mostly reported shorter-term goals with some notable big picture aspirations. The majority of goals could be clustered into one of four themes: (1) study, (2) job and career, (3) health, fitness, and sport, and (4) spirituality, happiness, and meaning.

Common obstacles to goal achievement reported included asthma and related illness (ie, allergies and anaphylaxis that are common comorbidities in people with asthma), lack of money, and lack of motivation. Common supports reported included family, friends, and personal traits such as determination (previous research has shown the importance of support to successful asthma management [37]). A list of these themes, along with illustrative quotations and how participants related goal obstacles to asthma (when they did), are included in Table 2. Tables 3 and 4 show the obstacles and supports participants reported, respectively.

Preferred Features

As part of a brainstorming activity, participants generated ideas for features that they believed would make the app practical,

useful, and engaging. These are presented in Table 5 in a categorized list beside the psychological need (SDT construct) most directly supported by each feature. *Autonomy* is indicated for features that provide the user with choices, or access to tools providing a greater sense of control. *Competence* was selected for features that provide users with knowledge or skills that increase their capacity to handle situations. *Relatedness* was selected where features connect the user to others or contribute to a sense of belonging. Of course, many features contribute to more than one psychological need. The list is not intended to be definitive, but rather to help make explicit how practical design features can contribute to psychological experience in health apps. Figure 2 shows example screen prototypes created by participants.

Table 2. Goals in life.

Goal theme	Examples of goals as written by participants ^a	Relationship of goal to asthma
Study	"Finish HSC ^b ," "School leadership," "Be a good student," "Get a defense scholarship"	Asthma sabotaged exams Missed class when in hospital for asthma Couldn't pass physical exam for defense scholarship
Job and career	"Become a professional footballer," "Become a manager at Maccas [sic]," "Become a doctor (ophthalmic surgeon)," "Get a job that I enjoy and love," "Become a child carer," "Be an archaeologist"	Career goal requires travel to countries with different asthma-inducing climates Exceptional fitness required to be a professional athlete
Health, fitness, and sport	"Be physically fit—able to play sport for a longer time without getting tired," "STAY ALIVE," "Get new hearing aids," "Get a good night's sleep," "National champion fencing and Olympics," "Be a good netball player"	Asthma limits athletic performance Asthma limits ability to maintain physical fitness Asthma interrupts sleep
Spirituality, happiness, and meaning	"Achieve happiness," "Live a long and happy life," "Becoming more like Jesus," "Caring for others in a genuine loving way," "Charity work"	No direct connections to asthma were reported for these goals

^aAll text quoted verbatim from the contributions made by participants during concept mapping.

^bHSC: Higher School Certificate.

Table 3. Obstacles in life.

Obstacle theme	Examples of obstacles as written by participants ^a
Asthma and health	"Asthma," "Exhaustion," "Bringing my Ventolin," "Emergency situations where I need to run," "Allergies"
Money and circumstances	"Money (lack thereof)," "Time (lack of)," "Constraints: (time-based, family-based, etc.)"
Lack of motivation	"Laziness," "Lack of motivation," "Not caring about certain [school] subjects"
Study and work	"Academic marks," "Experiments not working," "Inexperience"
Personal faults	"Not managing my time well," "Selfishness, envy"

^aAll text quoted verbatim from the contributions made by participants during concept mapping.

Table 4. Supports in life.

Support theme	Examples of supports as written by participants ^a
Other people	“Family or friends,” “Netball team,” “Coaches”
Medicine or health	“Medication,” “Puffer,” “Training my body to be fitter”
Circumstances	“My school,” “Australia,” “Opportunities”
Personal effort	“Working hard,” “Studying,” “Getting involved [in extra-curricular activities],” “Reminding myself to be proactive”
Personal traits	“Personal nature as a caring person,” “determination,” “perseverance”
Faith or spirituality	“Jesus”

^aAll text quoted verbatim from the contributions made by participants during concept mapping.

Although the features suggested by participants may address psychological needs, it is interesting to note that, with the exception of mood tracking, participants did not explicitly connect the psychological factors they expressed so strongly during the collage activity, to app feature ideas. This is probably because there is little precedent for providing psychological support among health apps, or apps for young people generally. However, when the development team added an explicit mental health feature (a “managing anxiety” section), participants during user testing expressed strong approval, requesting the section be expanded (detailed below).

Figure 2. Example screen designs showing: a) mood tracking b) personal story sharing.

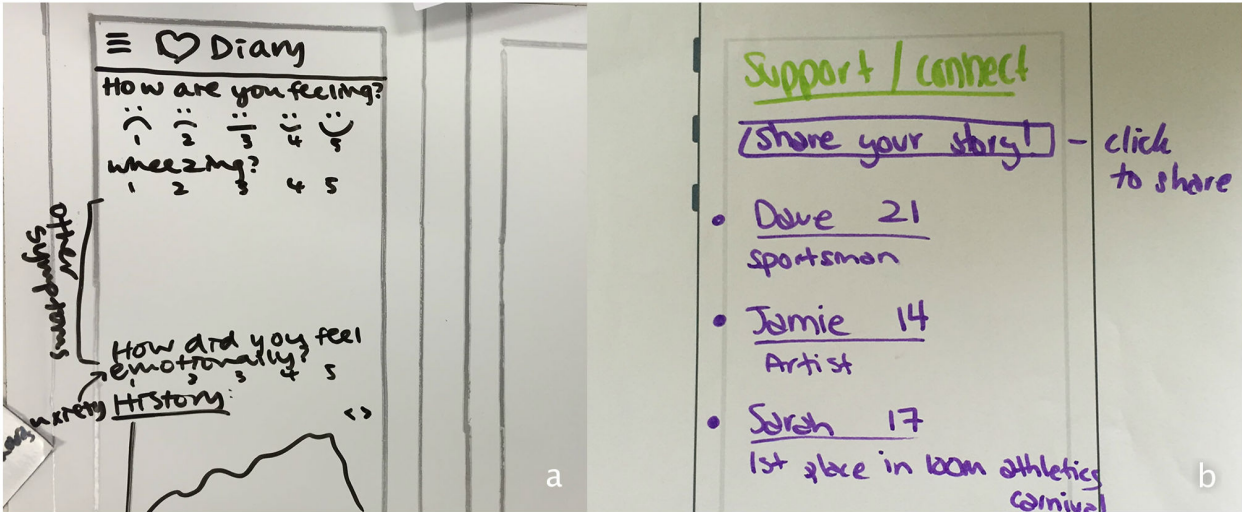


Table 5. List of suggested features linked to psychological needs.

Feature	Psychological need supported
Profile	
Medical history and treatment summary	Competence
Medication list	Competence
Customizable to the particular health characteristics or health needs of individual users	Competence
Reminders	
Medication alerts (eg, “It’s time for 2 puffs of preventer”)	Competence
Feature that allows users to find their inhaler “Find my puffer”	Autonomy or Competence
Location-based reminders (“alert! You are outside of your home. Do you have your puffer?”)	Autonomy or Competence
Tracking	
Symptom diary (wheezing scale, coughing, and so on)	Competence
Causes or triggers diary	Competence
Asthma attack diary (“log an attack”)	Competence
Mood diary	Competence
Health over time overview graph	Competence
Step count	Competence
Overall log feedback: “Your attacks have been more frequent lately, maybe you should check with your doctor”	Competence
Log reports to show your doctor	Autonomy or Competence
Social	
In-app chat with other people with asthma	Relatedness
Q&A forum	Competence
Share personal experiences with others	Relatedness
Share with friends (connect with friends to share achievements and encourage each other)	Relatedness
Share your story (inspiring stories from young people overcoming an asthma challenge)	Relatedness
Emergency support	
GPS ^a tool to find the nearest emergency service	Autonomy or Competence
“How to handle an asthma attack” screen to show others	Autonomy
Asthma attack management sequence for user	Autonomy or Competence
Emergency contacts	Autonomy or Competence
Off-line mode or not Internet-dependent (especially for emergency info)	Autonomy
Emergency info available from lock screen	Autonomy
Expert access and information	
Chat with a friendly professional	Competence
Automated chat (for off-hours)	Competence
Expert and how-to videos	Competence
List of specialists near you	Autonomy
Personalized asthma facts and statistics (based on your profile)	Competence
Information on how to prevent an attack	Competence
Feedback and motivational features	

Feature	Psychological need supported
Support messages, badges, points for good behavior (“you’ve been doing well—go reward yourself!”)	Competence
Daily motivations (quotes)	Competence
Competing with friends or anonymous strangers	Relatedness

^aGPS: Global Positioning System.

User Feedback on a Prototype: “More Support for Mental Health”

User testing of an app prototype based on the data gathered at the user workshops occurred 8 months after the initial codesign phase (the gap between phases was a result of the challenges of recruiting minors including ethics committee approval, parental consent, and logistics), with nine of the original workshop participants. Feedback from the user testing was highly positive with all participants reporting that they would recommend the app to a friend. Moreover, a number of testers independently expressed approval of the mental health support features in the app, in particular, the information section on managing anxiety, and went further to suggest the section be expanded with practical exercises and tools for self-management. For example, one female participant (aged 22 years) stated during an individual user test...

I think anxiety is quite a big thing when you have asthma...There’s not a lot of information out there about how to deal with it. As a child, when I had bad asthma I had to attend the hyperventilation clinic to teach me how to manage anxiety...So if there was a bit more information about it...even basic exercises that you could try, to keep your asthma, or even your anxiety, under control.

When you’re dealing with asthma and then it’s like “go speak with your doctor;” “go do this and that.” But if it’s small things you could learn to do yourself I think that probably is a good thing. Because anxiety is not a thing we spoke about with asthma especially when I was young. It’s becoming more understood now, but it’s actually quite a big problem. Anxiety in general is a big thing that we don’t often look at.

A second female participant (aged 19 years) provided similar advice:

Perhaps, on the managing anxiety thing, add other options. For example, something you enjoy, have a bath or listen to some music, do a craft. Something other than “talk to your doctor,” etc. All of those are great, but not when you’re having an asthma attack.

Finally, one male participant’s (aged 16 years) reference to stress provides evidence for how even esthetic design decisions can link to well-being factors:

Stress has been proven to exacerbate asthma and this relaxes me with its colorsthe blue and the cloud.

Discussion

Principal Findings

This participatory study revealed that the lived experience of young people with asthma includes a number of psychological factors that can—and should, according to participants—be supported via an asthma app.

Although the participatory approach posed challenges from a recruitment perspective, its benefits became clear when factors relating to psychological experience emerged. It was through generative participatory practices (in particular, collaborative collage and concept mapping) that difficult-to-articulate matters of psychological experience came to light. In other words, had user involvement not been facilitated through a participatory approach, we arguably would not have come to understand the importance to young asthma sufferers of psychological support needs, and their desire for these to be addressed via technology.

Indeed, the name “Kiss My Asthma,” which was elicited from young people with asthma, reflects a highly autonomous, if irreverent, attitude toward asthma as an adversary (a notion expressed by participants in workshops, eg, “Asthma is the guy that’s punching you in the face”). It also demonstrates the importance of user input when developing autonomy supportive apps since it seems unlikely that any researcher or clinician would have named the app in such a way.

The interest in psychological support features, including anxiety management and mood tracking, may result from the lack of attention given to mental health comorbidities in adolescents with asthma. Although it may be especially pertinent for adolescents with asthma due to the anxiety that breathing difficulties can engender, the desire for technology-based mental health support may also have important implications for adolescents with other chronic illnesses and requires further research.

The high user satisfaction feedback received during prototype testing provides early evidence for the success of the participatory methods, and support for the consideration of psychological factors in the app.

Translating Findings on Psychological Factors to Design Features

The psychological factors revealed during participatory design workshops required further analysis to translate into actionable design specifications. The literature in positive computing provided an approach to designing technology to support psychological well-being that includes structuring ideation around individual well-being determinants [26]. Mental health

factors reported in this study (eg, anxiety, panic, social disconnectedness) were clustered as impairments to autonomy, competence, and relatedness as consistent with SDT. SDT was useful both for categorizing user needs, as well as for linking these needs to evidence-based well-being determinants that could become the focus of design.

The research team, which includes two psychologists and an interaction and user interface designer, further developed the

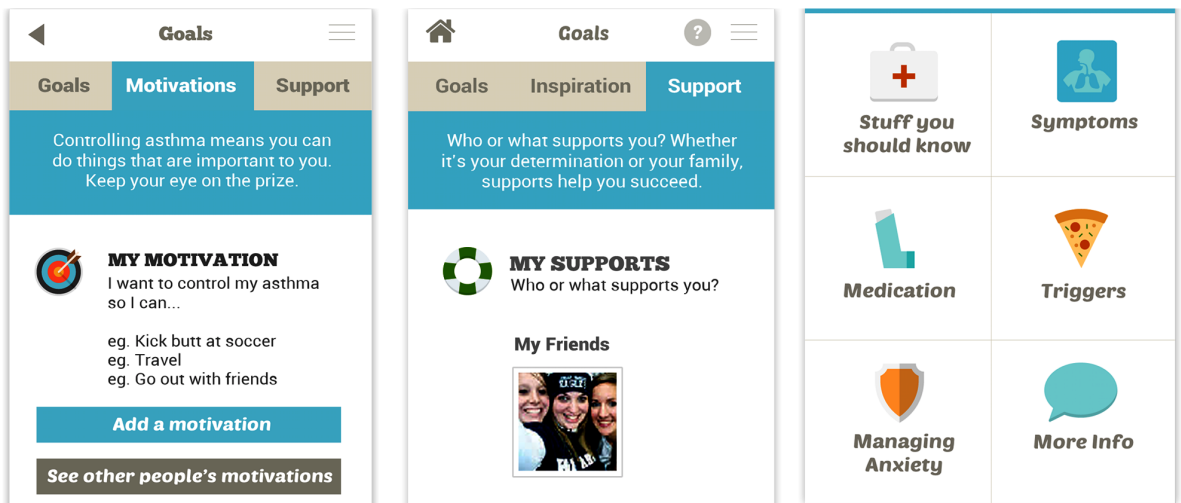
psychological support features for the needs for autonomy, competence, and relatedness expressed by the young people (Table 6). This example not only shows how the use of a motivational model and user-experience methods can work in synergy to guide design for health technologies, but also functions as an example of positive computing in which ideation is based on well-being determinants (in this case those identified by SDT). Figure 3 shows three example screens from the prototype app, which relate to these supportive design features.

Table 6. Psychological support features.

Psychological factor	Supportive design features
Relatedness (Make salient user belongingness to a larger community while maintaining privacy and safety)	Convey the message “you are not alone” by including text on the prevalence of mental health issues among young people with asthma. Provide users with a place to upload images of those who support them. Allow optional anonymous sharing of user motivations for asthma management. Allow users to browse a gallery of other users’ shared motivations (<i>these are currently moderated for safety</i>). Provide the app dialogs as if they were coming from a supportive friend (delivered by characters rather than by “the phone”).
Autonomy and competence (Provide tools that will allow users to feel more in control and better able to handle asthma impacts and emergencies; support autonomous motivation, based on personal values and goals [38])	Provide strategies and tips for better asthma control. Scaffold goal-setting to support agency and improve asthma control. Scaffold the identification of autonomous motivations for managing asthma (by allowing users to record and share their own autonomous motivations and to view others). Provide an information section on managing anxiety and panic attacks.
General support for mental health	Provide a calming esthetic (ie, light color palette, cloud metaphor, “Cute monster” characters). Incorporate subtle humor as part of dialogs to lighten the otherwise stressful context of asthma control. Provide links to mental health support lines and resources for young people.

The impact of practical features (eg, symptom tracking) on a sense of autonomy is further supported by recent research into using asthma app features to promote adolescent self-management through autonomy [39].

Figure 3. Example app screens: (left) “My Motivations,” (middle) “My Supports,” and (right) “Managing Anxiety” in “About Asthma”.



Limitations and Future Work

The participant cohort, while representative with respect to age, gender, and asthma control, was limited with respect to social

disadvantage (most participants lived in a socially advantaged location and were urban or suburban students of either a university or private high school). Further studies targeting other demographics would contribute to generalizability. Our study

volunteers may have been more engaged with their disease than the wider population; their preferences may not fully represent individuals less engaged with asthma.

A number of constraints intrinsic to app development projects, including limited access to developers and a limited time frame, meant that the data could only be analyzed by one researcher, and then reviewed by a team of researchers, before being implemented for development. Although this approach allowed research insights to inform design and development, additional analysts would provide greater reliability to the findings. Some confirmation of reliability was provided by the high satisfaction rates reported during prototype testing, but owing to the small sample size, further evaluation would strengthen these findings as well.

In a next phase of the project, we will evaluate the impact of app features on quality of life (and against measures of autonomy, competence, and relatedness) as well as on asthma control outcomes as part of a clinical trial. Contingent on positive outcomes from the clinical trial, we will release the app to the public on World Asthma Day, May 2, 2017. It is hoped that following public release, having an access to usage data from a significantly larger audience for the app will allow us to conduct a larger scale quantitative study on the use and impact of various features, including those that support intrinsic motivation, mental health, and well-being.

Comparison With Prior Work

Cushing et al [40], Roberts et al [41], and Burbank et al [42], recently reported on the use of mHealth for asthma support but each with specific foci, namely, a sensor device for adherence-support, communication with friends and caregivers, and asthma action plans, respectively. None of these employed participatory codesign methods. Furthermore, there are no reports of asthma apps that have taken into consideration the psychological factors associated with the disease.

Conclusions

In this study, participatory design methods were used to (1) elicit descriptions of the lived experience from young people with asthma, (2) identify psychological support features and characteristics for an app for asthma from the perspective of young people, and (3) provide young people with the opportunity to guide and cocreate an app to support their asthma self-management and well-being. The resulting prototype was highly rated during user feedback.

In addition to informing the development of asthma-support technologies, this research provides a practical example of how health technologies in general can be designed to support autonomy, competence, and relatedness. Moreover, our finding that there is a desire for technology-based mental health support among young people with asthma could have implications for adolescents with other chronic illnesses.

Acknowledgments

Funding was provided by a competitive research grant from Asthma Australia. Professor Calvo is supported by an Australian Research Council Future Fellowship.

Conflicts of Interest

In the last 3 years the Woolcock Institute of Medical Research has received independent research funding from AstraZeneca and GlaxoSmithKline for asthma research carried out by Dr JM Foster.

References

1. Blanchard M, Morris J, Birrell E, Stephens-Reicher J, Third A, Burns J. National Young People and Asthma Survey: Issues and Opportunities in Supporting the Health and Wellbeing of Young People Living with Asthma. Melbourne: Young and Well Cooperative Research Centre; 2014.
2. Thomas M, Bruton A, Moffat M, Cleland J. Asthma and psychological dysfunction. *Prim Care Respir J* 2011 Sep;20(3):250-256 [FREE Full text] [doi: [10.4104/pcrj.2011.00058](https://doi.org/10.4104/pcrj.2011.00058)] [Medline: [21674122](https://pubmed.ncbi.nlm.nih.gov/21674122/)]
3. Slade T, Johnston A, Teesson M, Whiteford H, Burgess P, Pirkis J, et al. Health. Canberra: Australian Government; 2009. The Mental Health of Australians 2: Report on the 2007 National Survey of Mental Health and Wellbeing URL: <http://www.health.gov.au/internet/main/publishing.nsf/content/mental-pubs-m-mhaust2> [accessed 2017-04-03] [WebCite Cache ID [6pScvXmZh](https://www.webcitation.org/6pScvXmZh)]
4. Turkel S, Pao M. Late consequences of chronic pediatric illness. *Psychiatr Clin North Am* 2007 Dec;30(4):819-835 [FREE Full text] [doi: [10.1016/j.psc.2007.07.009](https://doi.org/10.1016/j.psc.2007.07.009)] [Medline: [17938047](https://pubmed.ncbi.nlm.nih.gov/17938047/)]
5. 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;13:58 [FREE Full text] [doi: [10.1186/s12916-015-0303-x](https://doi.org/10.1186/s12916-015-0303-x)] [Medline: [25857569](https://pubmed.ncbi.nlm.nih.gov/25857569/)]
6. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. *BMC Med* 2012;10:144 [FREE Full text] [doi: [10.1186/1741-7015-10-144](https://doi.org/10.1186/1741-7015-10-144)] [Medline: [23171675](https://pubmed.ncbi.nlm.nih.gov/23171675/)]
7. Marcano Belisario JS, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. *Cochrane Database Syst Rev* 2013;11:CD010013. [doi: [10.1002/14651858.CD010013.pub2](https://doi.org/10.1002/14651858.CD010013.pub2)] [Medline: [24282112](https://pubmed.ncbi.nlm.nih.gov/24282112/)]

8. World Health Organization. WHO. 2001. The Second Decade: Improving Adolescent Health and Development URL: http://www.who.int/maternal_child_adolescent/documents/frh_adh_98_18/en/ [accessed 2017-04-03] [WebCite Cache ID 6pRS1g3BF]
9. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
10. Harrison AM, Goozee R. Psych-related iPhone apps. *J Ment Health* 2014 Feb;23(1):48-50. [doi: [10.3109/09638237.2013.869575](https://doi.org/10.3109/09638237.2013.869575)] [Medline: [24484194](https://pubmed.ncbi.nlm.nih.gov/24484194/)]
11. Cho MJ, Sim JL, Hwang SY. Development of smartphone educational application for patients with coronary artery disease. *Healthc Inform Res* 2014 Apr;20(2):117-124 [FREE Full text] [doi: [10.4258/hir.2014.20.2.117](https://doi.org/10.4258/hir.2014.20.2.117)] [Medline: [24872910](https://pubmed.ncbi.nlm.nih.gov/24872910/)]
12. Dunton GF, Dzibur E, Kawabata K, Yanez B, Bo B, Intille S. Development of a smartphone application to measure physical activity using sensor-assisted self-report. *Front Public Health* 2014;2:12 [FREE Full text] [doi: [10.3389/fpubh.2014.00012](https://doi.org/10.3389/fpubh.2014.00012)] [Medline: [24616888](https://pubmed.ncbi.nlm.nih.gov/24616888/)]
13. Jibb LA, Stevens BJ, Nathan PC, Seto E, Cafazzo JA, Stinson JN. A smartphone-based pain management app for adolescents with cancer: establishing system requirements and a pain care algorithm based on literature review, interviews, and consensus. *JMIR Res Protoc* 2014;3(1):e15 [FREE Full text] [doi: [10.2196/resprot.3041](https://doi.org/10.2196/resprot.3041)] [Medline: [24646454](https://pubmed.ncbi.nlm.nih.gov/24646454/)]
14. Sanders L, Stappers PJ. *Convivial Toolbox: Generative research for the front end of design*. Amsterdam: BIS; 2013.
15. Clemensen J, Larsen SB, Kyng M, Kirkevold M. Participatory design in health sciences: Using cooperative experimental methods in developing health services and computer technology. *Qual Health Res* 2007 Jan;17(1):122-130. [doi: [10.1177/1049732306293664](https://doi.org/10.1177/1049732306293664)] [Medline: [17170250](https://pubmed.ncbi.nlm.nih.gov/17170250/)]
16. Grönvall E, Kyng M. On participatory design of home-based healthcare. *Cognition, Technology & Work* 2013;15(4):389-401.
17. Wadley G, Lederman R, Gleeson J, Alvarez-Jimenez M. Participatory design of an online therapy for youth mental health. New York, NY, USA: ACM Press; 2013 Presented at: 25th Australian Computer-Human Interaction Conference on Augmentation, Application, Innovation, Collaboration, OzCHI 2013; November 25 - 29, 2013; Adelaide p. 517-526. [doi: [10.1145/2541016.2541030](https://doi.org/10.1145/2541016.2541030)]
18. Nicholas M, Hagen P, Rahilly K, Swainston N. Using participatory design methods to engage the uninterested. In: *Proceedings of the th Participatory Design Conference on Exploratory Papers Workshop Descriptions Industry Cases, Volume 2*. New York, NY, USA: ACM Press; 2012 Presented at: The 12th Participatory Design Conference; 12-16 August, 2012; Roskilde, Denmark p. 121-124.
19. Miller A, Pater J, Mynatt E. Design Strategies for Youth-Focused Pervasive Social Health Games. New York: IEEE; 2013 Presented at: 7th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) 2013; 5-8 May, 2013; Venice, Italy p. 9-16. [doi: [10.4108/icst.pervasivehealth.2013.252081](https://doi.org/10.4108/icst.pervasivehealth.2013.252081)]
20. 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](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
21. Ryan RM, Deci EL. On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annu Rev Psychol* 2001;52:141-166. [doi: [10.1146/annurev.psych.52.1.141](https://doi.org/10.1146/annurev.psych.52.1.141)] [Medline: [11148302](https://pubmed.ncbi.nlm.nih.gov/11148302/)]
22. Deci E, Vansteenkiste M. Self-determination theory and basic need satisfaction: understanding human development in positive psychology. *Ric di Psicol* 2004;27(1):23-40.
23. Williams GC, McGregor HA, Zeldman A, Freedman ZR, Deci EL. Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. *Health Psychol* 2004 Jan;23(1):58-66. [doi: [10.1037/0278-6133.23.1.58](https://doi.org/10.1037/0278-6133.23.1.58)] [Medline: [14756604](https://pubmed.ncbi.nlm.nih.gov/14756604/)]
24. Ryan R, Rigby C, Przybylski A. The motivational pull of video games: a self-determination theory approach. *Motiv Emot* 2006:344-360.
25. Rigby S, Ryan R. *Glued to Games: How Video Games Draw Us In and Hold Us Spellbound (New Directions in Media)*. Santa Barbara: Praeger; 2011.
26. Calvo R, Peters D. *Positive Computing: Technology for Wellbeing and Human Potential*. Cambridge: The MIT Press; 2014:9063692846.
27. Ryan RM. In: Deci EL, editor. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. 1st ed. New York: The Guilford Press; 2017.
28. Ryff CD, Keyes CL. The structure of psychological well-being revisited. *J Pers Soc Psychol* 1995 Oct;69(4):719-727. [Medline: [7473027](https://pubmed.ncbi.nlm.nih.gov/7473027/)]
29. Entwistle VA, Carter SM, Cribb A, McCaffery K. Supporting patient autonomy: the importance of clinician-patient relationships. *J Gen Intern Med* 2010 Jul;25(7):741-745 [FREE Full text] [doi: [10.1007/s11606-010-1292-2](https://doi.org/10.1007/s11606-010-1292-2)] [Medline: [20213206](https://pubmed.ncbi.nlm.nih.gov/20213206/)]
30. Calvo R, Peters D, Johnson D, Rogers Y. Autonomy in Technology Design. In: *CHI2014 Extended Abstracts on Human Factors in Computing Systems*. New York: ACM; 2014 Presented at: The ACM CHI Conference on Human Factors in Computing Systems; April 26-May 1; Toronto.
31. Muller MJ. Participatory design: the third space in HCI. In: Sears A, Jacko JA, editors. *Human-Computer Interaction: Development Process*. Boca Raton, FL: CRC Press; 2003:165-185.

32. Gray D, Brown S, Macanujo J. *Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers*. Sebastopol, CA: O'Reilly Media; 2010.
33. Juniper EF, Bousquet J, Abetz L, Bateman ED. Identifying 'well-controlled' and 'not well-controlled' asthma using the Asthma Control Questionnaire. *Respir Med* 2006 Apr;100(4):616-621 [FREE Full text] [doi: [10.1016/j.rmed.2005.08.012](https://doi.org/10.1016/j.rmed.2005.08.012)] [Medline: [16226443](https://pubmed.ncbi.nlm.nih.gov/16226443/)]
34. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
35. Choi J, Noh G, Park D. Smoking cessation apps for smartphones: content analysis with the self-determination theory. *J Med Internet Res* 2014;16(2):e44 [FREE Full text] [doi: [10.2196/jmir.3061](https://doi.org/10.2196/jmir.3061)] [Medline: [24521881](https://pubmed.ncbi.nlm.nih.gov/24521881/)]
36. Australian Bureau of Statistics. ABS. Canberra: Australian Bureau of Statistics; 2006. Socio-economic indexes for areas URL: <http://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0.55.001> [accessed 2017-04-03] [WebCite Cache ID 6pSb85asQ]
37. Kyngas H, Rissanen M. Support as a crucial predictor of good compliance of adolescents with a chronic disease. *J Clin Nurs* 2001 Nov;10(6):767-774. [Medline: [11822848](https://pubmed.ncbi.nlm.nih.gov/11822848/)]
38. Ryan R, Deci E. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp Educ Psychol* 2000 Jan;25(1):54-67. [doi: [10.1006/ceps.1999.1020](https://doi.org/10.1006/ceps.1999.1020)] [Medline: [10620381](https://pubmed.ncbi.nlm.nih.gov/10620381/)]
39. Carpenter DM, Geryk LL, Sage A, Arrindell C, Sleath BL. Exploring the theoretical pathways through which asthma app features can promote adolescent self-management. *Transl Behav Med* 2016 Dec;6(4):509-518. [doi: [10.1007/s13142-016-0402-z](https://doi.org/10.1007/s13142-016-0402-z)] [Medline: [27118115](https://pubmed.ncbi.nlm.nih.gov/27118115/)]
40. Cushing A, Manice MP, Ting A, Parides MK. Feasibility of a novel mHealth management system to capture and improve medication adherence among adolescents with asthma. *Patient Prefer Adherence* 2016;10:2271-2275 [FREE Full text] [doi: [10.2147/PPA.S115713](https://doi.org/10.2147/PPA.S115713)] [Medline: [27853357](https://pubmed.ncbi.nlm.nih.gov/27853357/)]
41. Roberts CA, Geryk LL, Sage AJ, Sleath BL, Tate DF, Carpenter DM. Adolescent, caregiver, and friend preferences for integrating social support and communication features into an asthma self-management app. *J Asthma* 2016 Nov;53(9):948-954. [doi: [10.3109/02770903.2016.1171339](https://doi.org/10.3109/02770903.2016.1171339)] [Medline: [27116540](https://pubmed.ncbi.nlm.nih.gov/27116540/)]
42. Burbank AJ, Lewis SD, Hewes M, Schellhase DE, Rettiganti M, Hall-Barrow J, et al. Mobile-based asthma action plans for adolescents. *J Asthma* 2015;52(6):583-586 [FREE Full text] [doi: [10.3109/02770903.2014.995307](https://doi.org/10.3109/02770903.2014.995307)] [Medline: [25494553](https://pubmed.ncbi.nlm.nih.gov/25494553/)]

Abbreviations

ACQ: Asthma Control Questionnaire
SDT: self-determination theory

Edited by K Dinakar; submitted 15.11.16; peer-reviewed by D Carpenter, B Ainsworth; comments to author 22.01.17; revised version received 09.03.17; accepted 12.03.17; published 11.04.17.

Please cite as:

Peters D, Davis S, Calvo RA, Sawyer SM, Smith L, Foster JM
Young People's Preferences for an Asthma Self-Management App Highlight Psychological Needs: A Participatory Study
J Med Internet Res 2017;19(4):e113
URL: <http://www.jmir.org/2017/4/e113/>
doi: [10.2196/jmir.6994](https://doi.org/10.2196/jmir.6994)
PMID: [28400353](https://pubmed.ncbi.nlm.nih.gov/28400353/)

©Dorian Peters, Sharon Davis, Rafael Alejandro Calvo, Susan M Sawyer, Lorraine Smith, Juliet M Foster. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 11.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Corrigenda and Addenda

Authorship Correction: Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles

Sohrab Saeb¹, PhD; Thaddeus R Cybulski², B Sc; Stephen M Schueller¹, PhD; Konrad P Kording², PhD; David C Mohr¹, MD, PhD

¹Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Northwestern University, Chicago, IL, United States

²Rehabilitation Institute of Chicago, Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL, United States

Corresponding Author:

Sohrab Saeb, PhD

Center for Behavioral Intervention Technologies

Department of Preventive Medicine

Northwestern University

Rubloff Bldg, 10th Fl.

750 N Lake Shore Dr.

Chicago, IL, 60611

United States

Phone: 1 312 503 4624

Fax: 1 312 908 9588

Email: s-saeb@northwestern.edu

Related Article:

Correction of: <http://www.jmir.org/2017/4/e118/>

(*J Med Internet Res* 2017;19(4):e143) doi:[10.2196/jmir.7932](https://doi.org/10.2196/jmir.7932)

The authors of the paper entitled “Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles” [*J Med Internet Res* 2017;19(4):e118] inadvertently omitted Stephen M Schueller, PhD (Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Northwestern University) from the list of authors. The author Stephen M Schueller should have been added after Thaddeus R Cybulsky in the published manuscript. Stephen M Schueller was not on the initial authors list at the time of the first submission, and was added during the revising of the article,

when he contributed to the writing of the revised manuscript. His name already appears in the authors' funding acknowledgements section in the published paper, “Author SMS was supported by the research grant K08MH102336 from the National Institute of Mental Health.” This omission has been corrected in the online version of the paper on the JMIR website on April 28, 2017, together with the publication of this correction notice. Because this was made after submission to PubMed, the correction notice has been submitted to PubMed. The corrected metadata have also been resubmitted to CrossRef.

Multimedia Appendix 1

Letter from authors requesting correction.

[[PDF File \(Adobe PDF File\), 288KB - jmir_v19i4e143_app1.pdf](#)]

Edited by G Eysenbach; submitted 27.04.17; this is a non-peer-reviewed article; accepted 27.04.17; published 28.04.17.

Please cite as:

Saeb S, Cybulski TR, Schueller SM, Kording KP, Mohr DC

Authorship Correction: Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles

J Med Internet Res 2017;19(4):e143

URL: <http://www.jmir.org/2017/4/e143/>

doi: [10.2196/jmir.7932](https://doi.org/10.2196/jmir.7932)

PMID: [30578218](https://pubmed.ncbi.nlm.nih.gov/30578218/)

©Sohrab Saeb, Thaddeus R Cybulski, Stephen M Schueller, Konrad P Kording, David C Mohr. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.04.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), 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.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>