Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 20 (2018), Issue 12 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Reviews

Original Papers

Viewpoint

Corrigenda and Addenda

Review

Theory-Based Digital Interventions to Improve Asthma Self-Management Outcomes: Systematic Review

Helen J Lycett[1*], BSc, MSc; Eva M Raebel[1*], BSc, MSc, PhD; Emilie K Wildman[1*], BSc, MSc; Jordi Guitart[1*], PhD; Thomas Kenny[1*], BM, MSc-PH(HSM), MFPH, MBA; Jon-Paul Sherlock[2*], BEng, PhD; Vanessa Cooper[3*], BSc, MSc, PhD

[1]Spoonful of Sugar Ltd, London, United Kingdom

[2]Pharmaceutical Technology & Development, AstraZeneca, Macclesfield, United Kingdom

[3]UCL School of Pharmacy, University College London, London, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Helen J Lycett, BSc, MSc
Spoonful of Sugar Ltd
UCL Business PLC
97 Tottenham Court Road
London, W1T 4TP
United Kingdom
Phone: 44(0)20 3515 5550
Email: helen@sos-adherence.co.uk

Abstract

Background: Asthma is a chronic disease requiring effective self-management to control it and prevent mortality. The use of theory-informed digital interventions promoting asthma self-management is increasing. However, there is limited knowledge concerning how and to what extent psychological theory has been applied to the development of digital interventions, or how using theory impacts outcomes.

Objective: The study aimed to examine the use and application of theory in the development of digital interventions to enhance asthma self-management and to evaluate the effectiveness of theory-based interventions in improving adherence, self-management, and clinical outcomes.

Methods: Electronic databases (CENTRAL, MEDLINE, EMBASE, and PsycINFO) were searched systematically using predetermined terms. Additional studies were identified by scanning references within relevant studies. Two researchers screened titles and abstracts against predefined inclusion criteria; a third resolved discrepancies. Full-text review was undertaken for relevant studies. Those meeting inclusion criteria were assessed for risk of bias using the Cochrane Collaboration tool. The review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. Study outcomes were classified as medication adherence, self-management, asthma control, clinical markers of health, quality of life, other quality of life outcomes, and health care utilization. Effectiveness was calculated as an average outcome score based on the study's reported significance. The Theory Coding Scheme (TCS) was used to establish the extent to which each intervention had applied theory and which theoretical constructs or behavioral determinants were addressed. Associations between TCS scores and asthma outcomes were described within a narrative synthesis.

Results: Fourteen studies evaluating 14 different digital interventions were included in this review. The most commonly cited theories were Social Cognitive Theory, Health Belief Model, and Self-Efficacy Theory. A greater use of theory in the development of interventions was correlated with effective outcomes (r=.657; P=.01): only the 3 studies that met >60% of the different uses of theory assessed by the TCS were effective on all behavioral and clinical outcomes measured. None of the 11 studies that met ≤60% of the TCS criteria were fully effective; however, 3 interventions were partially effective (ie, the intervention had a significant impact on some, but not all, of the outcomes measured). Most studies lacked detail on the theoretical constructs and how they were applied to the development and application of the intervention.

Conclusions: These findings suggest that greater use of theory in the development and application of digital self-management interventions for asthma may increase their effectiveness. The application of theory alone may not be enough to yield a successful

intervention, and other factors (eg, the context in which the intervention is used) should be considered. A systematic approach to the use of theory to guide the design, selection, and application of intervention techniques is needed.

(J Med Internet Res 2018;20(12):e293) doi:10.2196/jmir.9666

KEYWORDS

asthma; adherence; self-management; quality of life; digital interventions; psychological theory

Introduction

Background

Approximately 235 million people worldwide are living with asthma [1]. First-line treatment for this chronic disease consists of a combination of quick-reliever inhalers (short-acting beta-agonists) during exacerbations and daily use of preventer medication (mainly inhaled corticosteroids, ICS) to control the disease [2]. Asthma is usually self-managed at home by the patient or caregivers [3], therefore, its effective control depends upon the patient's behavior [4,5].

Efficient self-management involves active commitment to follow a written asthma action plan, self-monitoring symptoms, controlling environmental factors and, importantly, adhering to treatment [5-7]. Adherence to medication is a major determinant of treatment success in long-term conditions [8,9]. An adherence rate to ICS of >80% is needed to reduce asthma exacerbations [10], successfully control symptoms, and improve lung function [9,11]. This level of adherence has also been shown to decrease hospital admissions by 30% [9].

Despite these benefits, adherence rates to asthma treatment remain low [12] and variable [13]. In general, 30% to 70% of people on long-term preventer therapy do not maintain the high levels of adherence necessary for good asthma control. Suboptimal levels of adherence are found in adults [11], children [14,15], and adolescents [14-16].

Effective self-management of asthma is dependent on multiple factors, including consideration of patients' perceptual and practical barriers to their disease and treatment [4]. Patients adopt self-management and adherence behaviors to cope with their illness, and these are influenced by their perceptions of their condition [17]. Nonadherence to asthma medication is influenced by perceptual barriers such as patients' doubts about their need for treatment and treatment concerns (eg, fears about possible short- or long-term effects of treatment [18]) and/or as a result of practical barriers (eg, forgetting, bad inhaler technique).

Inadequate adherence to preventer medication can lead to overuse of relievers and the prescription of higher doses of medication than the patient needs [9]. Nonadherence has been associated with uncontrolled asthma, poor clinical outcomes, increased hospitalizations, decreased quality of life, absenteeism from work/school, and mortality in adults and children [8,19-21]. Most patients do not inform their health care professional when they stop treatment [8,22]; therefore, there may be limited opportunities to support patients to get the most from their medicines.

There is a clear need for effective self-management interventions, yet, to date, interventions have had varying degrees of success [23]. Digital support services (mobile and Web technologies) may increase the accessibility of interventions, given that most people now use electronic devices in their daily lives [24] and are willing to self-manage their disease using mobile technology interventions [25]. Digital support services can be highly scalable, personalized to increase medication adherence in targeted patient populations, can be applied in real time, and have the potential to provide consistency and delivery at low cost.

Digital adherence interventions, from electronic monitoring to short message service (SMS)–based programs, have been evaluated across long-term conditions with varying degrees of success [26-28]. However, the literature has been dominated by small-scale feasibility and exploratory studies and pilot evaluations that lack statistical power [26,29]. For patients with asthma, digital support services may provide a highly accessible and effective means of monitoring and improving adherence to treatment and disease control.

Recent systematic reviews have found that digital interventions can improve adherence to asthma preventer medication and asthma control when compared with standard treatment [12,30,31]. Miller et al [12] conducted a recent review and meta-analysis of mobile health (mHealth) interventions for the self-management of asthma comparing mHealth interventions with usual care and found a moderate effect on adherence, a large effect on quality of life, but no significant effect on lung function. The authors also found mHealth interventions to be as effective as paper-based monitoring on adherence and clinical outcomes. However, the findings of individual studies have been inconsistent. Although telemonitoring (text messaging, Web systems, etc) was not associated with better control of asthma symptoms when compared with usual care [32], internet-based self-management support has been shown to improve asthma quality of life and asthma control [33].

Guidelines for the development of interventions recommend the use of a theoretical framework or model of behavior change [34-37]. Theory can be used in various ways, for example, to identify modifiable determinants of health behaviors to be addressed within interventions (eg, illness perceptions), to select appropriate techniques to address behavioral determinants (eg, motivational interviewing), or to select people who are most likely to benefit from the intervention (eg, patients who have misconceptions about their illness or treatment). Many theory-based interventions used to explain health behavior have been based on social cognition theories [37,38]. These include Social Cognitive Theory (SCT) [39], the Health Belief Model (HBM) [40], Theory of Reasoned Action (TRA) [41], and Theory of Planned Behavior (TPB) [42], all of which are based

on the premise that people are rational decision makers who can weigh up the advantages and disadvantages of adopting a behavior.

Several reviews of behavior change interventions have shown that interventions that explicitly refer to a theoretical approach to their development are more effective than those that lack a theoretical base [43-45]. A systematic review of interventions to improve adherence to asthma medicines showed that the use of theory was more common among effective than ineffective interventions [46], and another study reviewed the application of behavior change theory and clinical guidelines on internet-based asthma interventions [47]. However, these reviews only indicated whether theory was cited within the paper, rather than the extent to which theory was used to guide the development of the intervention or its effect on clinical outcomes. A review of digital interventions across long-term conditions found that more extensive use of theory was associated with a larger effect on health-related behavior [48]. To date, no systematic reviews of asthma self-management interventions have assessed how the use of theory impacts their effectiveness; therefore, little is known about how and to what extent theory has been applied, which theoretical models show promise, or which components of these models are most effective.

Objectives

This review was designed to address the following questions about how best to use theory in the development of digital self-management interventions for asthma: (1) are theory-based digital interventions to enhance asthma self-management effective at changing behavior and improving clinical outcomes and quality of life?; (2) which theories have been applied to the development of digital interventions to enhance asthma self-management, and which theoretical constructs and behavioral determinants have been addressed?; (3) how and to what extent have theoretical models been applied to the development of digital interventions to enhance asthma self-management?

Methods

Literature Search

Searches were conducted using CENTRAL (The Cochrane Library), MEDLINE, EMBASE, and PsycINFO. Predetermined terms within titles, abstracts, and keywords were used to identify relevant studies. More detailed information about search terms used is available in Multimedia Appendix 1 Searches were completed on June 22, 2017. This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [49].

Selection of Papers

Titles, abstracts, and keywords from the electronic searches were screened independently by 2 researchers (HJL, EKW) and coded as “include” or “exclude” with both researchers screening all studies (100% overlap). Discrepancies were resolved by a third researcher (VC). Full texts of relevant papers were subjected to further scrutiny, and reference lists within relevant papers were hand-searched for significant titles, which were screened following the same process above. Final papers were selected based on the inclusion and exclusion criteria presented in Textbox 1. The selection process of papers for the review is summarized in Figure 1.

Quality Assessment

The Cochrane Collaboration tool [50] was used to assess bias in the studies reporting on randomized controlled trials (RCTs). The item *blinding participants and personnel* was excluded because it would not be possible to blind participants to the use of the digital intervention. Each of the remaining 6 items was rated independently (low/high/unclear) by 2 researchers (HJL, EKW). Any disagreements were resolved through discussion.

Data Extraction and Synthesis

Study Characteristics

Data were extracted by 2 independent researchers (HL, EKW). Data extracted on characteristics of the interventions included country, study design (RCT or pre-post design), inclusion criteria of participants, sample size, percentage of females, and mean age (or range). Details can be found in Multimedia Appendix 2.

Mode of Digital Delivery

Interventions were classified as fully digital or partly digital (digital and nondigital components). Data were extracted on the type of digital platform (eg, SMS, smart device app) and the type of nondigital component (eg, telephone call, paper-based). Full details are available in Multimedia Appendix 3.

Outcomes

To be able to compare the efficacy of the interventions on self-management, and as studies reported on different outcomes, only outcomes relevant to the study were extracted (EKW, EMR; eg, knowledge was not included) and classified under one of these overall themes: adherence to medication, self-management and asthma control, clinical markers of health, quality of life, other quality of life outcomes and health care utilization (Table 1). The intervention was considered to be effective on a specific outcome if the study reported a statistically significant ($P<.05$) improvement in the outcome. This included a significantly improved outcome in the intervention group relative to the control group for RCTs or a significant positive change in the outcome in pre-post studies. A score based on the study’s reported significance level was assigned to each outcome (2=if reported as a significant *P* value, 1=if reported as a marginally significant *P* value, and 0=if reported as not significant). An average score was applied when different suboutcomes of the same outcome were reported in the same study (eg, both symptom days and symptom nights were reported as clinical markers of health [51]). Finally, an average score was calculated for each study by adding the average outcome scores and dividing this result by the total number of outcomes. Therefore, interventions were deemed to be fully effective if they were associated with an outcome average score of 2.0, partially effective if they were in the range

of 1.0 to 1.9, and not effective if the score was in the range of 0 to 0.9.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Paper in English
- Patients with asthma
- Empirical study (pilot, feasibility, or evaluative study)
- Intervention focused on patient (rather than physician or carer)
- Digital intervention (eg, online intervention, smart phone app, electronic monitor, short message service (SMS), interactive voice recognition, or wearable
- Intervention designed to enhance adherence or persistence with asthma medication or self-management
- Explicit mention of the use of theory to design the self-management intervention or to increase engagement with the intervention

Exclusion criteria

- Conference abstracts
- Paper not in English
- Review or letter
- Intervention is delivered to parent(s) of children with asthma
- Not an empirical study
- Clinician focus (clinician attitude, behavior, or diagnostic tool)
- Intervention not designed to enhance self-management or adherence or persistence with asthma medication
- Intervention was not electronic
- Full-text paper not available

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the selection process of studies included in the review.

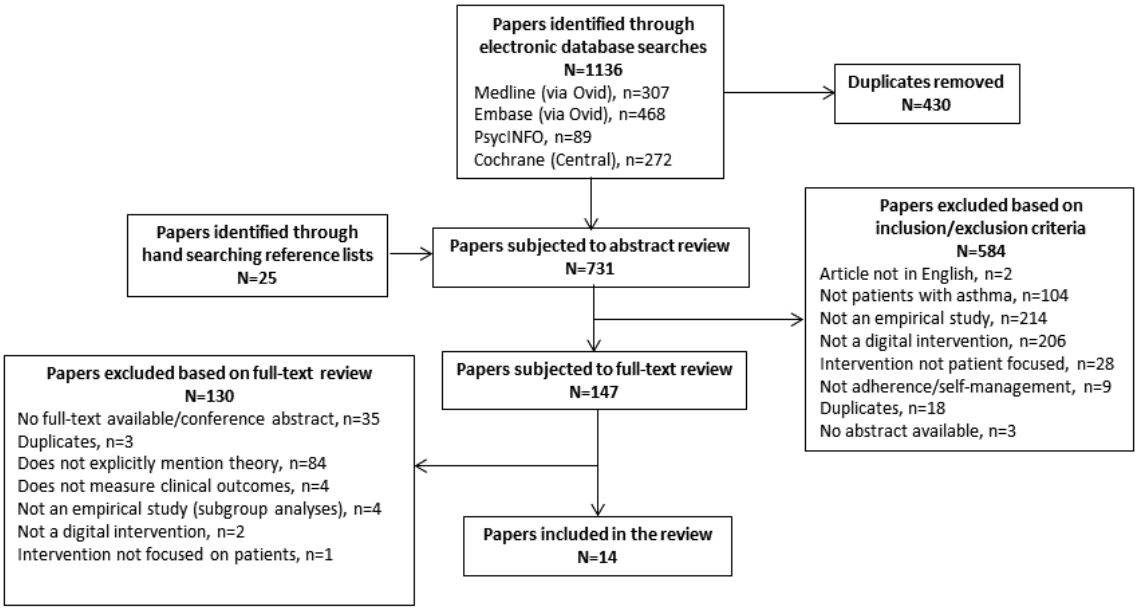


Table 1. Application of theory according to the Theory Coding Scheme (TCS) and effectiveness scores for study outcomes.

Authors, year	N	Behavior change model/theory	Theory Coding Scheme (Item number)	% theory applied	Adherence	Self-management and control	Clinical markers of health	Quality of life	Other Quality of life outcomes	Health care utilization	Outcomes average score
Bartholomew, 2000 [54]	133	SRM[a]; SCT[b]	1; 3; 5; 6	36	N/A[c]	2	2	N/A	N/A	1	1.67
Bartlett, 2002 [62]	16	SLT[d]	1; 2; 3; 5; 6; 7; 8; 9	73	2	N/A	N/A	N/A	N/A	N/A	2
Huss, 2003 [56]	101	PRECEDE-PROCEED model; Developmental; Social Support and learning theories	1; 2; 5	27	N/A	N/A	0	0	N/A	N/A	0
Krishna, 2003 [58]	228	SRT[e]	1	9	N/A	N/A	1.33	0	1	0.67	0.78
Joseph, 2007 [57]	315	Transtheoretical Model; HBM[f]	1; 2; 5; 6; 11	45	1	N/A	2	0	1.33	1.5	1.17
Bender, 2010 [55]	50	Benefit-Risk Model of Health Behavior	1; 2; 3; 5	36	2	0	N/A	0	N/A	N/A	0.68
Petrie, 2012 [60]	148	Extended SRM	1; 2; 3; 4; 5; 6; 7;11	73	2	N/A	N/A	N/A	N/A	N/A	2
Burns, 2013 [63]	51	TPB[g]	1; 3	18	N/A	1	N/A	2	N/A	N/A	1.5
Joseph, 2013 [51]	422	HBM	1; 2; 5; 6; 11	45	N/A	2	1	N/A	0.67	0	0.92
Lau, 2015 [59]	330	HBM; SCT; SET[h]; Transtheoretical change	1; 2; 5; 8; 11	45	N/A	0	0	N/A	0	0	0
Wiecha, 2015 [61]	58	SCT	1	9	1	N/A	0	N/A	0.25	0	0.31
Ahmed, 2016 [53]	98	Behavior change; SET; Motivational Theory	1; 2	18	N/A	0		1	N/A	0	0.33
Speck, 2016 [64]	44	SCT	1; 2; 3; 5; 6; 8; 11	64	N/A	2		2	N/A	N/A	2
Warren, 2016 [65]	12	SRT	1; 2; 3	27	N/A	0	2	0	N/A	N/A	0.67

[a]SRM: Self-Regulatory Model.

[b]SCT: Social Cognitive Theory.

[c]N/A: not applicable.

[d]SLT: Social Learning Theory.

[e]SRT: Self-Regulation Theory.

[f]HBM: Health Belief Model.

[g]TPB: Theory of Planned Behavior.

[h]SET: Self-Efficacy Theory.

Use of Theory

Data extracted included the theory(ies) reported in the intervention and the theoretical construct(s) addressed by the intervention. The Theory Coding Scheme (TCS) [52] was used to assess the extent to which theory had been applied. This instrument consists of 19 items, from which items 1 to 11 were relevant to this review, as items 12 to 19 do not measure the use of theory in the development of the interventions [48]. Items 1 to 11 assessed whether theory was mentioned in the paper, the use of theory to select participants, intervention techniques, or tailoring of the intervention and whether theoretical constructs or behavioral determinants were explicitly linked to intervention techniques [52]. For each study, a percentage score was calculated representing the proportion of relevant TCS items applied to the intervention ([number of TCS items applied divided by number of relevant TCS items] × 100).

Data Synthesis

Narrative synthesis was used to describe the impact of the interventions on the study outcomes and the application of theory in the development of the interventions. Pearson correlation coefficients were used to calculate the correlation between the effectiveness of interventions and the percentage score for the use of theory.

Results

Characteristics of the Interventions

From 1136 papers originally identified, 14 met the inclusion criteria (Figure 1). Multimedia Appendix 2 shows full details of the studies' design and population characteristics. Of the 14 studies, 71% (10/14) reported on RCTs [51,53-61]), and 29% (4/14) were feasibility studies employing a pre-post design [62-65]. In all, 71% (10/14) of studies were undertaken in the United States. Studies included children (36%, 5/14) [54,56,61,62,65]), adolescents (14%, 2/14) [51,57]), adults (43%, 6/14); [53,55,59,60,63,64]), and mixed samples (7%, 1/14) [58]). Between 35% and 82% of the samples were female. Sample sizes ranged from 16 to 422 and included a total of 1856 participants. Multimedia Appendix 3 shows details of the type of digital platforms, the frequency of the interventions, details of the nondigital component, if applicable, and control conditions. None of the included studies incorporated measures to prevent dropout, with details of adoption and engagement with the interventions shown in Multimedia Appendix 2. A total of 2 studies involved patients in the development of the interventions [55,63].

Effectiveness of Theory-Based Digital Interventions to Enhance Asthma Self-Management

Effect of Interventions on Behavioral Outcomes

Medication Adherence

Five studies (36%, 5/14) reported on adherence to preventer medications (Table 1), from which 3 studies measured adherence using electronic monitoring [55,61,62], and 2 used self-report [57,60]. This included 4 RCTs [55,57,60,61] and the single pre-post study [62]. A total of 3 studies reported a significant positive effect of the intervention on adherence [55,60,62]. Moreover, 2 studies were considered as having a partial effect, 1 reported controller medication adherence improved significantly from baseline for the subgroup of subjects with low (<75%) adherence on the intervention group only but also reported no significant differences in change between the intervention group and control group (P=.10) [61]; the other study [57] described their result as only marginally significant (P=.09; see Table 1).

Self-Management and Control

A total of 8 studies (57%; 8/14) measured self-management and control outcomes (Multimedia Appendix 4). Each of the 8 studies (5 RCTs and 3 pre-post studies) that measured self-management behavior and control [51,53-55,59,63-65] used a different measure. In terms of self-management, these included the Partners in Health Scale [63], a validated measure of self-management behaviors [54] and the Asthma Belief survey [65]. Asthma control was measured by the Asthma Control Questionnaire (ACQ) [59], the Asthma Control Test (ACT) [53,55,64], potential overuse of rescue fast-acting bronchodilators [53], indicators of uncontrolled asthma [51], and the Royal College of Physicians 3-questions screening tool [63]. In addition, 3 studies reported a significant positive effect of the intervention on self-management behavior [54], 2 studies reported a significant positive effect on asthma control [51,64], and 1 study [63] reported the intervention had a significant positive outcome on asthma control but not on self-management (Table 1). Only 1 [64] of the 2 pre-post studies showing a significant effect of the intervention on asthma control reported that the improvement of over 3 points on the ACT at 3 months was greater than the minimally important difference.

Effect of the Interventions on Clinical Outcomes

Clinical Markers of Health

A total of 8 studies (57%, 8/14: 7 RCTs and 1 pre-post study) reported on clinical markers of health (Multimedia Appendix 4). Measures included asthma symptoms, symptom days or symptom nights [51,54,57,58], forced expiratory volume [56], functional status measure [54], severe asthma exacerbation [58], worsening of asthma needing treatment changes [58], reported days of wheezing [61], peak expiratory flow rate [65], days of reliever use, and average daily dose of ICS [58]. Moreover, 3 studies reported a significant effect of the intervention on all of their clinical markers measured [54,57,65] (Table 1).

Quality of Life

A total of 8 studies (57%, 8/14: 5 RCTs and 3 pre-post studies) reported on quality of life [53,55-58,63-65] (Multimedia Appendix 4). Validated measures included the Asthma Quality of Life Questionnaire (AQLQ) [55,64], the Paediatric Asthma Quality of Life Questionnaire [56,58,65], and the mini AQLQ [53]. Two studies developed a quality of life measure specific to their study [57,63]. In addition, 2 studies [63,64] reported a significant positive effect of the intervention on quality of life (Table 1). One study [53] reported a significant improvement from baseline to 3 months, but this effect was not significant at 6- and 9-month follow-ups.

Other Quality of Life Outcomes

A total of 5 studies (36%; 5/14) reported on factors influencing quality of life [51,57-59,61] (Multimedia Appendix 4). These included nights of sleep disturbance or patient awakening [58,61], days of activity limitation/restricted activity [51,57,58,61], number of school days missed [51,57,58,61], number of work days missed [59], days of changed plans [51,57], and number of days the patient had to slow down [61]. Two studies were partly effective in improving these outcomes [57,58] (Table 1). For example, although in 1 study [58], days of activity limitation and number of school days missed significantly decreased in the intervention group only (P<.01), there were no significant differences between the control and intervention groups.

Health Care Utilization

A total of 7 studies (50%, 7/14, all being RCTs) reported on health care utilization [51,53,54,57-59,61] (Multimedia Appendix 4). All measured the number of emergency department visits or hospitalizations over a given time. A total of 3 studies also reported the total number of urgent visits to a health care professional, general practitioner, or physician [58,59,61]. In all, 2 studies reported a significant decrease in hospitalizations following the intervention but no significant differences in emergency room visits [54,57]. One study found a significant decrease in emergency department annual visits in the intervention group but not for the number of hospitalizations or urgent visits to physicians [58]. A total of 4 studies did not find any significant effect of the intervention on health care utilization outcomes (Table 1).

Theories That Have Been Applied to Intervention Development

Details of the theoretical basis of the intervention are shown in Table 1. Theories included Social Cognitive Theory [54,59,61,64], Health Belief Model [51,57,59], Theory of Planned Behavior [63]; Social Learning Theory [62], the Transtheoretical Model [57], the PRECEDE-PROCEED model [56], developmental and social support and learning theories [56], Behavior Change theory and Motivational theory [53], the Benefit-Risk Model of Health Behavior [55], and Self-Efficacy Theory [53,59]. A total of 5 of the interventions referenced the Self-Regulatory Model, Common Sense Model of Self-Regulation, Extended Common Sense Model of Self-Regulation, Illness Perceptions, or Necessity Concerns Framework in the development of the intervention [54,55,58,60,65].

Theoretical Constructs That Have Been Addressed

Theoretical constructs/behavioral determinants specified within the models and addressed in the interventions included illness perceptions, which specifically explored identity, consequences, timeline, personal control, treatment control, concern, understanding, and emotional response to the illness [60]; beliefs about medicines were addressed in 3 interventions by targeting patients' beliefs about the necessity of their medication and their concerns about taking their medication [53,55,60]. General control beliefs [54,64] and self-efficacy beliefs looked at how confident patients felt in areas such as self-management, which is taking medicines as prescribed, and self-awareness, which includes recognizing and acting on the symptoms [53,59,62,65].

The Extent Theoretical Models Have Been Applied

Responses to the TCS are shown in Table 1, and the frequency each item was reported in the studies is illustrated in Figure 2. In line with the study inclusion criteria, all studies (100%, 14/14) mentioned theory (item 1), and 10 studies (71%, 10/14) [51,53,55-57,59-62,64] mentioned a target construct as a predictor of behavior (item 2). Theory was explicitly used to select or develop intervention techniques (item 5) in 9 studies (64%, 9/14) [51,54-57,59,60,62,64]. A total of 7 studies (50%, 7/14) [54,55,60,62-65] referred to the application of a single theory rather than a combination of different theories (Item 3). A total of 6 studies (43%, 6/14) [51,54,57,60,62,64] used theory or predictors to tailor intervention techniques to participants (item 6). A total of 3 studies (21%, 3/14) [59,62,64] linked at least 1 intervention technique to a theory-relevant construct/predictor (item 8), 2 studies (14%, 2/14) [60,62] linked all intervention techniques to at least 1 theory-relevant predictor (item 7), and 1 study [62] (7%, 1/14) linked a group of techniques to a group of clusters/predictors (item 9). Only 1 study (7%, 1/14) [60] screened or selected participants based on a particular score or level on a theory-relevant construct or predictor (item 4). No studies linked every theoretical construct within a stated theory to an intervention technique (item 10); however, 5 studies (36%, 5/14) [51,57,59,60,64] linked at least 1 theoretical construct to at least 1 intervention technique (item 11).

The use of theory as assessed by the TCS ranged from 9% to 73%. Three studies applied >60% of the different uses of theory based on the items of the TCS (6 items) [60,62,64] (Table 1). All 3 of these studies (100%) showed a significant positive effect of the intervention on all behavioral and clinical outcomes measured (average score 2.0; Table 1). Comparably, from the 11 studies that incorporated ≤60% of theory, no study was fully effective, but 3 interventions were partially effective (average score range 1.0-1.9) [54,57,63]. All other studies yielded average scores of <1.0. There was a significant correlation between the percentage of theory applied to the interventions and the effectiveness of the intervention (outcomes average score) (r=.657; P=.01). To assess whether sample size had an influence on the results, correlations were recalculated excluding Bartlett et al [62], with a small sample size of n=16. Results showed correlations were still highly significant, indicating theory and effectiveness were not biased by sample size (r=.581; P=.04). None of the studies reported using theory to promote engagement with the intervention.

Figure 2. Frequency items from the Theory Coding Scheme used in the studies.

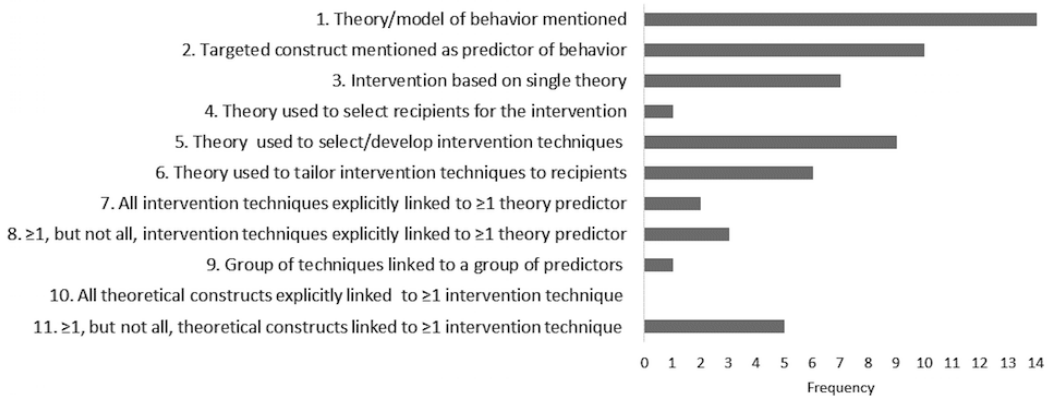
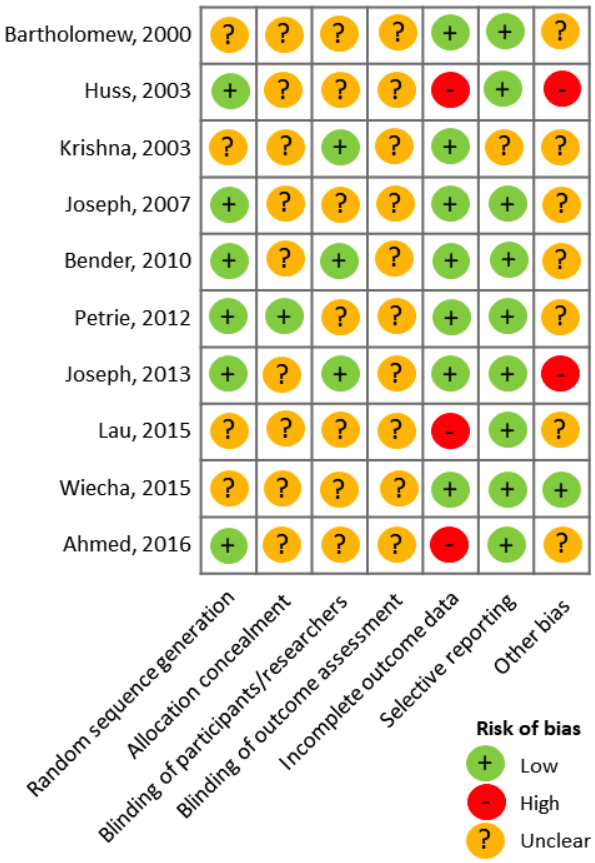


Figure 3. Risk of bias across interventions.



Risk of Bias Assessment

Risk of bias assessment was performed on the 10 RCTs included in the review. The results are shown in Figure 3, and complete details are provided in Multimedia Appendix 5. A total of 6 studies reported using appropriate random sequence generation methods; of these all used computer-generated random allocation [51,53,55-57,60]. Four studies did not specify the method of randomization [54,58,59,61]. Only 1 study reported concealment of allocation [60], while this was unclear for the remaining 9 studies. None of the studies specified whether there had been blinding of the outcome assessment. Three studies were considered to have high risk of incomplete outcome data [53,56,59] due to high rates of attrition, whereas the remainder were considered to have low risk. A total of 9 studies were assessed as having low-risk of selective reporting, while this was unclear in 1 study [58] as measures had not been stated at the outset.

Discussion

Summary of Findings

This review identified 14 studies that evaluated theory-based digital interventions in RCTs or pre-post studies. A range of different theories had been used in the development of these interventions, most frequently Social Cognitive Theory, the Health Belief Model, and the Common-Sense Model of Self-Regulation [51,54,55,57-61,64,65]. The findings indicate that the use of psychological theory can enhance the effectiveness of digital interventions, as interventions that incorporated a more extensive use of theory were more likely to achieve successful outcomes. These findings are consistent with those of previous systematic reviews showing that digital self-management interventions can be effective at improving clinical outcomes in asthma [12,30,66] and suggest that theory-based interventions may be more effective than interventions that have not used theory in their development [46]. A previous meta-analytic review of internet- based interventions also found that extensive use of theory was associated with larger effect sizes on health behavior change [48].

To our knowledge, this is the first systematic review to examine the extent to which theory has been applied to the development of digital self-management interventions for asthma. We found substantial differences between studies in terms of their use of theory. Although most of the studies that mentioned theory referred to the use of theory in the development of their interventions, fewer studies explicitly reported the use of theory to select recipients for the intervention or indicated how they had linked intervention techniques, relevant constructs or predictors. Our findings suggest that interventions that incorporated these items in their development were more likely to be effective; however, only a small number of studies utilized these constructs. Further research is, therefore, required to ascertain how the application of theory in the development of interventions impacts their effectiveness.

Other factors, such as the delivery channel (eg, via different digital platforms), the context in which the intervention is delivered (eg, via hospital or routine assessments), and the type of user (eg, children vs adults) may also influence outcomes. The fact that interventions that applied theory to a similar extent could have varying degrees of effectiveness implies that the use of theory is necessary, but not sufficient, for a successful intervention.

Strengths and Limitations

The strengths of this review include the systematic approach, inclusion of a range of interventions focusing on many different self-management behaviors, and the use of a reliable instrument to determine the extent to which theory had been used to inform the design of the interventions. The heterogeneity in outcomes measured precluded the use of meta-analysis, therefore, we were not able to determine the size of the effect. Although the findings indicate an increasing number of researchers are utilizing theory in the development of digital interventions for asthma, there were insufficient numbers of studies referencing each theoretical model to determine whether any one theory showed promise over another.

Limitations of the individual studies included a lack of information describing the interventions. Often it was not possible to determine which behavioral determinants had been targeted or how they had been addressed by the intervention. This could be improved in future studies through the use of a framework such as the Template for Intervention Description and Replication checklist [67] to describe the intervention. This would not only aid replication but also allow a more reliable and thorough assessment of the process by which digital self-management interventions exert their effect. In addition, there was a lack of information on methods of randomization and concealment in many of the studies, meaning that the risk of bias was often unclear. Eysenbach [58] stated there is a need to address the “law of attrition,” which relates to the dropout and nonengagement in electronic health users. A high dropout rate was observed in the interventions included within this review [53,59,60]. However, none of the included studies incorporated measures to engage participants in the intervention and prevent dropout, and none of the studies mentioned they used theory to increase engagement with the interventions. The short duration of some studies means that individual studies may have been underpowered or overpowered for individual outcomes. Further research is needed to explore how theory could specifically target engagement behavior to achieve effective engagement.

Implications of Our Findings for Clinical Care and Future Research

Our findings suggest that theory-based digital interventions to enhance asthma self-management can be effective at improving adherence and self-management and that more extensive use of theory in the development and application of digital interventions for asthma self-management may enhance their effectiveness. However, although a number of theories have been applied to the development of asthma digital interventions, it is not clear whether any particular theory is more effective. Furthermore, most studies lack details on the theoretical constructs used and behavioral determinants addressed by the intervention, and whether or how these have been applied to the design or application of the intervention. The systematic recording and reporting on the use of theory in the development of future interventions is, therefore, important. It is not sufficient to merely state theory has been used; there should be specific reference to exactly how it has informed the design of the intervention. The TCS can be used to inform the design of an intervention, ensuring that the theoretical basis of an intervention is adequately and clearly described so that the use of theory can be evaluated.

Acknowledgments

AstraZeneca commissioned and funded Spoonful of Sugar to undertake this review. The funders have had no influence on the content of the review.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms. Terms within columns were combined using the Boolean "OR" operator, terms between columns were then combined with "AND," that is, papers were retrieved if the title/abstract/keywords contained at least one term from each column.

[PDF File (Adobe PDF File), 111KB - jmir_v20i12e293_app1.pdf]

Multimedia Appendix 2

Study characteristics: Study design, population characteristics, and intervention engagement.

[PDF File (Adobe PDF File), 126KB - jmir_v20i12e293_app2.pdf]

Multimedia Appendix 3

Details of mode of delivery of the intervention.

[PDF File (Adobe PDF File), 118KB - jmir_v20i12e293_app3.pdf]

Multimedia Appendix 4

Behavioral (adherence, self-management, and control) and clinical outcomes of studies. Randomized control trials' (RCT) values reported in this table refer to the differences between intervention (IG) and control groups (CG); values for pre-post intervention studies report changes from baseline.

[PDF File (Adobe PDF File), 223KB - jmir_v20i12e293_app4.pdf]

Multimedia Appendix 5

Risk of Bias table.

[PDF File (Adobe PDF File), 158KB - jmir_v20i12e293_app5.pdf]

References

1. World Health Organization. 2017. Bronchial asthma URL: http://www.who.int/en/news-room/fact-sheets/detail/asthma [WebCite Cache ID 73EmCWAJP]
2. Shahidi N, Fitzgerald JM. Current recommendations for the treatment of mild asthma. J Asthma Allergy 2010;3:169-176 [FREE Full text] [doi: 10.2147/JAA.S14420] [Medline: 21437051]
3. Howell G. Nonadherence to medical therapy in asthma: risk factors, barriers, and strategies for improving. J Asthma 2008;45(9):723-729. [doi: 10.1080/02770900802395512] [Medline: 18972285]
4. Horne R, Weinman J. Self-regulation and self-management in asthma: exploring the role of illness perceptions and treatment beliefs in explaining non-adherence to preventer medication. Psychol Health 2002;17(1):17-32. [doi: 10.1080/08870440290001502]
5. Eakin MN, Rand CS. Improving patient adherence with asthma self-management practices: what works? Ann Allergy Asthma Immunol 2012;109(2):90-92 [FREE Full text] [doi: 10.1016/j.anai.2012.06.009] [Medline: 22840247]
6. Global Initiative for Asthma. Global strategy for asthma management and prevention. 2018. URL: https://ginasthma.org/ [WebCite Cache ID 73EmRQyD9]
7. Pinnock H, Epiphaniou E, Pearce G, Parke H, Greenhalgh T, Sheikh A, et al. Implementing supported self-management for asthma: a systematic review and suggested hierarchy of evidence of implementation studies. BMC Med 2015;13(1):127 [FREE Full text] [doi: 10.1186/s12916-015-0361-0] [Medline: 26032941]
8. Gamble J, Stevenson M, McClean E, Heaney LG. The prevalence of nonadherence in difficult asthma. Am J Respir Crit Care Med 2009;180(9):817-822. [doi: 10.1164/rccm.200902-0166OC] [Medline: 19644048]
9. Mokoka MC, Lombard L, MacHale EM, Walsh J, Cushen B, Sulaiman I, et al. In patients with severe uncontrolled asthma, does knowledge of adherence and inhaler technique using electronic monitoring improve clinical decision making? A

protocol for a randomised controlled trial. BMJ Open 2017 Jun 15;7(6):e015367 [FREE Full text] [doi: 10.1136/bmjopen-2016-015367] [Medline: 28619778]
10. Santos PD, D'Oliveira Júnior A, Noblat LD, Machado AS, Noblat AC, Cruz Á. Preditores da adesão ao tratamento em pacientes com asma grave atendidos em um centro de referência na Bahia. J Bras Pneumol 2008;34:995-1002. [doi: 10.1590/S1806-37132008001200003]
11. Kandane-Rathnayake RK, Matheson MC, Simpson JA, Tang ML, Johns DP, Mészáros D, et al. Adherence to asthma management guidelines by middle-aged adults with current asthma. Thorax 2009;64(12):1025-1031. [doi: 10.1136/thx.2009.118430] [Medline: 19703827]
12. Miller L, Schüz B, Walters J, Walters EH. Mobile technology interventions for asthma self-management: systematic review and meta-analysis. JMIR Mhealth Uhealth 2017;5(5):e57 [FREE Full text] [doi: 10.2196/mhealth.7168] [Medline: 28465281]
13. Bender BG, Pedan A, Varasteh LT. Adherence and persistence with fluticasone propionate/salmeterol combination therapy. J Allergy Clin Immunol 2006;118(4):899-904. [doi: 10.1016/j.jaci.2006.07.002] [Medline: 17030244]
14. Morton RW, Everard ML, Elphick HE. Adherence in childhood asthma: the elephant in the room. Arch Dis Child 2014;99(10):949-953. [doi: 10.1136/archdischild-2014-306243] [Medline: 24876303]
15. De Simoni A, Horne R, Fleming L, Bush A, Griffiths C. What do adolescents with asthma really think about adherence to inhalers? Insights from a qualitative analysis of a UK online forum. BMJ Open 2017;7(6):e015245 [FREE Full text] [doi: 10.1136/bmjopen-2016-015245] [Medline: 28615272]
16. Bender B, Wamboldt FS, O'Connor SL, Rand C, Szefler S, Milgrom H, et al. Measurement of children's asthma medication adherence by self report, mother report, canister weight, and Doser CT. Ann Allergy Asthma Immunol 2000;85(5):416-421. [Medline: 11101187]
17. Leventhal H, Diefenbach M, Leventhal EA. Illness cognition: Using common sense to understand treatment adherence and affect cognition interactions. Cogn Ther Res 1992;16(2):143-163. [doi: 10.1007/BF01173486]
18. Horne R. Compliance, adherence, and concordance: implications for asthma treatment. Chest 2006;130(1):65S-72S. [doi: 10.1378/chest.130.1_suppl.65S] [Medline: 16840369]
19. Williams LK, Pladevall M, Xi H, Peterson EL, Joseph C, Lafata JE, et al. Relationship between adherence to inhaled corticosteroids and poor outcomes among adults with asthma. J Allergy Clin Immunol 2004;114(6):1288-1293. [doi: 10.1016/j.jaci.2004.09.028] [Medline: 15577825]
20. Desai M, Oppenheimer JJ. Medication adherence in the asthmatic child and adolescent. Curr Allergy Asthma Rep 2011;11(6):454-464. [doi: 10.1007/s11882-011-0227-2] [Medline: 21968618]
21. Royal College of Physicians. Why asthma still kills: the National Review of Asthma Deaths (NRAD) Confidential Enquiry report. 2014. URL: http://eprints.uwe.ac.uk/23453/1/why-asthma-still-kills-full-report.pdf [WebCite Cache ID 73O4RCiIU]
22. Levine RJ. Monitoring for adherence: ethical considerations. Am J Respir Crit Care Med 1994;149(2):287-288. [doi: 10.1164/ajrccm.149.2.8306018] [Medline: 8306018]
23. Ring N, Malcolm C, Wyke S, Macgillivray S, Dixon D, Hoskins G, et al. Promoting the use of Personal Asthma Action Plans: a systematic review. Prim Care Respir J 2007;16(5):271-283 [FREE Full text] [doi: 10.3132/pcrj.2007.00049] [Medline: 17710351]
24. Gould G, Nilforooshan R. WhatsApp Doc? BMJ Innov 2016;2(3):109-110 [FREE Full text] [doi: 10.1136/bmjinnov-2016-000116] [Medline: 27547446]
25. Fonseca JA, Costa-Pereira A, Delgado L, Fernandes L, Castel-Branco MG. Asthma patients are willing to use mobile and web technologies to support self-management. Allergy 2006;61(3):389-390. [doi: 10.1111/j.1398-9995.2006.01016.x] [Medline: 16436151]
26. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
27. DeKoekkoek T, Given B, Given C, Ridenour K, Schueller M, Spoelstra S. mHealth SMS text messaging interventions and to promote medication adherence: an integrative review. J Clin Nurs 2015;24(19-20):2722-2735. [doi: 10.1111/jocn.12918] [Medline: 26216256]
28. Kannisto KA, Adams CE, Koivunen M, Katajisto J, Välimäki M. Feedback on SMS reminders to encourage adherence among patients taking antipsychotic medication: a cross-sectional survey nested within a randomised trial. BMJ Open 2015;5(11):e008574 [FREE Full text] [doi: 10.1136/bmjopen-2015-008574] [Medline: 26553830]
29. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014(11):CD000011. [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
30. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
31. McLean G, Murray E, Band R, Moffat KR, Hanlon P, Bruton A, et al. Interactive digital interventions to promote self-management in adults with asthma: systematic review and meta-analysis. BMC Pulm Med 2016;16(1):83 [FREE Full text] [doi: 10.1186/s12890-016-0248-7] [Medline: 27215329]

32. Kew KM, Cates CJ. Home telemonitoring and remote feedback between clinic visits for asthma. Cochrane Database Syst Rev 2016(8):CD011714. [doi: 10.1002/14651858.CD011714.pub2] [Medline: 27486836]
33. van Gaalen JL, Beerthuizen T, van der Meer V, van Reisen P, Redelijkheid GW, Snoeck-Stroband JB, SMASHING Study Group. Long-term outcomes of internet-based self-management support in adults with asthma: randomized controlled trial. J Med Internet Res 2013;15(9):e188 [FREE Full text] [doi: 10.2196/jmir.2640] [Medline: 24028826]
34. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
35. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Br Med J 2008;337:a1655 [FREE Full text] [Medline: 18824488]
36. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. Appl Psychol 2008;57(4):660-680. [doi: 10.1111/j.1464-0597.2008.00341.x]
37. Stavri Z, Michie S. Classification systems in behavioural science: current systems and lessons from the natural, medical and social sciences. Health Psychol Rev 2012;6(1):113-140. [doi: 10.1080/17437199.2011.641101]
38. Holmes EA, Hughes DA, Morrison VL. Predicting adherence to medications using health psychology theories: a systematic review of 20 years of empirical research. Value Health 2014;17(8):863-876 [FREE Full text] [doi: 10.1016/j.jval.2014.08.2671] [Medline: 25498782]
39. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice Hall; 1986.
40. Hochbaum G, Rosenstock I, Kegels S. Health Belief Model.: United States Public Health Service; 1952. URL: http://www.infosihat.gov.my/infosihat/artikelHP/bahanrujukan/HE_DAN_TEORI/DOC/Health%20Belief%20Model.doc [WebCite Cache ID 73MJVSm78]
41. Fishbein MA, Ajzen I. Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research. Reading, MA: Addison-Wesley; 1975.
42. Azjen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control: From Cognition to Behavior. Berlin, Heidelberg: Springer; 1985:11-39.
43. Ammerman AS, Lindquist CH, Lohr KN, Hersey J. The efficacy of behavioral interventions to modify dietary fat and fruit and vegetable intake: a review of the evidence. Prev Med 2002;35(1):25-41. [Medline: 12079438]
44. Legler J, Meissner HI, Coyne C, Breen N, Chollette V, Rimer BK. The effectiveness of interventions to promote mammography among women with historically lower rates of screening. Cancer Epidemiol Biomarkers Prev 2002;11(1):59-71 [FREE Full text] [Medline: 11815402]
45. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
46. McCullough AR, Ryan C, Macindoe C, Yii N, Bradley JM, O'Neill B, et al. Behavior change theory, content and delivery of interventions to enhance adherence in chronic respiratory disease: a systematic review. Respir Med 2016;116:78-84 [FREE Full text] [doi: 10.1016/j.rmed.2016.05.021] [Medline: 27296825]
47. Al-Durra M, Torio MB, Cafazzo JA. The use of behavior change theory in Internet-based asthma self-management interventions: a systematic review. J Med Internet Res 2015;17(4):e89 [FREE Full text] [doi: 10.2196/jmir.4110] [Medline: 25835564]
48. 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;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
49. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
50. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [Medline: 22008217]
51. Joseph CL, Ownby DR, Havstad SL, Saltzgaber J, Considine S, Johnson D, et al. Evaluation of a web-based asthma management intervention program for urban teenagers: reaching the hard to reach. J Adolesc Health 2013;52(4):419-426 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.07.009] [Medline: 23299008]
52. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychol 2010;29(1):1-8. [doi: 10.1037/a0016939] [Medline: 20063930]
53. Ahmed S, Ernst P, Bartlett SJ, Valois MF, Zaihra T, Paré G, et al. The effectiveness of web-based asthma self-management system, My Asthma Portal (MAP): a pilot randomized controlled trial. J Med Internet Res 2016;18(12):e313 [FREE Full text] [doi: 10.2196/jmir.5866] [Medline: 27908846]
54. Bartholomew LK, Gold RS, Parcel GS, Czyzewski DI, Sockrider MM, Fernandez M, et al. Watch, Discover, Think, and Act: evaluation of computer-assisted instruction to improve asthma self-management in inner-city children. Patient Educ Couns 2000;39(2-3):269-280. [Medline: 11040726]

55. Bender BG, Apter A, Bogen DK, Dickinson P, Fisher L, Wamboldt FS, et al. Test of an interactive voice response intervention to improve adherence to controller medications in adults with asthma. J Am Board Fam Med 2010;23(2):159-165 [FREE Full text] [doi: 10.3122/jabfm.2010.02.090112] [Medline: 20207925]
56. Huss K, Winkelstein M, Nanda J, Naumann PL, Sloand ED, Huss RW. Computer game for inner-city children does not improve asthma outcomes. J Pediatr Health Care 2003;17(2):72-78. [doi: 10.1067/mph.2003.28] [Medline: 12665729]
57. Joseph CL, Peterson E, Havstad S, Johnson CC, Hoerauf S, Stringer S, et al. A web-based, tailored asthma management program for urban African-American high school students. Am J Respir Crit Care Med 2007;175(9):888-895 [FREE Full text] [doi: 10.1164/rccm.200608-1244OC] [Medline: 17290041]
58. Krishna S, Francisco BD, Balas EA, König P, Graff GR, Madsen RW. Internet-enabled interactive multimedia asthma education program: a randomized trial. Pediatrics 2003;111(3):503-510. [Medline: 12612228]
59. Lau AY, Arguel A, Dennis S, Liaw ST, Coiera E. "Why didnt it work?" Lessons from a randomized controlled trial of a web-based personally controlled health management system for adults with asthma. J Med Internet Res 2015;17(12):e283 [FREE Full text] [doi: 10.2196/jmir.4734] [Medline: 26678294]
60. Petrie KJ, Perry K, Broadbent E, Weinman J. A text message programme designed to modify patients' illness and treatment beliefs improves self-reported adherence to asthma preventer medication. Br J Health Psychol 2012;17(1):74-84. [doi: 10.1111/j.2044-8287.2011.02033.x] [Medline: 22107110]
61. Wiecha JM, Adams WG, Rybin D, Rizzodepaoli M, Keller J, Clay JM. Evaluation of a web-based asthma self-management system: a randomised controlled pilot trial. BMC Pulm Med 2015;15(1):17-27 [FREE Full text] [doi: 10.1186/s12890-015-0007-1] [Medline: 25885418]
62. Bartlett SJ, Lukk P, Butz A, Lampros-Klein F, Rand CS. Enhancing medication adherence among inner-city children with asthma: results from pilot studies. J Asthma 2002;39(1):47-54. [Medline: 11883739]
63. Burns P, Jones SC, Iverson D, Caputi P. AsthmaWise - a field of dreams? The results of an online education program targeting older adults with asthma. J Asthma 2013;50(7):737-744. [doi: 10.3109/02770903.2013.799688] [Medline: 23745621]
64. Speck AL, Hess M, Baptist AP. An electronic asthma self-management intervention for young African American adults. J Allergy Clin Immunol Pract 2016;4(1):89-95. [doi: 10.1016/j.jaip.2015.08.007] [Medline: 26441151]
65. Warren CM, Dyer A, Blumenstock J, Gupta RS. Leveraging mobile technology in a school-based participatory asthma intervention: Findings from the Student Media-Based Asthma Research Team (SMART) study. J Health Educ 2016;47(2):59-70. [doi: 10.1080/19325037.2015.1133337]
66. Tran N, Coffman JM, Sumino K, Cabana MD. Patient reminder systems and asthma medication adherence: a systematic review. J Asthma 2014;51(5):536-543. [doi: 10.3109/02770903.2014.888572] [Medline: 24506699]
67. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687. [Medline: 24609605]

Abbreviations

ACT: Asthma Control Test
AQLQ: Asthma Quality of Life Questionnaire
HBM: Health Belief Model
ICS: inhaled corticosteroids
mHealth: mobile health
RCT: randomized controlled trials
SCT: Social Cognitive Theory
SET: Self-Efficacy Theory.
SLT: Social Learning Theory
SMS: short message service
SRM: Self-Regulatory Model
SRT: Self-Regulation Theory
TCS: Theory Coding Scheme
TPB: Theory of Planned Behavior

Edited by G Eysenbach; submitted 15.01.18; peer-reviewed by L Morrison, E Arden-Close, D Carpenter, A Kotlo; comments to author 22.03.18; revised version received 03.08.18; accepted 07.08.18; published 12.12.18.

Please cite as:
Lycett HJ, Raebel EM, Wildman EK, Guitart J, Kenny T, Sherlock JP, Cooper V
Theory-Based Digital Interventions to Improve Asthma Self-Management Outcomes: Systematic Review
J Med Internet Res 2018;20(12):e293
URL: https://www.jmir.org/2018/12/e293/
doi:10.2196/jmir.9666
PMID:30541741

Review

The Use and Effects of Electronic Health Tools for Patient Self-Monitoring and Reporting of Outcomes Following Medication Use: Systematic Review

Karla Lancaster[1], BSc, MSc; Aseel Abuzour[2], MPharm, PhD; Manmeet Khaira[2,3], BSc (Hons), PharmD; Annalise Mathers[2], BSc, MPH; April Chan[4], BSc, PharmD, ACPR; Vivian Bui[5], BSc, PharmD; Annie Lok[6], BA, MHE; Lehana Thabane[4], BSc, MSc, PhD; Lisa Dolovich[1,2], BSc, MSc, PharmD

[1]Department of Family Medicine, McMaster University, Hamilton, ON, Canada

[2]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

[3]School of Pharmacy, University of Waterloo, Waterloo, ON, Canada

[4]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

[5]Pharmacy Department, Sunnybrook Health Sciences, Toronto, ON, Canada

[6]Michael G. DeGroote National Pain Centre, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Lisa Dolovich, BSc, MSc, PharmD
Leslie Dan Faculty of Pharmacy
University of Toronto
144 College Street
Toronto, ON, M5S 3M2
Canada
Phone: 1 416 978 6608
Fax: 1 416 978 8511
Email: lisa.dolovich@utoronto.ca

Abstract

Background: Electronic health (eHealth) tools are becoming increasingly popular for helping patients' self-manage chronic conditions. Little research, however, has examined the effect of patients using eHealth tools to self-report their medication management and use. Similarly, there is little evidence showing how eHealth tools might prompt patients and health care providers to make appropriate changes to medication use.

Objective: The objective of this systematic review was to determine the impact of patients' use of eHealth tools on self-reporting adverse effects and symptoms that promote changes to medication use. Related secondary outcomes were also evaluated.

Methods: MEDLINE, EMBASE, and CINAHL were searched from January 1, 2000, to April 25, 2018. Reference lists of relevant systematic reviews and included articles from the literature search were also screened to identify relevant studies. Title, abstract, and full-text review as well as data extraction and risk of bias assessment were performed independently by 2 reviewers. Due to high heterogeneity, results were not meta-analyzed and instead presented as a narrative synthesis.

Results: A total of 14 studies, including 13 randomized controlled trials (RCTs) and 1 open-label intervention, were included, from which 11 unique eHealth tools were identified. In addition, 14 RCTs found statistically significant increases in positive medication changes as a result of using eHealth tools, as did the single open-label study. Moreover, 8 RCTs found improvement in patient symptoms following eHealth tool use, especially in adolescent asthma patients. Furthermore, 3 RCTs showed that eHealth tools might improve patient self-efficacy and self-management of chronic disease. Little or no evidence was found to support the effectiveness of eHealth tools at improving medication recommendations and reconciliation by clinicians, medication-use behavior, health service utilization, adverse effects, quality of life, or patient satisfaction. eHealth tools with multifaceted functionalities and those allowing direct patient-provider communication may be more effective at improving patient self-management and self-efficacy.

Conclusions: Evidence suggests that the use of eHealth tools may improve patient symptoms and lead to medication changes. Patients generally found eHealth tools useful in improving communication with health care providers. Moreover, health-related outcomes among frequent eHealth tool users improved in comparison with individuals who did not use eHealth tools frequently.

Implementation issues such as poor patient engagement and poor clinician workflow integration were identified. More high-quality research is needed to explore how eHealth tools can be used to effectively manage use of medications to improve medication management and patient outcomes.

(J Med Internet Res 2018;20(12):e294) doi:10.2196/jmir.9284

KEYWORDS

eHealth; mHealth; electronic health record; telemedicine; self-report; patient portals; patient-centered care; drug monitoring; adverse effects

Introduction

Rationale

Use of the internet has increased considerably since the early 1990s. The World Bank reports that almost 44% of people across the globe used the internet in 2015, compared with 0.25% in 1993 [1]. This number is expected to increase to over 50% by 2019 [2]. Nearly two-thirds of internet users are estimated to access health information on the Web [3]. With such demand for accessible health information, electronic health (eHealth) has become a popular way to provide patients with health information, recommendations to self-manage their health, and access to their health records and data [3,4]. eHealth is defined as "an overarching term used today to describe the application of information and communications technologies in the health sector. It encompasses a whole range of purposes from purely administrative through to health care delivery" [5]. eHealth tools, therefore, are technologies that may include electronic medical records (EMRs), personal health records (PHRs), mobile apps, patient portals, information repositories, and many other internet-based programs or software used to help patients monitor and manage their health. eHealth tools may help decrease fragmentation of care by compiling patient health information from multiple providers into 1 easily accessible location [6], while also streamlining patient-provider communication and promoting shared decision making [3,4].

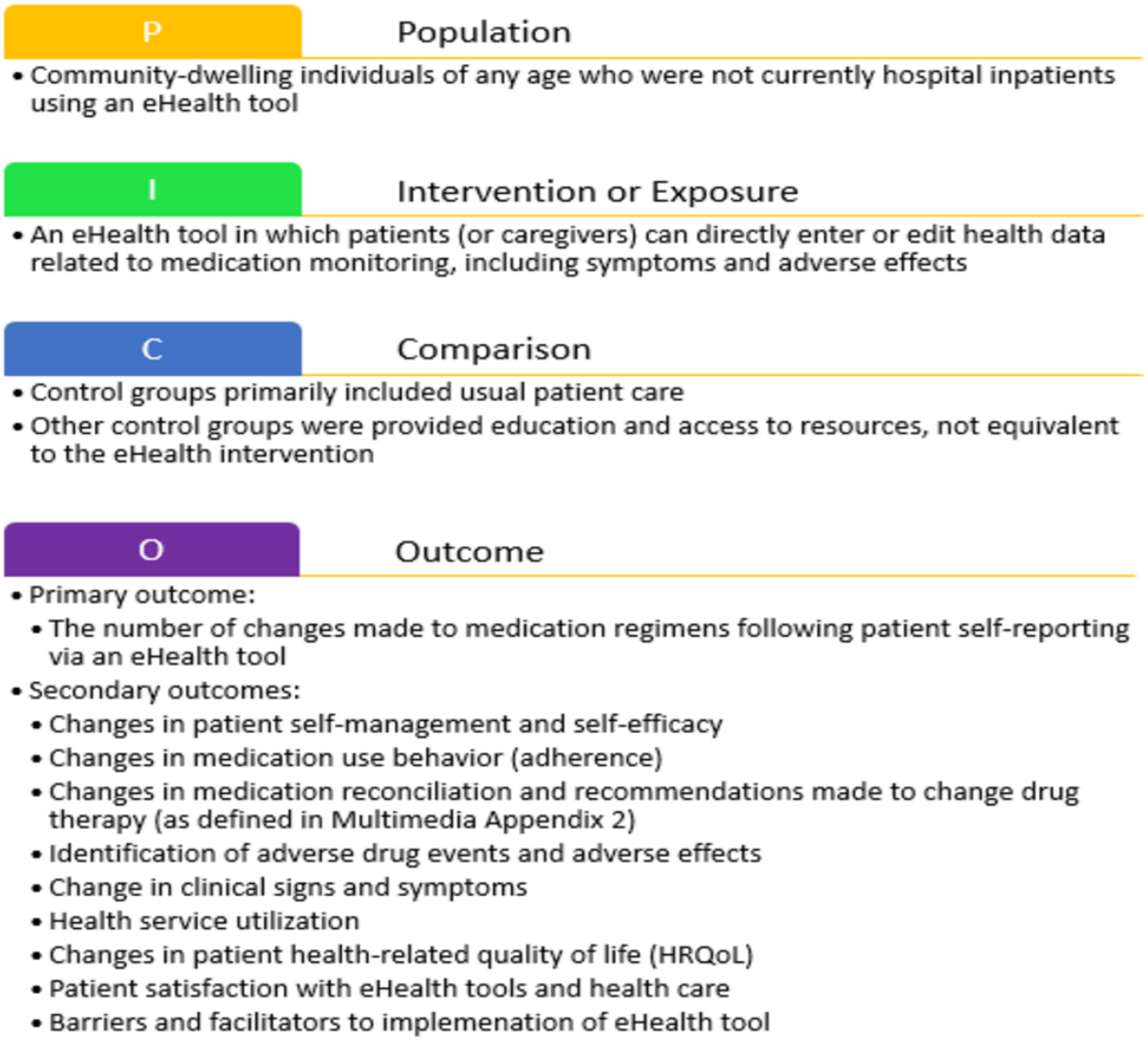
Well-functioning eHealth tools can help patients better understand their health [7] and may lead to improvements in patient-physician relationships [8]. eHealth tools can encourage patients to play a larger role in shared decision making and might increase focus on self-management and preventative care [8,9]. As technology advances, the use of eHealth tools can provide a level of convenience for both patients and providers [10]. These tools can generally be accessed from any internet-capable device and often provide a method of asynchronous communication such as emails and short message service (SMS) text messaging. These methods allow patients and providers to ask and answer questions at their convenience, creating less of a burden on physician workflow [8].

The ability of patients to use eHealth tools to better manage medication by reporting feedback on symptoms and use of medications directly to health care providers has not been comprehensively explored in the literature. Similarly, there is little evidence showing how eHealth tools might provide prompts to patients and health care providers to make appropriate changes to medication use based on this feedback. A synthesis of this literature will provide greater understanding of what eHealth tool design features may be helpful in patient reporting of medication-related experiences and outcomes.

Objective

The objective of this systematic review was to determine the impact of patients' use of eHealth tools on self-reporting adverse effects and symptoms that promote changes to medication use. The PICO model was used to focus the objective of the review, as seen in Figure 1.

Figure 1. Use of the PICO model in this systematic review. eHealth: electronic health.



Methods

Study Design and Study Selection

This systematic review was performed following steps outlined by Cochrane's Effective Practice and Organization of Care group and reported based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement [11]. A total of 3 biomedical and health science databases were searched: MEDLINE/Ovid, EMBASE/Ovid, and CINAHL. References of all included articles were also searched. All 3 databases were searched from January 1, 2000, to April 25, 2018. The search was limited to articles published in English using terms representing eHealth (eg, Web-based applications), symptoms and adverse drug reaction reporting (eg, drug-related adverse effects and adverse reactions), and patient self-monitoring (eg, self-management; see Multimedia Appendix 1 for full search strategy and Multimedia Appendix 2 [12-17] for definitions for terminology used). The search date began from 2000, which generally marks the start of scientific reporting of eHealth interventions that would have relevance to the current use of eHealth tools. As a result of the aforementioned search strategy, studies were included in this review if they determined the effectiveness and impact of changes to medication regimens as a result of using eHealth tools. As such, this review investigated these effects using a comparative quantitative methodological approach.

Criteria for Inclusion of Studies

For the purposes of this review, an eHealth tool was considered to be any internet-based intervention, including mobile health apps, used by patients for clinical purposes that focused on improving patient health and clinical outcomes. The term PHR refers to an eHealth tool wherein a patient has access to and can enter or edit their own health data. The population investigated was community-dwelling individuals of any age in an outpatient setting.

For a study to be included, the eHealth tool must have allowed patients (or caregivers) to enter information directly (as opposed to information being entered by a health care provider); included self-reporting functionalities focusing on medication monitoring, contain a medication monitoring or use component, or specifically incorporating the option for the patient or caregiver to enter symptoms including adverse effects; and needed to focus specifically on medication use, clinical outcomes, or

symptom reporting following use of the eHealth tool. Any eHealth tools involving changes in medication reconciliation and recommendations made to changes in drug therapy were also included.

Exclusion criteria were conference abstracts; qualitative studies; articles without a comparator group; articles that did not report on at least one medication-related outcome; articles where self-management strategies focused on lifestyle modification, behavioral interventions, or nondrug interventions; articles focused solely on the validation of an eHealth tool; articles focused on methodological or technical aspects of eHealth interventions; articles containing nonempirical information; articles that synthesized information about multiple eHealth tools in an article (ie, review articles); and eHealth tools used by regulatory agencies to report adverse drug events (ADEs).

Article Selection

All potentially relevant articles were uploaded into DistillerSR software, which was used throughout the selection process. Potentially relevant articles underwent title, abstract, and full-text review. Articles that met inclusion criteria proceeded to data abstraction and risk of bias assessment. Articles not meeting inclusion criteria were excluded at both levels. Figure 2 represents the flow of articles through the selection process.

Title and abstract review were performed independently by 2 reviewers from a pool of 5 reviewers. Of these, 1 reviewer went through the reference lists of all the articles included in this study. Another reviewer went through reference lists of relevant systematic reviews identified during the literature search. Potentially relevant articles were identified. These articles went through abstract review by 2 reviewers. Studies found not to fit inclusion criteria after abstract review were excluded. Full-text review was performed independently by 5 reviewers. The kappa scores were calculated to determine agreement among reviewers who conducted review of titles and abstracts. All kappa scores calculated were greater than .93. Conflicts were resolved by consensus.

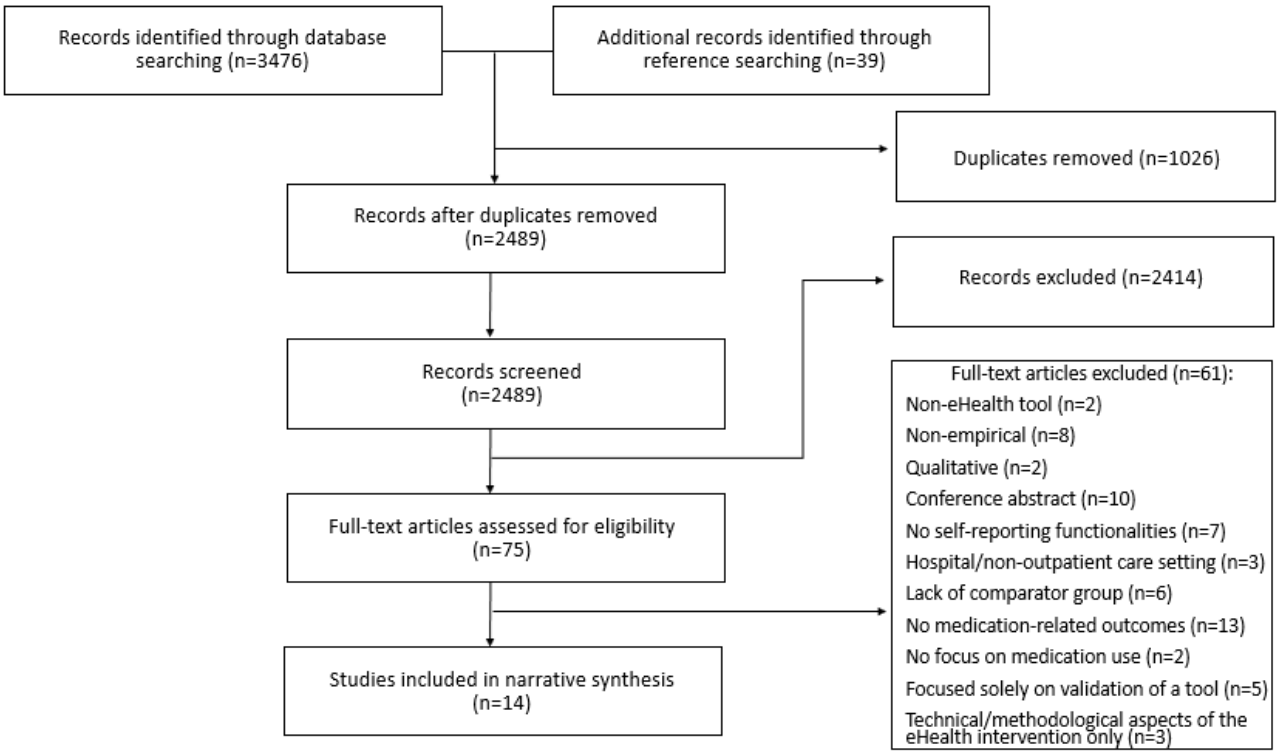
Data Extraction and Risk of Bias Assessment

Data extraction and risk of bias assessment were performed for each study independently by 2 reviewers. Data extracted included study design and setting, participant demographics, number of participants in each group, intervention components, comparator group components, eHealth tool functionality measured, and results and significance levels for each outcome measure. Conflicts in data extraction were resolved by consensus. Risk of bias assessment used questions recommended by the Agency for Healthcare Research and Quality's 2014 publication Methods Guide for Effectiveness and Comparative Effectiveness Reviews [18] and was performed for each study independently by 2 reviewers. All conflicts were resolved by consensus. The risk of bias assessment questions are presented in Multimedia Appendix 3.

Outcomes and Analysis

The primary and secondary outcomes are listed in Figure 1. The primary outcome was the number of changes made to medication regimens following patient self-reporting via an eHealth tool. The included studies varied considerably in populations, eHealth tool functionality, outcomes measured, and study design. Due to high heterogeneity, meta-analysis of outcomes was not feasible. Therefore, results for each outcome were synthesized descriptively and presented as narrative. Available data on barriers to implementation were extracted from the article text and summarized qualitatively so as to heighten awareness of implementation issues.

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses style study inclusion flowchart. eHealth: electronic health.



A Priori Subgroup and Sensitivity Analyses

Subgroup analyses were performed to investigate differences in treatment effect present because of (1) age of participants, (2) patients with specific conditions targeted by intervention, and (3) different features and functionalities of the included eHealth tools.

Results

Included Studies

A total of 3515 articles were generated from database and reference searching, resulting in 2489 potential articles that were screened based on their titles and abstracts, after duplicates were removed. Furthermore, 75 full-text articles were assessed for eligibility, of which 14 were included in this systematic review (see Figure 2 for more details).

Of the included articles, 13 were randomized controlled trials (RCTs) [19-31] and 1 was an open-label intervention [32]. A total of 10 studies were conducted in the United States [20-29], and 1 study was conducted in each of South Korea [19], Canada [30], Finland [31], and Denmark [32]. Dates of publication ranged from 2006 to 2017. The majority were published in 2007 or later (n=13). This distribution mirrors the increase in both internet and eHealth tool usage beginning in the late 2000s [1,33]. Further details on the characteristics of these studies can be seen in Table 1. Details regarding the design and outcomes of included studies are presented in Multimedia Appendix 4.

Table 1. Characteristics of included studies.

First author (year)	Country	Study design	Sample size	Patient age group study population
Cho (2006) [19]	South Korea	RCT[a]	80	Adults; adults (aged >30 years) with type II diabetes
Chrischilles (2014) [20]	United States	RCT	1075	Elderly (aged >65 years); patients using a computer in the past month to visit websites or to send or receive email
Fiks (2015) [21]	United States	RCT	60	Children (aged 6 to 12 years) and parents
Grant (2008) [22]	United States	Cluster RCT	11 sites; 244 patients	Adults; adults with type II diabetes, A1c ≥7% or ≥1 diabetes medication, with ≥1 primary care visit in the last year and enrolled in Patient Gateway
Gustafson (2012) [23]	United States	RCT	305	Children (aged 4 to 12 years); patients with poorly controlled asthma and parents
Joseph (2007) [24]	United States	RCT	314	Children and young adults (ninth to eleventh grade); students with an asthma diagnosis or meeting asthma criteria
Joseph (2013) [25]	United States	RCT	422	Children and young adults (ninth to eleventh grade); students meeting asthma criteria or with an asthma diagnosis
Schnipper (2012) [26]	United States	Cluster RCT	11 sites; 541 patients	Adults; adults with ≥1 primary care visit and enrolled in Patient Gateway
Simon (2011) [27]	United States	RCT	208	Adults (aged >18 years); depressive disorder diagnosis with new antidepressant treatment
Weingart (2013) [28]	United States	RCT	738	Adults (aged 18 to 87 years); patients enrolled in PatientSite and received at least one new medication
Mooney (2017) [29]	United States	RCT	6 sites; 358 participants	Adults, seniors; English-speaking adults with a life expectancy of ≥3 months, beginning chemotherapy consisting of at least three cycles with daily access to a telephone
Ahmed (2016) [30]	Canada	RCT	2 sites; 100 participants	Adults (aged 18 to 69 years); French- or English-speaking adults diagnosed with asthma, prescribed at least one rescue medication, have poor asthma control, access to internet, and smoking <20 pack-years
Karhula (2015) [31]	Finland	RCT	517 participants (267 heart disease and 250 diabetes)	Adults, seniors; ability to complete questionnaires in Finnish, use the RPM system/devices, adequate cognition, able to walk; type 2 diabetes (diagnosed at least 3 months earlier) with hemoglobin A1c >6.5% within 1 year before screening; heart disease group (ischemic heart disease or heart failure)
Carlsen (2017) [32]	Denmark	Open-label	One site; 50 participants (29 electronic health tools, 21 control)	Children, adolescents; aged 10 to 17 years with ulcerative colitis or Crohn disease on maintenance infliximab treatment at the Department of Pediatrics, Hvidovre Hospital

[a]RCT: randomized controlled trial.

Of the 13 RCTs included in this review, 2 studies were cluster RCTs [22,26]. The remaining 11 RCTs used participants as the unit of randomization [19-21,23-25,27-31].

A total of 4 RCTs focused on pediatric and adolescent asthma patients [21,23-25]; 1 study focused on adult asthma patients [30]. Moreover, 1 trial included only elderly patients (aged 65 years and older) and focused on medication self-management and safety [20]. The remaining 7 trials all included both adult and elderly participants. From these studies, 1 focused on patients with depression [27]; 3 on patients with type 2 diabetes [19,22,31]; 2 on medication safety and use, including identification of ADEs [26,28] using eHealth tools; and 1 on identifying adverse effects in patients receiving chemotherapy [29].

In 3 studies, use of an eHealth tool in the intervention group was compared with usual care plus links to relevant websites [24,25,28]. Simon et al [27] compared their Web-based messaging eHealth tool for depression and Web-based messaging system to usual care with Web-based messaging between patients and health care providers. Gustafson et al [23] used nearly identical interventions in both groups; the control group was restricted from accessing the eHealth tool but participated in other aspects of the intervention (clinical visits, interviews, patient education, etc). Fiks et al [21] also used a usual care group with no access to the intervention Web portal; however, all health care providers had access to a computerized decision support system. Moreover, 2 studies [22,26] utilized a *double-dummy* style intervention, where both groups used Web-based PHRs to record information that differed only in content. Cho et al [19] compared an electronic blood glucose (BG) monitoring system with an informal paper-based monitoring system, with both groups receiving diabetes education and regular clinical visits. Chrischilles et al [20] utilized a conventional usual care group without supplementary information or resources. Mooney et al [29] used a self-monitoring tool to manage chemotherapy symptoms. Ahmed et al [30] developed an asthma portal to view patient's personal health information, monitor patients, and provide feedback on self-management strategies. Karhula et al [31] used a management system for patient self-monitoring of diabetes. Only 1 study was identified as an open-label intervention study, which included a comparator group [32]. Carlsen et al [32] used an eHealth tool to monitor responses of patients with inflammatory bowel disease to determine the need to adjust treatment interval or dose.

Quality of Included Studies

Figure 3 displays a summary of the risk of bias assessment. The studies, overall, were of moderate quality, with studies ranging from poor to good. Common issues included small numbers of participants, lack of blinding, poor description of interventions, and contamination of intervention. Many trials relied only on patient self-reported data (as would be expected based on the topic), which can introduce bias if methods to ensure validity and reliability are not demonstrated.

Types of Electronic Health Tools

From the 14 included studies, 11 unique eHealth tools were described. The 2 RCTs by Joseph et al [24,25] utilized the same asthma management eHealth tool. The cluster RCT by Schnipper et al [26] used a Web-based PHR to record information that differed only in content, which was nested within the larger RCT by Grant et al [22]. Two studies by Fiks and Ahmed used a Web-based portal for asthma symptom management [21,30]. Each study and eHealth tool is described in Multimedia Appendix 4. Features and functionalities of the eHealth tools are also presented in Table 2. Although Schnipper [26] and Grant [22] use the same eHealth tool, Schnipper's study [26] investigates a specific medication module. Thus, they have been counted separately here.

All 11 eHealth tools from all 14 studies included a component where patients could self-report medication management information or changes, including symptoms, health data, adverse effects, or ADEs. A total of 12 studies included Web-based patient questionnaires or surveys [20,21,23-32]. Many studies used validated patient questionnaires but several developed their own. A list of patient questionnaires utilized by each study is presented in Table 3. In addition, 10 eHealth tools included patient educational resources [20,21,23-26,28-31]. Details of these resources are also listed in Table 3. Cho et al [19] measured outcomes by having a patient record their BG readings in the eHealth tool. Patients were also interviewed in person by clinicians every 3 months. Schnipper et al [26] had participants complete medication electronic journals, where they would note discrepancies between a Web-based medication list and their actual medications as taken. Karhula et al [31] used a 36-question survey for patients to report their health-related quality of life (HRQoL) score at baseline and post intervention.

Outcomes of Included Studies

The results of each study by outcome can be seen in Multimedia Appendix 5.

Primary Outcome: Changes in Use of Medications and Other Therapies

A total of 6 RCTs [19-22,24,27] and 1 open-label intervention [32] measured this outcome. Moreover, 5 RCTs [20,22,24,27,32] found significant increases in medication changes as a result of using eHealth tools. All medication change outcomes reported were consistent with more appropriate prescribing and use of medications.

Chrischilles et al [20] found a significant reduction in use of more than 2 nonsteroidal anti-inflammatory drugs in the intervention group (14.1% vs 19%, P=.035). A nonsignificant trend approaching significance was seen for decreased number of over-the-counter medications used in the eHealth tool group (P=.05) [20]. Grant et al [22] found a significant increase in the number of diabetes mellitus–related medication changes in the intervention group (43.5 vs 6.2, P<.001). They also found that a significantly higher proportion of patients in the intervention group had medications initiated or dosages changed for hypertension (13% vs 0%, P=.02) and hyperlipidemia (11% vs 0%, P=.03).

Figure 3. Summary chart of risk of bias assessment for included studies. Risk of bias summary: Green "+" symbols indicate a low risk of bias, yellow "?" symbols indicate an unknown risk of bias, and red "-" symbols indicate a high risk of bias.

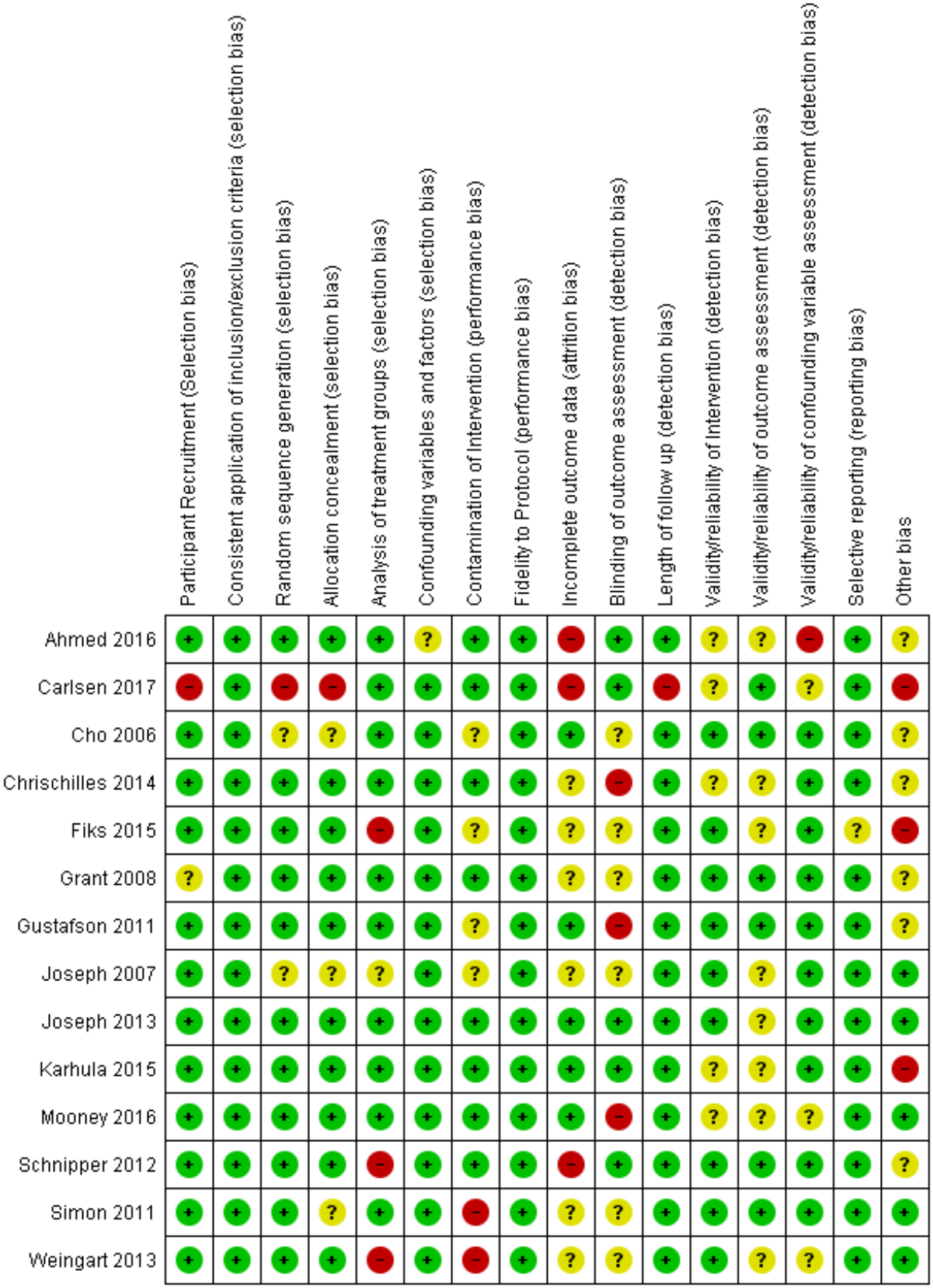


Table 2. Features and functionalities of electronic health tools.

Topic; study author, year	Linked to electronic medical record	Function as personal health record	Clinicians can view self-reported information	Messaging between patient and clinician	Web-based surveys or questionnaires	Web-based drug list	Web-based access to lab results	Patient prompts or reminders	Patient educational resources
Focus on medication safety and usage									
Chrischilles, 2014 [20]	✗[a]	✓[b]	✗	✗	✓	✓	✗	✓	✓
Schnipper, 2012 [26]	✓	✓	✓	✓	✗	✓	✗	✗	✓
Weingart, 2013 [28]	✓	✓	✓	✓	✓	?[c]	✓	✓	✓
Focus on pediatric and adolescent asthma patients									
Fiks, 2015 [21]	✓	✓	✓	✗	✓	✓	✗	✓	✓
Gustafson, 2012 [23]	✗	✓	✓	✓	✓	✗	✗	✓	✓
Joseph, 2007 and 2013 [24,25]	✗	✗	✗	✗	✓	✗	✗	✗	✓
Focus on adult asthma patients									
Ahmed, 2016 [30]	✓	✓	✓	✓	✓	✓	✗	✓	✓
Focus on cancer patients									
Mooney, 2017 [29]	✓	✓	✓	✗	✓	✓	✓	✓	✓
Focus on diabetic patients									
Cho, 2006 [19]	?	✓	✓	✓	✗	✓	✓	✗	✗
Grant, 2008 [22]	✓	✓	✓	✓	✗	✓	✓	✗	✗
Other									
Simon, 2011 [27]	✓	✓	✗	✓	✓	?	✓	✗	✗
Karhul, 2015 [31]	✓	✓	✓	✓	✗	✗	✗	✓	✓
Carlsen, 2017 [32]	✗	✓	✓	?	✓	✗	✓	✓	✗

[a]✗ is used to demonstrate that the feature or functionality is not present and mentioned in the article.

[b]✓ is used to demonstrate that the feature or functionality is present and mentioned in the article.

[c]? is used to demonstrate that the feature or functionality is not discussed in the article.

Table 3. Use of patient questionnaires and educational resources in included studies.

First author (year)	Patient questionnaires used	Patient educational resources
Cho (2006) [19]	N/A[a]	N/A
Chrischilles (2014) [20]	Morisky adherence measure for medication adherence (modified); Assessing Care of Vulnerable Elders (ACOVE-3) medication-use quality indicators (modified); 12-item short form health survey (SF-12) for health status; other surveys developed by the study team [34,35]	ACOVE-3 adapted into patient medication safety messages [36]
Fiks (2015) [21]	Parent Patient Activation Measure; Integrated Therapeutics Group Child Asthma Short Form; Asthma Control Tool (ACT); other questions developed by the study team [37-39]	Handouts and videos available, but source and items used not reported
Grant (2008) [22]	Not reported by the study team, focused on medication adherence barriers	N/A
Gustafson (2012) [23]	Asthma Control Questionnaire; other questionnaires developed by the study team [40,41]	On the basis of the National Asthma Education and Prevention Program guidelines [42-44]
Joseph (2007) [24]	Lung Health Survey, developed by the study team, using items from the International Survey of Asthma and Allergies in Childhood questionnaire (ISAAC), and National Asthma Education and Prevention Program guidelines "Guidelines for the Diagnosis and Management of Asthma: Expert Panel Report II" (EPRII; adapted) [45,46]	EPRII; resources identified by Croft et al [46,47]
Joseph (2013) [25]	Multidimensional Scale of Perceived Social Support (adapted); Diagnosis Interview Schedule for Children Predictive Scales; Lung Health Survey, developed by the study team, using items from ISAAC; EPRII (adapted); EPRIII (adapted) [44-49]	EPRII; resources identified by Croft et al [46,47]
Schnipper (2012) [26]	Questionnaires developed by study team (adjudicated by physicians)	Source and items used not reported
Simon (2011) [27]	Hopkins Symptom Checklist; Patient Health Questionnaire (PHQ) Depression questionnaire; other questions developed by the study team [50-52]	N/A
Weingart (2013) [28]	Questions developed by the study team regarding new prescriptions and symptoms or adverse drug events	National Patient Safety Foundation website [53]
Mooney (2017) [29]	Questionnaire about symptoms severity for 11 symptoms related to chemotherapy	Self-management coaching provided based on symptoms; nurse practitioner follow-up, if required within 4 hours
Ahmed (2016) [30]	Mini-Asthma Quality of Life Questionnaire ; Chronic Disease Self-Efficacy Scale; ACT; Beliefs about Medicines Questionnaire; 9-item PHQ; EuroQol visual analog scale	MyAsthma Portal
Karhula (2015) [31]	SF-36 health survey	Patients provided a self-management guide. Additionally received health coaching phone calls
Carlsen (2017) [32]	IMPACT III (pediatric inflammatory bowel disease health-related quality of life measure); Total Inflammation Burden Score: (pediatric ulcerative colitis activity index/abbreviated Pediatric Crohn Disease Activity Index + fecal calprotectin)	N/A

[a]N/A: not applicable.

Joseph et al [24] found that a significantly higher proportion of those in the active group had a rescue inhaler available (39% vs 32%, *P*=.01). Simon et al [27] found that a significantly higher proportion of participants in the active group used antidepressants for an appropriate length of time (≥90 days; χ^2_1=10.5, *P*=.001). Carlsen et al [54] showed that eHealth tools might help identify instances where medication changes may be appropriate. Moreover, Carlsen et al [54] showed that the eHealth tool used resulted in at least one significant effect on changes to medication use; a significant difference was found between intervals of treatment for the eHealth group relative to the control group (2.35; 95% CI 1.5 to 3.2; *P*<.001) [54].

In contrast, Cho et al [19] found no significance in terms of total occasions of drug modification through the use of their eHealth tool intervention, internet-based glucose monitoring system (4.7 vs 5.5, *P*=.36). Fiks et al [21] provided descriptive evidence regarding the mean number of medications per child in both the intervention and control groups, yet no between-groups comparisons were made.

Secondary Outcome: Changes in Signs and Symptoms of Health Conditions

A total of 9 RCTs [19,21-25,27,30,31] and 1 open-label intervention [32] measured changes in signs and symptoms of health conditions. Of these, 8 reported at least one significant improvement in signs and symptoms [19,21,23-25,27,30,31].

Moreover, 4 studies found improvements in asthma symptoms [21,23-25], 1 found a significant reduction in hemoglobin A1c (HbA_{1c}) [19], and 2 found a significant improvement in depression score [27,30].

Fiks et al [21] reported 17 instances of uncontrolled asthma in 13 children. They found that parents of active group children missed fewer days of work (mean of <0.1 vs 1.5, *P*=.001) and that the active group had less frequent flare-ups (mean of 1.4 vs 3.8, *P*=.02). Gustafson et al [23] found an increase in asthma control in the active group (mean change of −0.42 vs −0.11 on a 7-point Likert scale, *P*=.01).

Joseph et al [24] found that the active group had a lower risk for number of symptom nights (risk ratio [RR]=0.4, 95% CI 0.2 to 0.8, *P*=.009), symptom days (RR=0.5, 95% CI 0.4 to 0.8, *P*=.003), days of restricted activity (RR=0.5, 95% CI 0.3 to 0.8, *P*=.02), and school days missed (RR=0.3, 95% CI 0.1 to 0.7, *P*=.006). In another study, Joseph et al [25] reported a lower risk in the active group for symptom days (RR=0.8, 95% CI 0.6 to 1.0, *P*=.01). Following subgroup analysis, it was found that teenagers with moderate to severe asthma had fewer symptom days (RR=0.6, 95% CI 0.5 to 0.9, *P*=.01), total school days missed (RR=0.5, 95% CI 0.3 to 0.8, *P*=.009), school days missed because of asthma (RR=0.4, 95% CI 0.2 to 0.8, *P*=.007), and days of restricted activity (RR=0.6, 95% CI 0.4 to 0.9, *P*=.03).

Simon et al [27] reported a significant between-groups difference in mean depression score favoring the intervention group (mean score of 0.95 vs 1.17, *P*=.04). Ahmed et al [30] reported a statistically significant difference in depression using the Patient Health Questionnaire scale, as scores improved at 6 months (mean change −0.27, 95% CI −0.37 to −0.18 for a change of 5 units). Karhula et al [31] found no significant between-group difference in HbA_{1c} (change −0.106, 95% CI −0.33 to 0.11, *P*=.34); however, they did find a statistically significant decrease in waist circumference between intervention and control (change −1.711, 95% CI −3.042 to −0.38, *P*=.01). Cho et al [19] did find a decrease in mean HbA_{1c} in the active group (mean of 6.7% vs 7.4%, *P*=.006) at 30 months. Grant et al [22] found no differences between groups for HbA_{1c} and for percentage of patients at target HbA_{1c} levels. Similarly, Carlsen et al [32] found no significant difference for trough infliximab concentration when controlling for treatment intervals in the study (change of −2.19, 95% CI −5.37 to 0.99, *P*=.18).

Secondary Outcome: Patient Self-Management and Efficacy

A total of 5 RCTs measured this outcome [20,21,23,26,30]. In addition, 3 of the studies found increases in patient self-management or self-efficacy as a result of using eHealth tools [23,26,30].

Gustafson et al [23] found that self-efficacy had a positive, significant effect on Asthma Control Questionnaire (ACQ) scores (beta=.48, *P*=.01). They also found a positive significant effect of intervention on ACQ score when mediated by information competence (τ=−.235, *P*=.02). Schnipper et al [26] found that significantly more participants in the intervention group always disclosed drug therapy problems or new symptoms to clinicians (97.9% vs 87.1%, *P*<.001). Ahmed et al [30] found that a significant change in minimum asthma-related quality of life questionnaire adjusted for self-efficacy in adult asthma patients (0.24, 95% CI 0.16 to 0.32). Fiks et al [21] found that parents of children with asthma who used the eHealth tool improved their ability to manage asthma, although their findings were not statistically significant. In addition, they became more aware of the importance of ongoing attention to treatment. Chrischilles et al [20] found no difference between groups in ability to recognize adverse effects; however, in their as-treated analysis, they did find that high-frequency users had higher odds of recognizing symptoms and adverse effects (odds ratio, OR=1.76; 95% CI 1.08 to 2.86).

Secondary Outcome: Medication Use Behavior (Adherence)

A total of 3 RCTs measured this outcome, all using measures of medication adherence as a surrogate for medication use behavior [20,23,24]. None of these studies reported improvements over the 6 [20] and 12 months [23,24] studied.

Secondary Outcome: Medication Reconciliation and Recommendations to Change Drug Therapy

A total of 3 RCTs reported on this outcome [19,20,26]. Only Schnipper et al [26] found improvements in determining medication discrepancies when linking documented and patient-reported medication regimens using eHealth tools. Schnipper et al [26] explored the effects of a PHR medication module on medication accuracy and safety, reporting significantly lower odds of having discordant medications in the active group (OR=0.71, 95% CI 0.54 to 0.94, *P*=.01). In addition, Schnipper et al [26] found a significantly lower risk of discrepancies with the potential to cause severe harm in the active group (RR=0.31, 95% CI 0.10 to 0.92, *P*=.04). The number of medication discrepancies per patient with the potential for harm approached significance as a result of using eHealth tools (*P*=.05) [26].

Cho et al [19] acknowledged, using descriptive data, that a small percentage of individuals may have recommendations made for modification of drug therapy as a result of using their eHealth tool. Chrischilles et al [20] reported several instances of recommendations being made to alter drug therapy through medication reconciliation, none of which were significant.

Secondary Outcome: Adverse Effects and Adverse Drug Events

A total of 1 open-label intervention [32] and 5 RCTs measured this outcome [20,21,26,28,29]. Only Mooney et al [29] reported identification of adverse effects in favor of using eHealth tools, as there was a significant reduction in 10 of the 11 chemotherapy adverse effects in the intervention group (*P* value: .02 to <.001) relative to usual care. Descriptive evidence from Fiks et al [21] showed 1 instance of medication-related adverse effects. Carlsen et al [32] provided descriptive evidence that eHealth tools may lead to the identification of ADEs. The remaining 3 studies reporting on the identification of adverse effects [20] or ADEs [26,28] found no significant difference between intervention and control.

Secondary Outcome: Health Services Utilization

A total of 6 RCTs reported health service utilization outcomes [21,24,25,27,28,30]. Of these, only Joseph et al [24] found a significantly lower risk of hospitalizations as a result of using eHealth tools (RR=0.20, 95% CI 0.2 to 0.9, *P*=.01). The remaining 5 studies [21,25,27,28,30] found no difference between intervention and control groups in terms of health service utilization

Secondary Outcome: Patient Self-Reported Overall Health Status

A total of 4 studies measured this outcome [24,30-32], and none of the studies found differences in cumulative quality of life score between groups or a significant effect on overall health status.

Secondary Outcome: Patient Satisfaction With Health Care

A total of 2 RCTs [21,27] and 1 open-label intervention [32] measured this outcome. Only Simon et al [27] found that patient satisfaction improved as a result of using an eHealth tool. Simon et al [27] found that a significantly larger proportion of participants in the intervention group reported being very satisfied with the quality of their depression-related care (χ^2_1=8.38, *P*=.004). Fiks et al [21] found no significant changes when measuring this outcome, whereas Carlsen et al [32] provided positive descriptive evidence of patient and parent satisfaction using eHealth tools.

Subgroup Analyses

A total of 4 RCTs investigated the use of eHealth tools in children and teens with asthma [21,23-25]. There is evidence that eHealth tools may have the potential to reduce symptoms of asthma, frequency of asthma flare-ups, and number of days of school or work missed because of asthma [21,23-25]. They may also promote better asthma control, availability and use of rescue inhalers, and may have the potential to improve asthma symptoms in vulnerable groups (ie, African-American adolescents living in urban centers) [24,25].

Subgroup analysis also found that multifaceted interventions combining use of eHealth tools with clinician support or case management and eHealth tools utilizing direct patient-provider communication might be more effective at improving some aspects of patient self-management and self-efficacy [23,26,30,31]. Both studies utilizing multifaceted interventions and direct patient-provider communication that measured these outcomes found positive significant results [23,26], whereas both studies using only eHealth tools with no patient-provider communication found no significant differences [20,21]. Detailed results from the subgroup analyses can be found in Multimedia Appendix 4.

Barriers to Implementation

Many studies reported barriers to the implementation of eHealth tools. The most common barrier was lack of participant engagement, resulting in low eHealth tool utilization rates. This was reported by 9 of the 14 studies [20,22-27,31,32]. A total of 3 studies noted distinct differences between high and low eHealth tool users [19,20,26], with high users generally seeing more improvements in health-related outcomes. Chrischilles et al's study [20] was the only study to investigate the use of eHealth tools specifically in patients aged more than 65 years, and they found that patient engagement was negatively associated with age. Grant el al [22] found that patients with poor metabolic control were less likely to participate. The authors of several studies reported that a small sample size, high level of missing data, reduced power, and lower generalizability were observed as a result of low eHealth tool utilization and patient engagement [20,22,23,25-28,31,32]. Another important barrier, reported by 3 studies [22,26,29], was lack of clinician engagement and poor clinician training. This was generally because of time and workflow constraints [22,29] and lack of motivation [26,29]. Other implementation issues noted included lack of access to the internet [30], time burden of entering information [30], poor usability of eHealth tools [26], difficulties obtaining informed consent [24,25], and dilution of the intervention effect by the control group [23,26,28].

Discussion

Summary of Evidence

Evidence from 4 RCTs and 1 open-label intervention [20,22,24,27,32] show that eHealth tools focusing on symptom and adverse effect self-reporting can prompt positive changes in medication prescribing and use. In addition, the majority of eHealth tools studied were able to improve patient symptoms, regardless of functionalities, complexity, and differences in intervention. Those eHealth tools were particularly studied and found to be beneficial for improving signs and symptoms in children and adolescents with asthma [21,23-25]. This review supports the inclusion of patient entry or editing of symptoms into eHealth tools for the purposes of monitoring and reporting outcomes from the use of medications.

Evidence was found that eHealth tools improved the outcome of patient self-management and self-efficacy. Subgroup analysis found that eHealth tools that allow patients and clinicians to communicate directly, and multifaceted interventions combining eHealth tools with clinician support and case management might lead to greater increases in patient self-management and self-efficacy. It is notable that more significant improvements were found for more objective outcome measures, such as number of medication changes and clinical signs, and less were found for more subjective outcome measures such as self-management and self-efficacy. It may also be that sample sizes were too small to detect differences, particularly if this was not the primary objective for these studies. It is likely that the eHealth tools under investigation either did not provide effective content or functionalities to help participants improve self-management and medication management in participants or the tools used to measure these outcomes were not able to detect any differences between groups. Another possibility is the lack of patient understanding of chronic disease and poor perception of health goals.

It was thought that eHealth tools that focus on improvement of patient self-efficacy and self-management might lead to improved medication-use behavior, which in turn may lead to

changes in medication use, identification of real or potential ADEs, improvement in signs and symptoms, and overall improvement in HRQoL. However, there is not enough evidence to draw conclusions as to the effectiveness of eHealth tools for identification of adverse effects, improving medication-use behavior, increasing recommendations to medication therapies and improving medication reconciliation, improving health service utilization, and improving overall health status and patient satisfaction. Only a small number of included studies investigated these outcomes; it is likely that with such a small overall sample size, it was not possible to find differences between groups.

How Do These Results Compare With Other Reviews?

As with most systematic reviews on the subject of eHealth tools [7,8,33,54-57], this review found at best moderate evidence that patient reporting via eHealth tools can lead to improved clinical outcomes such as symptom reduction.

A 2012 systematic review by Ammenwerth et al found that use of patient portals linked to a PHR led to significant increases in medication adjustments in diabetic patients [55]. Other reviews and primary articles have also indicated that the use of eHealth tools may be more effective in specific patient populations such as patients with cancer [4,29,34,58]. This review found evidence that use of eHealth tools might increase the number of medication adjustments in diabetic patients [19,22]. In addition, patient-reported symptoms and adverse effects were used to identify toxicities in cancer patients, and in several instances, it lead to medication changes. It was also found that eHealth tools might improve signs and symptoms of asthma in children and teens [21,23-25].

Overall, this review supports findings by Ammenwerth et al [55] that interventions may be more effective at improving health outcomes if they combine eHealth tool features such as patient-provider communication and interactive coaching with eHealth tool use (see, eg, [21,22,26,28] as well as Table 2 and Multimedia Appendix 4 considering subgroup analyses). Evidence from this review indicates that eHealth tools in combination with clinician support or case management, and eHealth tools that encourage provider-patient communication may improve patient self-management and self-efficacy when compared with tools without these features [22,23,26].

Strengths and Limitations

To our knowledge, this is the first systematic review of eHealth interventions focusing on patient self-reporting of symptoms and adverse effects. The review's search strategy was augmented by reference searching. This review was limited to studies that included medication-related outcomes. The majority of studies included in this review were RCTs, most of which were of moderate quality.

This review also has a number of weaknesses. It was limited to studies published in English, which may have excluded relevant articles. No searching of grey literature was performed, as we focused on empirical work published in academic journals, and so it is possible that some early non–peer-reviewed reports may have been missed. We acknowledge the lack of definitional clarity surrounding the term *eHealth* and believe future research should focus on establishing better consensus for this term. There was considerable variety among the interventions in the studies, some of which included features such as direct health care provider follow-up, thus making it more complicated to determine which outcomes could be specifically attributable to using an eHealth tool. As this review also examined different populations of varying sample sizes and medical conditions for eHealth tools, it may be difficult to detect differences and generalize findings and conclusions. We did not include qualitative studies in our review because the goal of our study was to better understand the effectiveness and impact of changes to medication regimens based on quantifiable differences in using eHealth tools for self-reporting adverse effects and symptoms that promote changes to medication use versus a comparator. We value the insight of qualitative studies that have been investigated elsewhere [59,60]. Additional exploration of qualitative literature to better understand how use of these types of eHealth tools can generate impacts on medication use and health would be helpful.

Implications for Practice and Future Research

Where possible, health care providers should encourage patient use of eHealth tools for symptom and adverse effect self-reporting. eHealth tools may be especially useful for reducing symptoms in certain populations, for example, children and teenagers with asthma. eHealth tools might also encourage patients to improve self-management behaviors and participate in shared decision making with clinicians. Having information from the EMR entered directly into the eHealth tool may reduce the burden on the patient to routinely update their clinical information (something that only highly motivated patients are likely to do regularly) [6]. Clinicians should be encouraged to communicate with patients via eHealth tools where possible, especially where patients are experiencing worsening of symptoms or medication-related adverse effects. Evidence suggests that using technologies such as mobile apps and SMS text messaging may improve patient engagement by allowing quick, convenient communication without a computer or internet connection [23,27]. Clinicians should be supported in their eHealth tool use, and interventions should focus on clinician training and engagement. Ensuring that interventions can be successfully incorporated into physician workflow is important [22,26].

There is a paucity of primary research articles investigating eHealth tools and their impact on medication use. Studies are generally small and of moderate quality. Large-scale RCTs focusing on the use of eHealth tools for medication and symptom management should be undertaken to establish more high-quality evidence. This is especially important given how ubiquitous the use of medication is. Furthermore, the effects of patient self-management and self-efficacy on medication use and symptom experience are not well studied; more research in this area could help drive creation of medication-focused eHealth tools. Low patient engagement and eHealth tool utilization were commonly noted implementation barriers; it could be that patients were not engaged in eHealth tool use enough for them to feel an impact on their satisfaction with health care or overall quality of life. Descriptive evidence shows low proportions of patients felt that eHealth tools improved

their care or communication with providers, indicating that development of eHealth tools should focus on functionalities and outcomes that are important to the patient. This may be achieved by utilizing research on patient motivation and behavior change to increase patient engagement [20,24,25].

Conclusions

The results of this review show initial and promising findings that specialized eHealth tools can be used for reporting and monitoring of symptoms and medication-related adverse effects and some evidence that use of eHealth tools have the potential to identify instances where changes in medication use may be appropriate. A modest amount of mixed evidence was found, demonstrating that eHealth tools can improve patient self-management and self-efficacy. Very little or no evidence was found to demonstrate that use of eHealth tools could increase numbers of medication recommendations or improve medication-taking behavior, health services utilization, identification of adverse effects, overall health status, and patient satisfaction. eHealth tools may be more effective at promoting medication changes and improving patient self-management and self-efficacy if they provide mechanisms for direct patient-provider communication and may be more effective in certain populations such as children and teenagers with asthma.

Acknowledgments

This study results from an Applied Health Research Question submitted by Canada Health Infoway to the Ontario Pharmacy Research Collaboration (OPEN). OPEN is funded by a grant from the Government of Ontario. The views expressed in this paper are those of the authors and do not necessarily reflect those of the funder.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search Strategy.

[PDF File (Adobe PDF File), 44KB - jmir_v20i12e294_app1.pdf]

Multimedia Appendix 2

Definitions.

[PDF File (Adobe PDF File), 122KB - jmir_v20i12e294_app2.pdf]

Multimedia Appendix 3

Risk of Bias Assessment Questions.

[PDF File (Adobe PDF File), 32KB - jmir_v20i12e294_app3.pdf]

Multimedia Appendix 4

Supplemental Tables for Subgroups and Details about Outcomes.

[PDF File (Adobe PDF File), 132KB - jmir_v20i12e294_app4.pdf]

Multimedia Appendix 5

Breakdown of Study Results by Outcome.

[PDF File (Adobe PDF File), 33KB - jmir_v20i12e294_app5.pdf]

References

1. World Bank. Individuals using the Internet (% of population) URL: https://data.worldbank.org/indicator/IT.NET.USER.ZS [accessed 2017-07-27] [WebCite Cache ID 6sHDSGfud]
2. The Broadband Commission for Digital Development. UNESCO. The state of broadband 2015: broadband as a foundation for sustainable development URL: http://unesdoc.unesco.org/images/0023/002346/234674e.pdf [accessed 2017-07-28] [WebCite Cache ID 6sIKmhYlS]
3. Caligtan CA, Dykes PC. Electronic health records and personal health records. Semin Oncol Nurs 2011 Aug;27(3):218-228. [doi: 10.1016/j.soncn.2011.04.007] [Medline: 21783013]
4. Government Of Canada. 2010 Aug 09. eHealth URL: https://www.canada.ca/en/health-canada/services/health-care-system/ehealth.html [accessed 2017-08-10] [WebCite Cache ID 6scWTCxLt]

5. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc 2011;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
6. 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] [Medline: 16357345]
7. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
8. de Lusignan S, Mold F, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic interpretative review. Br Med J Open 2014;4(9):e006021 [FREE Full text] [doi: 10.1136/bmjopen-2014-006021] [Medline: 25200561]
9. Hordern A, Georgiou A, Whetton S, Prgomet M. Consumer e-health: an overview of research evidence and implications for future policy. Health Inf Manag 2011;40(2):6-14. [Medline: 21712556]
10. Dean BB, Lam J, Natoli JL, Butler Q, Aguilar D, Nordyke RJ. Review: use of electronic medical records for health outcomes research: a literature review. Med Care Res Rev 2009 Dec;66(6):611-638. [doi: 10.1177/1077558709332440] [Medline: 19279318]
11. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-269. [Medline: 19622511]
12. Garets D, Davis M. HIMSS Analytics Database. Chicago, IL: HIMMS Analytics; 2006 Jan 26. Electronic medical records vs electronic health records: yes, there is a difference. A HIMSS Analytics white paper URL: https://www.aao.org/asset.axd?id=8e9b1f20-0ed6-4d2b-92f8-e28fbaf378ec&t=634962799822530000 [accessed 2017-08-25] [WebCite Cache ID 6sywQy3sY]
13. Upsalla Monitoring Centre. 2017 Feb 13. Common concepts and terms in pharmacovigilance URL: https://www.who-umc.org/safer-use-of-medicines/safer-use-of-medicines-the-basics/common-concepts-and-terms [accessed 2017-08-24] [WebCite Cache ID 6sxQPnTk6]
14. Barlow J. How to use education as an intervention in osteoarthritis. Best Pract Res Clin Rheumatol 2001 Oct;15(4):545-558. [doi: 10.1053/berh.2001.0172] [Medline: 11567538]
15. Wood R, Bandura A. Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. J Pers Soc Psychol 1989 Mar;56(3):407-415. [doi: 10.1037/0022-3514.56.3.407] [Medline: 2926637]
16. Canadian Patient Safety Institute. 2016. Medication Reconciliation (MedRec) URL: http://www.patientsafetyinstitute.ca/en/Topic/Pages/medication-reconciliation-(med-rec).aspx [accessed 2018-10-31] [WebCite Cache ID 73ZjwSpPB]
17. Council of Europe. World Health Organization. 2006 Oct 05. Glossary of terms related to patient and medication safety URL: http://www.who.int/patientsafety/highlights/COE_patient_and_medication_safety_gl.pdf [accessed 2018-10-31] [WebCite Cache ID 73Zjzpiok]
18. Agency for Healthcare Research Quality. Rockville, MD: Agency for Healthcare Research Quality; 2014 Jan. Methods Guide for Effectiveness and Comparative Effectiveness Reviews URL: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/cer-methods-guide_overview.pdf [accessed 2018-10-31] [WebCite Cache ID 73ZhL9gWU]
19. Cho J, Chang S, Kwon H, Choi Y, Ko S, Moon S, et al. Long-term effect of the internet-based glucose monitoring system on HbA1c reduction and glucose stability: a 30-month follow-up study for diabetes management with a ubiquitous medical care system. Diabetes Care 2006 Dec;29(12):2625-2631. [doi: 10.2337/dc05-2371] [Medline: 17130195]
20. Chrischilles EA, Hourcade JP, Doucette W, Eichmann D, Gryzlak B, Lorentzen R, et al. Personal health records: a randomized trial of effects on elder medication safety. J Am Med Inform Assoc 2014;21(4):679-686. [doi: 10.1136/amiajnl-2013-002284] [Medline: 24326536]
21. Fiks AG, Mayne SL, Karavite DJ, Suh A, O'Hara R, Localio AR, et al. Parent-reported outcomes of a shared decision-making portal in asthma: a practice-based RCT. Pediatrics 2015 Apr;135(4):e965-e973. [doi: 10.1542/peds.2014-3167] [Medline: 25755233]
22. Grant RW, Wald JS, Schnipper JL, Gandhi TK, Poon EG, Orav EJ, et al. Practice-linked online personal health records for type 2 diabetes mellitus: a randomized controlled trial. Arch Intern Med 2008 Sep 8;168(16):1776-1782 [FREE Full text] [doi: 10.1001/archinte.168.16.1776] [Medline: 18779465]
23. Gustafson D, Wise M, Bhattacharya A, Pulvermacher A, Shanovich K, Phillips B, et al. The effects of combining Web-based eHealth with telephone nurse case management for pediatric asthma control: a randomized controlled trial. J Med Internet Res 2012 Jul 26;14(4):e101 [FREE Full text] [doi: 10.2196/jmir.1964] [Medline: 22835804]
24. Joseph CL, Peterson E, Havstad S, Johnson CC, Hoerauf S, Stringer S, Asthma in Adolescents Research Team. A web-based, tailored asthma management program for urban African-American high school students. Am J Respir Crit Care Med 2007 May 01;175(9):888-895 [FREE Full text] [doi: 10.1164/rccm.200608-1244OC] [Medline: 17290041]
25. Joseph CL, Ownby DR, Havstad SL, Saltzgaber J, Considine S, Johnson D, Research team members. Evaluation of a web-based asthma management intervention program for urban teenagers: reaching the hard to reach. J Adolesc Health 2013 Apr;52(4):419-426 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.07.009] [Medline: 23299008]

26. Schnipper JL, Gandhi TK, Wald JS, Grant RW, Poon EG, Volk LA, et al. Effects of an online personal health record on medication accuracy and safety: a cluster-randomized trial. J Am Med Inform Assoc 2012;19(5):728-734 [FREE Full text] [doi: 10.1136/amiajnl-2011-000723] [Medline: 22556186]
27. Simon GE, Ralston JD, Savarino J, Pabiniak C, Wentzel C, Operskalski BH. Randomized trial of depression follow-up care by online messaging. J Gen Intern Med 2011 Jul;26(7):698-704 [FREE Full text] [doi: 10.1007/s11606-011-1679-8] [Medline: 21384219]
28. Weingart SN, Carbo A, Tess A, Chiappetta L, Tutkus S, Morway L, et al. Using a patient internet portal to prevent adverse drug events: a randomized, controlled trial. J Patient Saf 2013 Sep;9(3):169-175. [doi: 10.1097/PTS.0b013e31829e4b95] [Medline: 23965840]
29. Mooney KH, Beck SL, Wong B, Dunson W, Wujcik D, Whisenant M, et al. Automated home monitoring and management of patient-reported symptoms during chemotherapy: results of the symptom care at home RCT. Cancer Med 2017 Dec;6(3):537-546 [FREE Full text] [doi: 10.1002/cam4.1002] [Medline: 28135050]
30. Ahmed S, Ernst P, Bartlett S, Valois M, Zaihra T, Paré G, et al. The Effectiveness of Web-Based Asthma Self-Management System, My Asthma Portal (MAP): A Pilot Randomized Controlled Trial. J Med Internet Res 2016;18(12):e313 [FREE Full text] [doi: 10.2196/jmir.5866] [Medline: 27908846]
31. Karhula T, Vuorinen A, Rääpysjärvi K, Pakanen M, Itkonen P, Tepponen M, et al. Telemonitoring and mobile phone-based health coaching among Finnish diabetic and heart disease patients: randomized controlled trial. J Med Internet Res 2015 Jun 17;17(6):e153 [FREE Full text] [doi: 10.2196/jmir.4059] [Medline: 26084979]
32. Carlsen K, Houen G, Jakobsen C, Kallemose T, Paerregaard A, Riis LB, et al. Individualized infliximab treatment guided by patient-managed eHealth in children and adolescents with inflammatory bowel disease. Inflamm Bowel Dis 2017 Dec;23(9):1473-1482. [doi: 10.1097/MIB.0000000000001170] [Medline: 28617758]
33. Eland-de KP, van Os-Medendorp H, Vergouwe-Meijer A, Bruijnzeel-Koomen C, Ros W. A systematic review of the effects of e-health on chronically ill patients. J Clin Nurs 2011 Nov;20(21-22):2997-3010. [doi: 10.1111/j.1365-2702.2011.03743.x] [Medline: 21707807]
34. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [Medline: 3945130]
35. Shrank WH, Polinski JM, Avorn J. Quality indicators for medication use in vulnerable elders. J Am Geriatr Soc 2007 Oct;55(Suppl 2):S373-S382. [doi: 10.1111/j.1532-5415.2007.01345.x] [Medline: 17910560]
36. MacLean CH, Pencharz JN, Saag KG. Quality indicators for the care of osteoarthritis in vulnerable elders. J Am Geriatr Soc 2007 Oct;55(Suppl 2):S383-S391. [doi: 10.1111/j.1532-5415.2007.01346.x] [Medline: 17910561]
37. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
38. Bukstein DA, McGrath MM, Buchner DA, Landgraf J, Goss TF. Evaluation of a short form for measuring health-related quality of life among pediatric asthma patients. J Allergy Clin Immunol 2000 Feb;105(2 Pt 1):245-251. [doi: 10.1016/S0091-6749(00)90072-1}] [Medline: 10669843]
39. Zorc JJ, Pawlowski NA, Allen JL, Bryant-Stephens T, Winston M, Angsuco C, et al. Development and validation of an instrument to measure asthma symptom control in children. J Asthma 2006 Dec;43(10):753-758. [doi: 10.1080/02770900601031615] [Medline: 17169827]
40. Gustafson DH, Hawkins R, Pingree S, McTavish F, Arora NK, Mendenhall J, et al. Effect of computer support on younger women with breast cancer. J Gen Intern Med 2001 Jul;16(7):435-445 [FREE Full text] [Medline: 11520380]
41. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999 Oct;14(4):902-907 [FREE Full text] [Medline: 10573240]
42. Wise M, Gustafson DH, Sorkness CA, Molfenter T, Staresinic A, Meis T, et al. Internet telehealth for pediatric asthma case management: integrating computerized and case manager features for tailoring a Web-based asthma education program. Health Promot Pract 2007 Jul;8(3):282-291 [FREE Full text] [doi: 10.1177/1524839906289983] [Medline: 16928987]
43. Bloom B, Cohen RA, Freeman G. Summary health statistics for US children: National Health Interview Survey, 2009. Vital Health Stat 10 2010 Dec(247):1-82 [FREE Full text] [Medline: 21563639]
44. National Heart, Lung, and Blood Institute, National Asthma Education and Prevention Program. US Department of Health and Human Services, National Institutes of Health. National Heart, Lung, and Blood Institute. 2007. Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma Full Report 2007 URL: https://www.nhlbi.nih.gov/files/docs/guidelines/asthgdln.pdf [accessed 2017-07-28] [WebCite Cache ID 6sII7gaS4]
45. Weiland SK, Björkstén B, Brunekreef B, Cookson WO, von Mutius E, Strachan DP, International Study of AsthmaAllergies in Childhood Phase II Study Group. Phase II of the International Study of Asthma and Allergies in Childhood (ISAAC II): rationale and methods. Eur Respir J 2004 Sep;24(3):406-412 [FREE Full text] [doi: 10.1183/09031936.04.00090303] [Medline: 15358699]
46. Croft DR, Peterson MW. An evaluation of the quality and contents of asthma education on the World Wide Web. Chest 2002 Apr;121(4):1301-1307. [doi: 10.1378/chest.121.4.1301] [Medline: 11948066]

47. National Asthma Education and Prevention Program. National Heart, Lung, and Blood Institute. Hyattsville, MD: U.S. Department of Health and Human Services; 1997. Expert Panel Report 2: Guidelines for the Diagnosis and Management of Asthma URL: https://www.nhlbi.nih.gov/files/docs/guidelines/asthgdln_archive.pdf [accessed 2017-07-28] [WebCite Cache ID 6sIIRlh00]
48. Lucas CP, Zhang H, Fisher PW, Shaffer D, Regier DA, Narrow WE, et al. The DISC Predictive Scales (DPS): efficiently screening for diagnoses. J Am Acad Child Adolesc Psychiatry 2001 Apr;40(4):443-449. [doi: 10.1097/00004583-200104000-00013] [Medline: 11314570]
49. Dahlem NW, Zimet GD, Walker RR. The multidimensional scale of perceived social support: a confirmation study. J Clin Psychol 1991 Nov;47(6):756-761. [Medline: 1757578]
50. Derogatis L, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. In: Pichot P, Olivier-Martin R, editors. Psychological measurements in psychopharmacology. Oxford, England: Karger; 1974:79-110.
51. 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]
52. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. J Am Med Assoc 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
53. National Patient Safety Foundation. 2017. For Patients and Families URL: http://www.npsf.org/?page=patientsandfamilies2 [accessed 2017-07-28] [WebCite Cache ID 6sIJNsECc]
54. Ross SE, Lin C. The effects of promoting patient access to medical records: a review. J Am Med Inform Assoc 2003 Apr;10(2):129-138 [FREE Full text] [Medline: 12595402]
55. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
56. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: a systematic review of the evidence. J Diabetes Complications 2014;28(2):243-251 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.07.003] [Medline: 24332469]
57. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
58. Kaelber D, Pan EC. The value of personal health record (PHR) systems. AMIA Annu Symp Proc 2008:343-347 [FREE Full text] [Medline: 18999276]
59. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. JMIR Mhealth Uhealth 2016 Sep 09;4(3):e107 [FREE Full text] [doi: 10.2196/mhealth.5720] [Medline: 27613084]
60. Linn A, Vervloet M, van Dijk L, Smit EG, Van Weert JC. Effects of eHealth interventions on medication adherence: a systematic review of the literature. J Med Internet Res 2011 Dec 05;13(4):e103 [FREE Full text] [doi: 10.2196/jmir.1738] [Medline: 22138112]

Abbreviations

ACQ: Asthma Control Questionnaire
ADE: adverse drug event
BG: blood glucose
eHealth: electronic health
EMR: electronic medical record
HbA_{1c}: hemoglobin A1c
HRQoL: health-related quality of life
OR: odds ratio
PHR: personal health record
RCT: randomized controlled trial
RR: risk ratio
SMS: short message service

Edited by G Eysenbach; submitted 02.11.17; peer-reviewed by N Yuskel, A Motulsky, Y Wang, P Wark; comments to author 15.03.18; revised version received 18.06.18; accepted 14.09.18; published 18.12.18.

Please cite as:
Lancaster K, Abuzour A, Khaira M, Mathers A, Chan A, Bui V, Lok A, Thabane L, Dolovich L
The Use and Effects of Electronic Health Tools for Patient Self-Monitoring and Reporting of Outcomes Following Medication Use: Systematic Review
J Med Internet Res 2018;20(12):e294
URL: https://www.jmir.org/2018/12/e294/
doi:10.2196/jmir.9284
PMID:30563822

Original Paper

The Use of Social Networking Sites in Mental Health Interventions for Young People: Systematic Review

Brad Ridout[1], MAPS, PhD; Andrew Campbell[1], MAPS, PhD

Cyberpsychology Research Group, Faculty of Health Sciences, The University of Sydney, Sydney, Australia

Corresponding Author:
Brad Ridout, MAPS, PhD
Cyberpsychology Research Group
Faculty of Health Sciences
The University of Sydney
City Road
Sydney, 2006
Australia
Phone: 61 2 9351 7089
Email: brad.ridout@sydney.edu.au

Abstract

Background: The onset of mental health problems peaks between adolescence and young adulthood; however, young people face barriers to treatment and are often reluctant to seek professional help. Many are instead seeking support and information regarding their mental health via the Web, especially via social networking sites (SNSs), and hence, there is a promising opportunity to use SNSs to deliver or integrate with youth-focused online mental health interventions. Previous reviews have evaluated the effectiveness of SNSs for specific disorders in young people; however, none of the reviews have covered the breadth of SNS–based youth mental health interventions available across all mental health issues.

Objective: This review aimed to systematically identify available evidence regarding the use of SNS–based interventions to support the mental health of young people aged up to 25 years, to evaluate their effectiveness, suitability, and safety, and identify gaps and opportunities for future research.

Methods: The PubMed and PsycINFO databases were searched using Medical Subject Headings terms and exploded keywords and phrases. Retrieved abstracts (n=974) were double screened, yielding 235 articles for screening at the full-text level. Of these, 9 articles met the review inclusion criteria. Given the small number of studies, and the variety of outcome measures used, a quantitative meta-analysis was not possible.

Results: The 9 articles (quantitative studies, qualitative studies, and descriptions of the iterative design process) covered 5 separate interventions. Of the 5 interventions, 2 interventions used purpose-built platforms based on the moderated online social therapy (MOST) model, 2 used Facebook, and 1 evaluated a purpose-built mobile app. The 2 MOST interventions targeted specific mental health issues (depression and psychosis), whereas the others focused on improving mental health literacy, social support, and general well-being. Only 3 quantitative studies were identified, and all used a pre-post design (without a control group) to establish *proof of concept*. Of the outcome variables assessed, there were significant improvements in mental health knowledge and number of depressive symptoms but no improvement in anxiety or psychosis symptoms. Acceptability of and engagement with the SNS platforms were generally high, as were perceptions of usefulness and safety. Moderation by clinical experts was identified as a key component of the more successful interventions. When offered a choice, users showed a preference for mobile apps over Web-based interfaces.

Conclusions: The evidence reviewed suggests young people find SNS–based interventions highly usable, engaging, and supportive. However, future studies need to address the current lack of high-quality evidence for their efficacy in reducing mental health symptoms. Given young people are already turning to SNSs to engage in knowledge seeking and peer-to-peer support, SNS–based youth mental health interventions provide an opportunity to address some of the barriers young people face in accessing qualified mental health support and information.

(J Med Internet Res 2018;20(12):e12244) doi:10.2196/12244

KEYWORDS

social media; social networking; mental health; social support; support groups

Introduction

Supporting the mental health of young people is a major public health challenge, with mental disorders accounting for almost half of the nonfatal burden of disease among people aged 10 to 25 years [1]. Adolescence is a particularly vulnerable period of development, with the onset of mental health problems peaking between adolescence and young adulthood [2]. However, many problems are not detected until later in life, as young people are often reluctant to seek professional help [3] and face barriers to treatment such as cost, poor mental health literacy, confidentiality concerns, stigma, and inaccessibility to or lack of knowledge of resources [4,5].

Given that internet-enabled mobile devices have become a near-ubiquitous element of adolescence, with 45% of teens admitting that they are online *almost constantly* [6], it is not surprising that young people are increasingly seeking support and information regarding their mental health online [7].

Over the past decade, social media has become an important element of communication for young people, with virtually all having at least one active social media account [6]. People with mental illness are often among the highest users [8], with many reporting that social media fosters community among users and makes them feel supported and accepted [9]. Furthermore, a recent study found that actively engaging with peers online about their mental health concerns was associated with an increased likelihood of seeking formal mental health care [10].

Social networking sites (SNSs), a subset of social media, have become the predominant context for communication and social support–seeking behaviors online among adolescents [11]. SNS users create a profile within a bounded system, which they use to make and display connections with other users [12]. Posting of user-generated and Web-based content and functions such as liking, commenting, and tagging are the lifeblood of SNSs and differentiate SNSs from Web 1.0 communication tools such as message boards and online support groups [13].

Given the barriers to mental health support young people face and the fact that they are naturally turning to SNSs to engage in knowledge seeking and peer-to-peer support, there is a promising opportunity to use SNSs to deliver or integrate with youth-focused online mental health interventions. Compared with other online mental health resources such as online counseling, mobile apps, and online support groups, research into the use of SNSs to support and treat young people with mental health issues is only in its infancy and is highly fragmented. Although there have been reviews evaluating the effectiveness of SNSs for specific mental health disorders in young people [14,15] and online peer-to-peer support for young people more broadly [16], none of the reviews have covered the breadth of SNS–based youth mental health interventions available across all mental health issues. A systematic review of the literature regarding the use of SNS–based interventions to support the mental health of young people is, therefore, required to evaluate their effectiveness, suitability, and safety and identify gaps and opportunities for future research.

Methods

Search Strategy

This systematic review was performed using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines [17]. A PRISMA checklist is available in Multimedia Appendix 1. PubMed was searched using Medical Subject Headings terms, and PsycINFO was searched using exploded keywords and phrases (see Multimedia Appendix 2). Searches were conducted in June 2018 and restricted to English-language articles published in peer-reviewed journals between January 2000 and June 2018.

In total, the database searches yielded 1020 records (592 from PubMed and 428 from PsycINFO), of which 60 duplicates were removed. Additionally, 14 records were identified through manual searches of previous reviews, key journals, and reference lists of key articles.

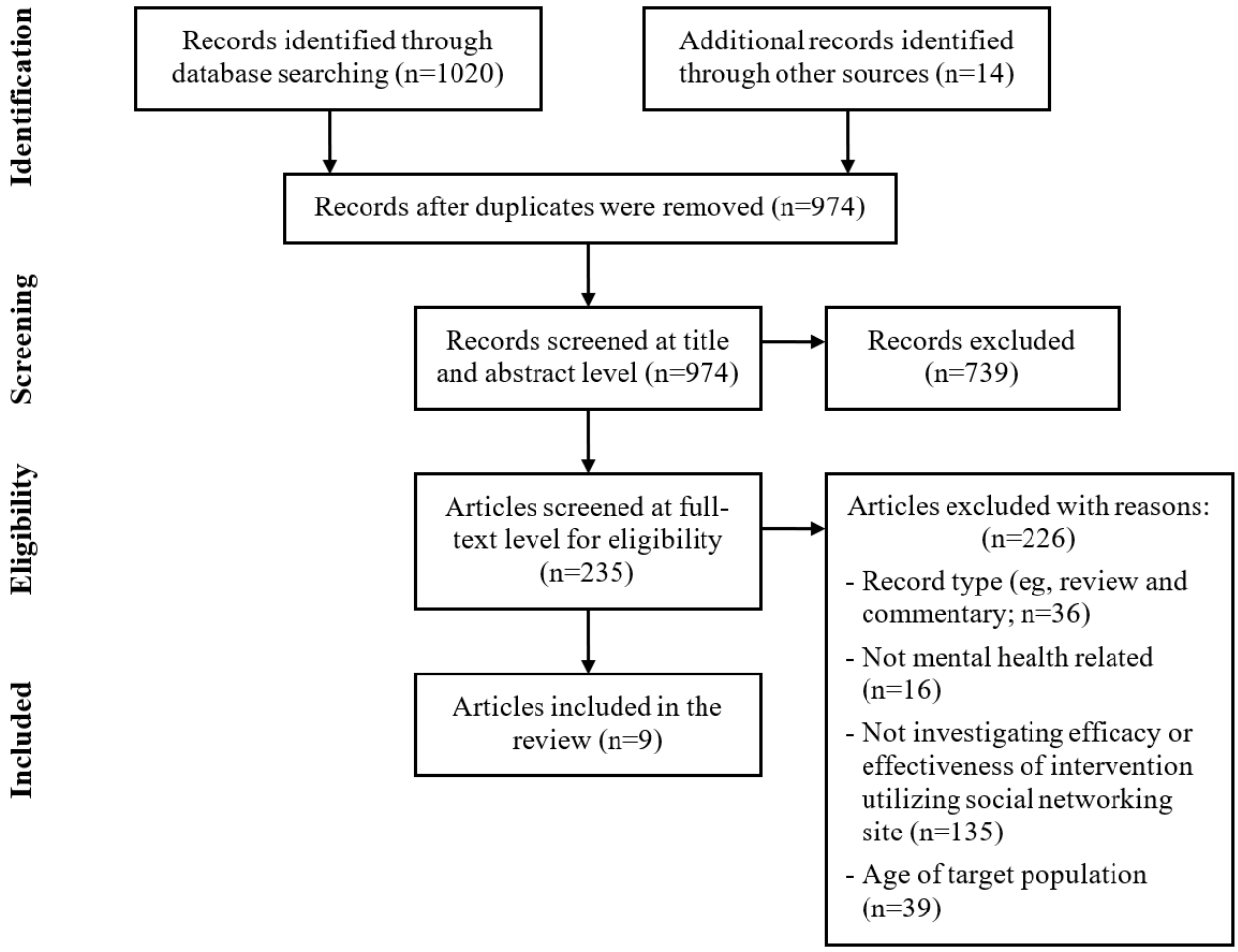
Screening Process

Figure 1 presents a PRISMA flow diagram of the screening process, which involved 2 stages: (1) title and abstract exclusion and (2) full-text exclusion. All records were independently screened by 2 reviewers (BR and AC) to establish relevance for inclusion at both stages. Any discrepancies between the reviewers were resolved by discussion. Of the 974 records identified (after duplicates were removed), 739 were removed because their titles and abstracts indicated they were not relevant to the topic of using SNS to support youth mental health. This left 235 articles to be assessed for eligibility according to predefined inclusion criteria.

The inclusion criteria were as follows: (1) the record must be an original empirical study (ie, not a review or commentary), (2) the primary aim of the study must be to address either a specific mental health condition or improve mental health and well-being generally, (3) the study must investigate the efficacy or effectiveness of a specific intervention utilizing an SNS (as defined by Boyd and Ellison [12]) to improve youth mental health (ie, not the impact of naturally occurring SNS support groups), and (4) the target population of the intervention must be young people aged up to 25 years.

A total of 36 articles were excluded during the second screening stage based on record type (eg, review and commentary) and 16 were excluded because they were not mental health related. A total of 135 articles were excluded as they were not investigating the efficacy or effectiveness of an intervention utilizing an SNS to improve youth mental health, and 39 articles were excluded based on the age of the target population being other than young people aged up to 25 years.

Figure 1. PRISMA flow diagram.



Data Analysis

Given the small number of studies included in this review, their exploratory nature, and the variety of outcome measures used, a quantitative meta-analysis was not possible. Primary and secondary outcome measures related to mental health are, therefore, reported (with effect sizes where possible) along with the characteristics of and usability and engagement data regarding the social networking components of the interventions.

Results

Study Characteristics

Detailed characteristics of the included articles (n=9) are provided in Multimedia Appendix 3. Of the 9 articles, 3 articles reported uncontrolled pilot studies that utilized a pre-post design [18-20] and 4 reported qualitative evaluations (2 of these evaluating 1 of the aforementioned pilot studies each) [21-24]. The 2 remaining articles were descriptions of the iterative design process of 2 of the already included studies [25,26]. In summary, there were 5 separate studies covered by the 9 included articles.

The articles were categorized according to the mental health issue they were focused on: psychotic disorder or mood disorder with psychotic features (n=3) [18,21,25], depression (n=2) [19,22], and health literacy and well-being (n=4) [20,23,24,27].

Origin

A total of 7 articles were from Australia, and the remaining articles were from the United States and Hong Kong. There were many studies from the United States and Europe included at the first screening stage; however, the majority of those that were focused on youth mental health were excluded at the second screening stage because they were investigating the impact of naturally occurring SNS groups, rather than purpose-built interventions for supporting youth mental health (an area that Australia is currently pioneering).

Interventions

Overall, 3 separate purpose-built SNSs were the focus of 7 of the articles (Horyzons: 3, Rebound: 2, and MindMax: 2), whereas the remaining articles evaluated interventions that used Facebook (one using a purpose-built Facebook game and the other using a closed Facebook group). Both Horyzons and Rebound were based on the moderated online social therapy (MOST) model developed by members of their research team [26], whereas MindMax took a modular approach combining well-being science, video games, and personal experience and stories of professional Australian Football League (AFL) players. Of the purpose-built apps, only MindMax was made available to the public.

Participants

Most of the samples consisted exclusively of participants with self-reported mental health concerns relevant to the focus of the respective study (n=5). Of the remaining articles, 2 articles used nonclinical samples of university students, and the 2 articles investigating the design of MindMax used convenience samples from the target audience of the app. As per the inclusion criteria, all studies aimed to support the mental health of children and/or young people, with the youngest participants across the studies being aged 15 years. The mean age of participants for all studies fell within a range of 18 to 21 years, except for the qualitative evaluation of MindMax [23], which was based on focus interviews of 7 participants (6 males and 1 female) with an age range of 24 to 49 years (average of 35 years). Although the convenience sample for this initial usability study was mostly made up of participants aged older than 25 years, it was included in this review as the target population of the MindMax app is young people aged 16 to 25 years. Gender was relatively balanced in all other studies (42%-50% male), except for the YBMen study, which focused on students who identified as black men. Participants were recruited from a range of sources and methods. The Horyzons and Rebound studies recruited from early intervention clinics, the Facebook studies from an online network of university students, and the MindMax studies were based on participatory design workshops and focus interviews with AFL fans, players, gamers, mental health and well-being consumers, clinicians, researchers, and academics.

Outcome Measures

The Horyzons study focused on mood disorders with psychotic features and used the Brief Psychiatric Rating Scale, the Calgary Depression Scale for Schizophrenia, and the Beck Anxiety Inventory as primary outcome measures to determine reduction in symptoms of psychosis, depression, and anxiety, respectively. The Rebound study focused on depression used the interviewer-rated Montgomery-Asberg Depression Rating Scale (MADRS) as its primary outcome measure, along with other secondary measures of anxiety, social and occupational functioning, strengths use, social support, and social connectedness. The Facebook game study focused on mental health literacy used a self-assessment questionnaire developed by the researchers to assess this primary outcome and modified questions from the Motivated Strategies for Learning Questionnaire to assess learning motivation as the secondary outcome. The authors of the closed Facebook group study, which also focused on improving mental health literacy (in addition to providing social support), mentioned that quantitative outcome measures were collected; however, only results of the qualitative interviews have been published to date. The MindMax study used participatory design workshops and focus interviews to evaluate usability and initial experiences of using the app, whereas outcomes related to the app's main aim to improve mental health literacy and well-being were being studied at the time of this review.

Study Quality

As all studies included in this systematic review were uncontrolled pilots or exploratory studies of acceptability or usability, no formal assessment of quality was performed. All 3 of the quantitative articles included used a pre-post design without a control group and aimed to provide *proof of concept*, rather than causal inferences about efficacy or effectiveness. The 2 MOST studies reported using completer analyses only (one of these had no dropouts and the other reported a dropout rate of 7.1%). The Facebook game study had a dropout rate of 42.5% and used multiple imputations to address loss of follow-up data for the 54 dropouts so that they could be included in an intention-to-treat (ITT) analysis. The result of the ITT analysis was consistent with the completer analysis conducted, which used only the 73 participants who completed both the pre- and posttest questionnaires.

The qualitative articles for the 2 MOST studies conducted semistructured interviews at the conclusion of the trial only, whereas the MindMax and closed Facebook group studies conducted structured interviews at multiple time points throughout the trial (although the latter only reported results of interviews conducted postintervention).

All 4 qualitative studies transcribed the interviews and coded them for thematic analysis using either QSR NVivo software or spreadsheet techniques. The Rebound and MindMax studies followed the established thematic analysis guidelines of Braun and Clarke [28], whereas the Horyzons study followed accepted qualitative methods [29,30] that recommend conducting multiple parses of the qualitative data with different levels of coding. The closed Facebook group study applied a data reduction technique developed by the lead author called the *rigorous and accelerated data reduction* technique [31].

Characteristics of Social Networking Functions

The social networking environment of the Horyzons and Rebound platforms was known as *The Café*. The platforms included a newsfeed where participants and moderators could post text, pictures, and videos and *like* or comment on posts of other users, similar to well-known Facebook functions. The newsfeed incorporated categories to organize discussion threads into themes (eg, *what's on your mind*, *I'm loving right now*, *cheer me up*, and *strength news*). The system also included a *homepage*, showing all the activity and notifications relevant to the participant. Participants could also view the *wall* of others, displaying that participant's individual activity (similar to what Facebook now refers to as the *timeline*), and their own *network* (similar to Facebook's *friends* function).

On the basis of the MOST model [26], the Horyzons and Rebound platforms were designed to reinforce the therapeutic content of the interventions and ensure constant flow between the therapy and social networking components. This was achieved by integrating questions within the therapeutic content to promote discussion and encourage users to share their own experiences, which then become discussion threads within The Café. Both platforms also featured an online group–based problem-solving space (known as *Talk It Out* in Rebound), guided by moderators within the social networking environment. Using an evidence-based problem-solving framework [32,33], moderators guided participants through the structured phases of problem definition, brainstorming solutions, identifying pros and cons, and summarizing possible choices. Offered solutions and participants' experiences were then saved in a database for

participants to refer to throughout the intervention. The Horyzons intervention also included a *job zone*, where users could access information regarding training and vocational recovery, or *ask Gina*, an expert in this area.

The MindMax platform also featured a familiar Facebook-style interface, with the main social networking function being a newsfeed where participants could post text, pictures, and animated graphics interchange formats (but not videos) and *like* or comment on posts of other users. Participants could also use hashtags to make their posts searchable to other participants (eg, #gratitude, #values, and #fitminds) or share their MindMax posts on other SNS such as Facebook. Integrated alongside the newsfeed were the *Train* well-being education modules (*Fit Minds*, *Values*, and *Thoughts*) and *Play*, where participants could use the *footies* awarded for completed activities within the education modules to play the *Flick Footy* video game and then post their score in the newsfeed for other participants to view, *like*, and comment on. There was also a *Me* function where participants could update their profile and view their activities and saved posts. The *Train* modules often included text and videos from AFL players and short questionnaires and activities. Completed activities were automatically posted to the newsfeed with appropriate hashtags by default (this could be turned off in profile settings). Being a mobile phone app, MindMax utilized push notifications to alert users about activities relevant to them when not using the app.

The purpose-built game *Ching Ching Story* was an app within Facebook, and as such, its social networking functions interacted with those of Facebook. For example, *task completed* notifications were posted on players' Facebook walls (now known as *timelines*) to acknowledge achievements in the game and encourage interaction between players via Facebook *likes* and comments. Additional functions within the game itself included the ability to send friends greetings, gifts, and special *tools* needed to accomplish certain tasks. Gifts were also offered to players for sending invitations to their Facebook friends to join the game. A leaderboard was also available within the game to create an atmosphere of competition.

The other Facebook study, YBMen, used a closed Facebook group to post educational material (taken from gender and culturally relevant popular culture references) and daily prompts for group discussion about the importance of mental health, social support, and the challenges associated with rigid adherence to masculine norms. Participants communicated with each other and the study team using standard Facebook functions (comments, *likes*, and posting and sharing content). Group facilitation techniques included group problem solving, action planning and feedback, and individual decision making to improve mental health behaviors and outcomes.

Moderation

The inclusion of expert moderators with clinical experience was a key feature of the MOST platforms. The moderators in Horyzons were clinical psychologists and vocational workers, identified within the social networking platform as *coach*. They moderated the site daily for 1 to 2 hours, and their role was to "guide, but not censor, the interaction to ensure a safe and supportive environment" [18]. The moderators in Rebound were experienced youth mental health clinicians who monitored the site daily. In addition, both platforms used an auto-detect risk management system to identify keywords associated with risk of relapse, self-harm, or suicide, which would then trigger crisis protocol and risk assessment.

The Rebound study also featured peer moderators known as *Super Users* —young people with recent lived experience of mental illness who were given training and supervision to provide peer support to other users of the site. However, most users were not aware of the Super Users (identified only by a distinct symbol on their avatar) and, therefore, did not recall interacting with them. Those who were aware of Super Users thought that they were useful as role models and gave them hope that they could also recover. The previously conducted Horyzon study did not include peer moderators; however, most users believed that including previous users of the site as peer moderators would be beneficial, and 90% reported interest in becoming one themselves.

The moderator in the YBMen Project was the lead author, an African American female researcher with 13 years' experience in research and community interventions on the mental health of black men. The moderator and her team (male and female graduate students) were responsible for not only monitoring the site but also for posting the daily educational material along with questions to generate group discussion and facilitate engagement, in contrast to the MOST platforms where therapeutic content was included in modules for users to work through at their own pace. It was not clear whether the YBMen moderator and her team were individually identifiable with the Facebook group or whether they all used the same Facebook user account.

The MindMax and Ching Ching Story studies did not report whether moderation was a feature of their platform.

Intervention Efficacy

Depression, Anxiety, and Psychosis Symptoms

A moderate to large reduction in participants' depressive symptoms (d=0.6) was found after using Horyzons for 1 month [18]. A small reduction in anxiety symptoms was found but failed to reach significance, and there was no reduction in psychosis symptoms between pre- and postintervention.

A similar effect size for reduction in depressive symptoms was reported by Rice et al [19] in their pilot study of Rebound, with a significant improvement in interviewer-rated depression scores on the MADRS after 2 months (d=0.45). There was no improvement in anxiety, social and occupational functioning, social support, or social connectedness; however, there was a trend (P<.1; d=0.29) for improved strength use.

Mental Health Literacy

The study measuring health literacy demonstrated a moderate to large improvement in performance on their 31-question knowledge test (d=0.65) following a 3-week period of using the purpose-built Facebook game [20]. Intrinsic goal orientation was identified as the primary factor in learning motivation. Self-efficacy for learning and performance significantly predicted learning outcomes, whereas test anxiety was found

to be negatively associated with learning outcomes. The MindMax and closed Facebook group studies also aimed to improve mental health literacy; however, studies published to date have not reported this as an outcome variable.

Engagement

User engagement data for the Horyzons platform suggested high use among participants, with 60% using the system in each of the 4 weeks (70% in at least 3 out of 4 weeks). The social networking component was used by 95% of participants, with a median of 192 social page views/actions per participant. Module use varied across participants; however, 95% completed at least one full therapy module and 60% completed at least three modules (of 7 available). There was a median of 65 therapy-based page views per participant.

Engagement with the Rebound platform was also high, with 70% logging in weekly (78.5% in at least 2 of 3 months). The social networking component was used by all participants, with a mean of 51.1 social posts per participant. In terms of therapeutic content, 42.9% completed 5 or more therapy modules (of 56 available) and 26.2% reported completing 5 or more *actions* (applying therapy content in the offline world).

The structured interviews conducted to investigate initial experiences with the MindMax platform revealed that participants found the gaming and sporting elements to be the most engaging aspects of the app, with many returning to use the game even after completing all the well-being training. Statistics of engagement with the social networking functions are anticipated to be reported in a future report of the currently ongoing evaluation trial of MindMax.

Engagement with the YBMen intervention was assessed via quantitative data collected about the level of Facebook activity recorded. Most participants (67.3%) viewed Facebook postings at least weekly and around half (50.9%) actively contributed each week by commenting or posting new material.

Assessment of engagement with the purpose-built Facebook game was not reported.

Usability

Horyzons was considered a useful long-term treatment option beyond discharge by 70% of participants, with a majority reporting that it significantly increased their social connectedness (60%) and empowered them in their own recovery process (55%). The social networking component of Horyzons was perceived as useful by 70% of participants. Moreover, 90% of participants considered moderation to be supportive and 85% thought it would be beneficial to include peer moderators who were previous users of Horyzons (with 90% reporting they would like to become online peer moderators). There were no incidents (ie, adverse events or inappropriate usage) during the study.

User experience data for the Rebound platform were collected via a standardized industry tool for benchmarking websites called Web Analysis and Measurement Inventory (WAMMI) [34]. Rebound rated above average on all 5 WAMMI domains (attractiveness, controllability, efficiency, helpfulness, and learnability), achieving a global utility percentile rank of 59.6. In addition, ratings on a 1- to 5-point scale were collected to assess safety (4.7), helpfulness (3.6), and perceived benefits related to social connectedness (3.5). Impressions of the moderators were assessed on 1- to 7-point scales, with participants rating their agreement that moderators encouraged open discussion (6.0), accepted them (5.8), provided them with choices (5.5), and listened to how they would like to use Rebound (5.6).

Initial usability testing of MindMax with 3 users revealed an appreciation for the gamification of content and shared use by known AFL players [27]. Concerns were expressed regarding privacy and the possibility that users may only post or *like* content to get points, rather than meaningfully engage with the app.

Interviews conducted at the conclusion of the YBMen study revealed that the use of Facebook as the intervention platform was well-liked by participants, as they appreciated being able to receive notifications alerts on their mobile phones. Many participants also liked that Facebook facilitated conversations that they would not feel as comfortable having face-to-face. Barriers to engagement identified included not being able to understand some of the language used by the moderator.

Assessment of the usability of the purpose-built Facebook game was not reported.

Discussion

Principal Findings

The aim of this systematic review was to identify studies investigating the use of SNS to support the mental health of children and youth. A total of 9 articles reporting on 5 separate studies were identified. Of the 9 studies, 2 studies targeted specific mental health issues (depression and psychosis), whereas the other studies focused on improving mental health literacy, social support, and general well-being. Only 3 quantitative studies were identified and all used a pre-post design (without a control group) to establish *proof of concept*, rather than causal inferences about efficacy. Although this precluded any meta-analysis or assessment using Effective Practice and Organization of Care quality criteria, some of the outcome measures produced encouraging results, with significant reductions in depressive symptoms and significant improvements in mental health knowledge. However, there was no significant reduction in anxiety or psychosis symptoms. Acceptability and usability of the platforms reviewed were generally high, as were perceptions of usefulness and safety. There were no adverse incidents reported in any of the studies. When offered a choice, users showed a preference for mobile apps over Web-based interfaces and appreciated receiving notification alerts on their mobile phones. Overall, this review found evidence for the potential for SNS–based interventions to support the mental health of young people.

Engagement with the SNS platforms was high in most studies, with low dropout rates, and most users logging in and actively posting and engaging with content, moderators, and other users, on at least a weekly basis. Moderation was identified as a key component of the success of the interventions. The therapeutic

interventions that were most favorably viewed by users were those that were guided by moderators within the social networking environment, with users generally finding moderators to be friendly, supportive, and caring. There was also initial support for the inclusion of peer moderators to act as role models and support the experience of other users; however, it is not suggested that these should replace the role of expert moderators with clinical experience.

Positive feedback on the benefit of giving and receiving peer-to-peer support was also received, consistent with established literature [35,36]. Users of the MOST platforms reported that the most valued characteristic of the intervention was the ability to connect with other young people of a similar age with shared experiences, backgrounds, and mental health issues. Users felt safe because the sites could only be accessed by clients of the mental health services from which they had been recruited, which also contributed to feelings of belonging to a group of peers with similar experiences. There were indications that users felt understood, supported, more socially connected, and more willing to discuss their issues as a result of interacting with peers who were facing challenges similar to them.

However, not all users were active in their use of the social networking functions, with qualitative feedback revealing that some users preferred to *eavesdrop* on discussions taking place or *lurk*. Santesteban-Echarri et al [22] identified 2 clear subgroups of low interactors in the Rebound study. The first subgroup did not like online interaction and/or had sufficient offline support, supporting the typology of social media users by Fergie et al [37] that suggests the more offline support someone has, the less regularly they engage with health-related content on social media. The second subgroup of low interactors simply felt too shy, indicating that either the activity on the site was not high enough for them to feel comfortable to initiate a conversation or that not *knowing* fellow users was a barrier (despite anonymity being one of the obvious advantages and aims of closed SNSs for supporting youth mental health). This desire to know other users was also raised by participants in the YBMen project, who suggested that having occasional face-to-face meetings would have benefited the intervention (although the authors note that this may have been influenced by the project's association with an existing offline group).

Although it is possible that having a less than positive regard for anonymous online interactions may be a potential barrier to gaining benefit from SNS–based youth mental health interventions, more research is needed to establish whether eavesdropping on discussions may still be beneficial for low interactors. Recent research suggests that having a strong sense of community and inclusive culture are important factors for deriving positive outcomes among lurkers of online health support groups [38,39]. The design of SNSs for supporting youth mental health should, therefore, engage in strategies to create a sense of community and promote regular contributions from users [40], given it appears that variable levels of interactivity and engagement over time are features of these platforms.

Overall, the integration of the social networking components with the psychoeducation and therapy modules in the MOST interventions was considered successful, as evidenced by a high level of engagement with both and positive qualitative feedback from users. The MindMax and Ching Ching Story platforms also aimed to integrate social networking functions with the online education activities around mental health literacy, for example, by encouraging users to post about their successes in completing activities and comment on the successes of others. However, there was no evidence provided to suggest that the social networking functions were well utilized during the trials of these 2 platforms (although it was stated that *social connectedness* of MindMax users will be reported in a future evaluation of a naturalistic trial).

The need to integrate therapeutic and social networking functions was not an issue for the YBMen project, as all activities took place within the closed Facebook group. This had the additional benefit of locating the intervention within a platform that most users were already familiar with and using daily on multiple devices, including mobile devices, which was something that users appreciated. Although using naturally occurring SNS such as Facebook to deliver interventions could be a way to address the difficulty that purpose-built platforms may face in creating the norms, dynamics, and atmosphere of naturally occurring online communities [41], more evaluation is needed regarding the potential benefits and risks of using such widely used SNSs for this purpose [42].

In their commentary on the future of peer-to-peer support on social media, Naslund et al [43] identified several risks that should be considered in the design of any platform that enables peer-to-peer support. First, there are risks inherent with obtaining advice from nonexpert peers who may unwittingly pass on misleading or unreliable information. Although research shows that many users of online health forums are aware of the need to evaluate the accuracy of advice received and whether it applies to their own circumstances [44], it is not known whether young people with mental health concerns do so routinely. Second, similar to all online environments, there is the potential to be exposed to hostile or derogatory comments from others, which could have a negative impact on the mental health of users. These key risks can be largely mitigated against on closed SNSs by having clinically trained moderators regularly review posts made by users so that they can clarify, correct, or potentially remove any posts that may be problematic for other users. Although none of the studies in this review reported the need to address any problematic posts, the MOST and YBMen interventions did have this ability, as their expert moderators were actively engaged with all content posted. The presence of expert moderators greatly contributed to users' perception of safety of the platforms.

Limitations

It is important to acknowledge the limitations of this systematic review. First, searches were conducted in 2 databases only, limited to English-language publications, and excluded grey literature. Although the selected databases contain the largest number of health, medical, and psychological journals, this search strategy was complemented by hand searches of previous

reviews, key journals, and reference lists of key articles, which yielded an additional 14 articles. Searches were current at June 2018, but as mentioned above, some of the interventions evaluated in the included articles were either ongoing or had collected additional data that were intended to be published in the future (eg, a 5-year randomized controlled trial of Horyzons was recently completed); therefore, it is likely that further articles will soon appear in scholarly journals. Finally, it is possible that this review was subject to publication bias should authors have failed to publish studies with null or negative findings.

Conclusions and Implications for Future Research

This review updates and expands previous reviews of the use of SNS for supporting youth mental health, which have, to date, only focused on specific disorders. By broadening the scope to include all aspects of mental health, including mental health literacy, this review shows that SNSs may play a useful role in providing mental health support to both clinical and nonclinical populations. It has also highlighted the importance of involving end users across all stages of intervention and platform design development according to participatory design principles [45] and suggests that users prefer to be able to access SNS interventions on their mobile devices.

The evidence reviewed suggests that young people find SNS–based interventions highly usable, engaging, and supportive. However, high-quality evidence for their efficacy in reducing mental health symptoms is currently lacking. Furthermore, the majority of data collected in the reviewed studies came from participants aged over 18 years; therefore, there is a particular need for further investigation into the suitability of SNS–based interventions for adolescents aged less than 18 years. Now that proof-of-concept is established for some of the SNS interventions reviewed here, higher quality studies are required (ie, randomized controlled trials over longer periods), with populations that focus on adolescents as well as young adults, to build the evidence base in this field and address the following unanswered questions: Which aspects of SNS interventions are most beneficial for users and how do they mediate mental health outcomes?, Do skills gained online translate to sustained improvements in offline functioning and well-being?, Are some mental health issues and/or phases of the users' journey better suited to SNS interventions than others?, What level of participation is required from users to gain benefit?, Are mobile apps and mobile-friendly interfaces more beneficial for users?, and Is there an optimum user group/community size? There are also methodological challenges to address such as those associated with evaluating multicomponent interventions, collecting objective measures of mental health outcomes online, and dealing with variable levels of engagement and retention over longer periods.

Overall, the evidence reviewed suggests that both clinical and nonclinical users found SNS–based interventions to be safe, engaging, supportive, and useful. When moderated, ideally by mental health professionals, the benefits of SNS–based interventions for youth mental health appear to outweigh any potential risks. Given that young people are already turning to SNSs to engage in knowledge seeking and peer-to-peer support, SNS–based youth mental health interventions present a promising opportunity to help address some of the barriers young people face in accessing qualified mental health support and information. They also provide an opportunity to combine the well-established benefits of peer-to-peer support with accessible and cost-effective online interventions.

Acknowledgments

This project was resourced by yourtown's Kids Helpline via a grant from the Future Generations Investment Company. The authors would like to acknowledge Julia Schaefer for assisting in collecting and reviewing abstracts in the initial stage of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[PDF File (Adobe PDF File), 135KB - jmir_v20i12e12244_app1.pdf]

Multimedia Appendix 2

Database search strategy.

[XLSX File (Microsoft Excel File), 9KB - jmir_v20i12e12244_app2.xlsx]

Multimedia Appendix 3

Study characteristics.

[XLSX File (Microsoft Excel File), 12KB - jmir_v20i12e12244_app3.xlsx]

References

1. Australian Institute of Health and Welfare. 2016. Australian burden of disease study: impact and causes of illness and death in Australia 2011 URL: https://www.aihw.gov.au/getmedia/d4df9251-c4b6-452f-a877-8370b6124219/19663.pdf [accessed 2018-08-30] [WebCite Cache ID 723YBWMS4]
2. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry 2015 Mar;56(3):345-365. [doi: 10.1111/jcpp.12381] [Medline: 25649325]
3. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people's help-seeking for mental health problems. Aust E J Adv Mental Health 2005 Dec;4(3):218-251. [doi: 10.5172/jamh.4.3.218]
4. Wright A, Jorm AF. Labels used by young people to describe mental disorders: factors associated with their development. Aust N Z J Psychiatry 2009 Jan;43(10):946-955. [doi: 10.1080/00048670903179129]
5. Brown A, Rice SM, Rickwood DJ, Parker AG. Systematic review of barriers and facilitators to accessing and engaging with mental health care among at-risk young people. Asia Pac Psychiatry 2016 Mar;8(1):3-22. [doi: 10.1111/appy.12199] [Medline: 26238088]
6. Anderson M, Jiang J. Pew Research Center. Teens, social media & technology 2018 URL: http://publicservicesalliance.org/wp-content/uploads/2018/06/Teens-Social-Media-Technology-2018-PEW.pdf [accessed 2018-08-30] [WebCite Cache ID 723XLv9Gr]
7. Rickwood DJ, Mazzer KR, Telford NR. Social influences on seeking help from mental health services, in-person and online, during adolescence and young adulthood. BMC Psychiatry 2015;15:40 [FREE Full text] [doi: 10.1186/s12888-015-0429-6] [Medline: 25886609]
8. Birnbaum ML, Rizvi AF, Confino J, 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 2017 Dec;11(4):290-295 [FREE Full text] [doi: 10.1111/eip.12237] [Medline: 25808317]
9. Dyson MP, Hartling L, Shulhan J, Chisholm A, Milne A, Sundar P, et al. A systematic review of social media use to discuss and view deliberate self-harm acts. PLoS One 2016;11(5):e0155813 [FREE Full text] [doi: 10.1371/journal.pone.0155813] [Medline: 27191728]
10. Lawlor A, Kirakowski J. Online support groups for mental health: a space for challenging self-stigma or a means of social avoidance? Comput Human Behav 2014 Mar;32:152-161. [doi: 10.1016/j.chb.2013.11.015]
11. Australian Communications and Media Authority. Communications report 2014–15 URL: https://www.acma.gov.au/~/media/Research%20and%20Analysis/Report/pdf/ACMA%20Communications%20report%202014-15%20pdf.pdf [accessed 2018-08-30] [WebCite Cache ID 723VXw3ce]
12. Boyd D, Ellison N. Social network sites: definition, history, and scholarship. J Comput-Mediat Comm 2007;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
13. Cormode G, Krishnamurthy B. Key differences between Web 10 and Web 2.0. First Monday 2008;13(6):-. [doi: 10.5210/fm.v13i6.2125]
14. Rice SM, Goodall J, Hetrick SE, Parker AG, Gilbertson T, Amminger GP, et al. Online and social networking interventions for the treatment of depression in young people: a systematic review. J Med Internet Res 2014;16(9):e206 [FREE Full text] [doi: 10.2196/jmir.3304] [Medline: 25226790]
15. Välimäki M, Athanasopoulou C, Lahti M, Adams CE. Effectiveness of social media interventions for people with schizophrenia: a systematic review and meta-analysis. J Med Internet Res 2016;18(4):e92 [FREE Full text] [doi: 10.2196/jmir.5385] [Medline: 27105939]
16. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. JMIR Ment Health 2015;2(2):e19 [FREE Full text] [doi: 10.2196/mental.4418] [Medline: 26543923]
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64. [Medline: 19622511]
18. Alvarez-Jimenez M, Bendall S, Lederman R, Wadley G, Chinnery G, Vargas S, et al. On the HORYZON: moderated online social therapy for long-term recovery in first episode psychosis. Schizophr Res 2013 Jan;143(1):143-149. [doi: 10.1016/j.schres.2012.10.009] [Medline: 23146146]
19. Rice S, Gleeson J, Davey C, Hetrick S, Parker A, Lederman R, et al. Moderated online social therapy for depression relapse prevention in young people: pilot study of a 'next generation' online intervention. Early Interv Psychiatry 2016 Jun 17;12(4):613-625. [doi: 10.1111/eip.12354] [Medline: 27311581]
20. Li TM, Chau M, Wong PW, Lai ES, Yip PS. Evaluation of a Web-based social network electronic game in enhancing mental health literacy for young people. J Med Internet Res 2013;15(5):e80 [FREE Full text] [doi: 10.2196/jmir.2316] [Medline: 23676714]
21. Gleeson JF, Lederman R, Wadley G, Bendall S, McGorry PD, Alvarez-Jimenez M. Safety and privacy outcomes from a moderated online social therapy for young people with first-episode psychosis. Psychiatr Serv 2014 Apr 1;65(4):546-550. [doi: 10.1176/appi.ps.201300078] [Medline: 24687106]
22. Santesteban-Echarri O, Rice S, Wadley G, Lederman R, D'Alfonso S, Russon P, et al. A next-generation social media-based relapse prevention intervention for youth depression: qualitative data on user experience outcomes for social networking,

safety, and clinical benefit. Internet Interv 2017 Sep;9:65-73 [FREE Full text] [doi: 10.1016/j.invent.2017.06.002] [Medline: 30135839]
23. Peever N, Vella K, Johnson D, Ploderer B, Klarkowski M, Mitchell J. Understanding initial experiences with Mindmax, an mHealth app that draws on shared interests in sports and video games. 2017 Presented at: Proceedings of the 29th Australian Conference on Computer-Human Interaction; Nov 28-Dec 1; Brisbane, Australia.
24. Watkins DC, Allen JO, Goodwill JR, Noel B. Strengths and weaknesses of the Young Black Men, Masculinities, and Mental Health (YBMen) Facebook project. Am J Orthopsychiatry 2017;87(4):392-401. [doi: 10.1037/ort0000229] [Medline: 27977287]
25. Lederman R, Wadley G, Gleeson J, Bendall S, Alvarez-Jimenez M. Tochi. Moderated online social therapy: designing and evaluating technology for mental health URL: https://people.eng.unimelb.edu.au/gwadley/pubs/ToCHI-MOST.pdf [accessed 2018-11-20] [WebCite Cache ID 744FeEjg8]
26. Gleeson JF, Alvarez-Jimenez M, Lederman R. Moderated online social therapy for recovery from early psychosis. Psychiatr Serv 2012 Jul;63(7):719. [doi: 10.1176/appi.ps.20120p719] [Medline: 22752039]
27. Mitchell J, Vella K, Johnson D, Peever N, Wan SC, Davenport T, et al. MindMax: using videogames and sport to engage young men and improve wellbeing. 2017 Presented at: Proceedings of the 2nd Symposium of Computing and Mental Health; 2017 May 6; Denver, Co URL: http://mentalhealth.media.mit.edu/wp-content/uploads/sites/46/2017/05/CMH_2017_paper_16.pdf
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. Glaser B, Strauss A. The Discovery of Grounded Theory: Strategies for Qualitative Research. London: Wiedenfeld & Nicholson; 1999.
30. Neuman W. Social Research Methods: Qualitative and Quantitative Approaches. Boston, Ma: Allyn & Bacon, Inc; 2010.
31. Watkins D. Mixed Methods Research (pocket Guides To Social Work Research Methods). New York, Ny: Oxford University Press; 2015.
32. McFarlane W, Beels CC, Lefley HP. Multifamily Groups in the Treatment of Severe Psychiatric Disorders. New York, Ny: Guilford Publications, Inc; 2004.
33. McFarlane WR, Lukens E, Link B, Dushay R, Deakins SA, Newmark M, et al. Multiple-family groups and psychoeducation in the treatment of schizophrenia. Arch Gen Psychiatry 1995 Aug;52(8):679-687. [Medline: 7632121]
34. Kirakowski J, Claridge N, Whitehand R. Human centered measures of success in web site design. 1998 Presented at: Proceedings of the 4th Conference on Human Factors & the Web; 1998 Jun 5; Basking Ridge, Nj URL: https://zing.ncsl.nist.gov/hfweb/att4/proceedings/kirakowski/index.html
35. Corrigan PW. Impact of consumer-operated services on empowerment and recovery of people with psychiatric disabilities. Psychiatr Serv 2006 Oct;57(10):1493-1496. [doi: 10.1176/ps.2006.57.10.1493] [Medline: 17035571]
36. Davidson L, Chinman M, Sells D, Rowe M. Peer support among adults with serious mental illness: a report from the field. Schizophr Bull 2006 Jul;32(3):443-450 [FREE Full text] [doi: 10.1093/schbul/sbj043] [Medline: 16461576]
37. Fergie G, Hunt K, Hilton S. Social media as a space for support: young adults' perspectives on producing and consuming user-generated content about diabetes and mental health. Soc Sci Med 2016 Dec;170:46-54 [FREE Full text] [doi: 10.1016/j.socscimed.2016.10.006] [Medline: 27750067]
38. Fullwood C, Chadwick D, Keep M, Attrill-Smith A, Asbury T, Kirwan G. Lurking towards empowerment: explaining propensity to engage with online health support groups and its association with positive outcomes. Comput Human Behav 2019 Jan;90:131-140. [doi: 10.1016/j.chb.2018.08.037]
39. Sun N, Rau PP, Ma L. Understanding lurkers in online communities: a literature review. Comput Human Behav 2014 Sep;38:110-117. [doi: 10.1016/j.chb.2014.05.022]
40. Attrill A, Fullwood C. Online support groupsnhancing the user experience with cyberpsychological theory. In: Applied Cyberpsychology. London: Palgrave Macmillan; 2018.
41. Lawlor A, Kirakowski J. Online support groups for mental health: a space for challenging self-stigma or a means of social avoidance? Comput Human Behav 2014 Mar;32:152-161. [doi: 10.1016/j.chb.2013.11.015]
42. Alvarez-Jimenez M, Gleeson JF, Rice S, Gonzalez-Blanch C, Bendall S. Online peer-to-peer support in youth mental health: seizing the opportunity. Epidemiol Psychiatr Sci 2016 Apr;25(2):123-126. [doi: 10.1017/S2045796015001092] [Medline: 26880239]
43. 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] [Medline: 26744309]
44. Armstrong N, Powell J. Patient perspectives on health advice posted on Internet discussion boards: a qualitative study. Health Expect 2009 Sep;12(3):313-320. [doi: 10.1111/j.1369-7625.2009.00543.x] [Medline: 19555377]
45. Schuler D, Namioka A. Participatory Design: Perspectives on Systems Design. Hillsdale, Nj: CRC Press; 1993.

Abbreviations

AFL: Australian Football League
ITT: intention-to-treat
MADRS: Montgomery-Asberg Depression Rating Scale
MOST: Moderated Online Social Therapy
PRISMA: preferred reporting items for systematic reviews and meta-analyses
SNSs: social networking sites
WAMMI: Web Analysis and Measurement Inventory

Edited by G Eysenbach; submitted 19.09.18; peer-reviewed by S D'Alfonso, M Dyson, J Constantin; comments to author 08.10.18; revised version received 21.10.18; accepted 25.10.18; published 18.12.18.

Please cite as:
Ridout B, Campbell A
The Use of Social Networking Sites in Mental Health Interventions for Young People: Systematic Review
J Med Internet Res 2018;20(12):e12244
URL: https://www.jmir.org/2018/12/e12244/
doi:10.2196/12244
PMID:30563811

Original Paper

Electronic Consultation Services Worldwide: Environmental Scan

Justin Joschko[1], MA; Erin Keely[2,3], MD; Rachel Grant[4], MSN; Isabella Moroz[1], PhD; Matthew Graveline[5], MA; Neil Drimer[5], MHSc; Clare Liddy[1,6], MD, MSc

[1]CT Lamont Primary Health Care Research Centre, Department of Family Medicine, University of Ottawa, Ottawa, ON, Canada

[2]Department of Medicine, University of Ottawa, Ottawa, ON, Canada

[3]Division of Endocrinology/Metabolism, The Ottawa Hospital, Ottawa, ON, Canada

[4]Faculty of Education, University of Ottawa, Ottawa, ON, Canada

[5]Canadian Foundation for Healthcare Improvement, Ottawa, ON, Canada

[6]Bruyère Research Institute, Ottawa, ON, Canada

Corresponding Author:
Clare Liddy, MD, MSc
CT Lamont Primary Health Care Research Centre
Department of Family Medicine
University of Ottawa
Ottawa, ON,
Canada
Phone: 1 613 562 6262 ext 2928
Fax: 1 613 562 6734
Email: CLiddy@bruyere.org

Abstract

Background: Excessive wait times for specialist care pose a serious concern for many patients, leading to duplication of tests, patient anxiety, and poorer health outcomes. In response to this issue, many health care systems have begun implementing technological innovations designed to improve the referral-consultation process. Among these services is electronic consultation (eConsult), which connects primary care providers and specialists through a secure platform to facilitate discussion of patients' care.

Objective: This study aims to examine different eConsult services available worldwide and compare the strategies, barriers, and successes of their implementation in different health care contexts.

Methods: We conducted an environmental scan comprising 3 stages as follows: literature review; gray literature search; and targeted, semistructured key informant interviews. We searched MEDLINE and EMBASE (literature review) and Google (gray literature search). Upon completing the search, we generated a list of potential interview candidates from among the stakeholders identified. Potential participants included researchers, physicians, and decision makers. The maximum variation sampling was used to ensure sufficient breadth of participant experience. In addition, we conducted semistructured interviews by telephone using an interview guide based on the RE-AIM framework. Analyses of transcripts were conducted using a thematic synthesis approach.

Results: A total of 53 services emerged from the published and gray literature. Respondents from 10 services participated in telephonic interviews. The following 4 major themes emerged from the analysis: service structure; benefits of eConsult; implementation challenges; and implementation enablers.

Conclusions: eConsult services have emerged in a variety of countries and health system contexts worldwide. Despite differences in structure, platform, and delivery of their services, respondents described similar barriers and enablers to the implementation and growth and reported improved access and high levels of satisfaction.

(J Med Internet Res 2018;20(12):e11112) doi:10.2196/11112

KEYWORDS

electronic consultation; interviews; primary care; referral-consultation process; telemedicine; quality of care; specialist care

Introduction

Excessive wait times for specialist care pose a serious concern for many patients, leading to duplication of tests, patient anxiety, and poorer health outcomes [1-3]. In response to this issue, many health care systems have begun implementing technological innovations designed to improve the referral-consultation process [4-8]; among these are electronic consultation (eConsult) services—secure Web-based applications that facilitate asynchronous communication between primary care providers (PCP) and specialists, allowing PCPs to ask questions to specialists directly about a patient's care and, in some cases, avoid the need for a face-to-face consultation.

In 2009, our team launched the Champlain Building Access to Specialists through eConsultation (BASE) eConsult service in the Champlain health region of Ontario. As our service grew, we wanted to gain a better understanding of whether other such services were operating in Canada. To this end, we conducted an environmental scan of services across Canada to ascertain the status of eConsult in each province. Our study found no other eConsult services in the country; only 2 other services emerged besides our own, both of which were exclusively electronic referral (eReferral) systems [9]. Unlike eConsult, which can supplement or replace the in-person referral in some cases, eReferral is simply a platform that lets PCPs submit or schedule patient referrals electronically.

Since then, interest in eConsult has expanded in many countries [7,8]. Champlain BASE has likewise grown, reaching its 50,000th case. Building on its regional success, the service is in the process of expanding province-wide, with money for its implementation earmarked in Ontario's 2017 budget. In addition, the service is expanding beyond provincial borders. Partnerships with provincial and national groups have resulted in services informed by the BASE model emerging in Alberta, Manitoba, Quebec, and Newfoundland and Labrador.

Given our service's forthcoming growth, we have endeavored to update our previous scan, making 2 key changes to its scope. First, we have expanded our search for services available outside of Canada to capture a broader range of experiences. Second, we focused our current scan exclusively on eConsult services, as eReferral services address different issues and are not directly comparable to eConsult. These changes allowed us to examine the success and barriers faced by eConsult services in a wide array of different contexts, providing invaluable insight into which elements are most vital and which may—or indeed, should—be adapted to fit the individual circumstances of the region in which they are implemented.

Methods

Design

This study follows the methodology used in our previous environmental scan modified to expand from a Canadian to an international focus [9]. Our process was implemented in 3 stages—a literature review, gray literature search, and key informant interviews.

Population

Our environmental scan targeted any documentation pertaining to the development, implementation, or expansion of eConsult services. We defined eConsult services as asynchronous, directed communication between providers over a secure electronic medium that involved sharing of patient-specific information and sought clarification or guidance regarding clinical care. Although services based in any country were eligible for inclusion, only literature published in English and French were reviewed.

Literature Review

We conducted a literature search of MEDLINE and EMBASE databases on April 5, 2017 to identify existing eConsult services. Our search strategy built on the keyword combinations and variants used in our previous scan, with modifications to expand the scope beyond Canadian services to include services implemented internationally and focus exclusively on eConsult services (Multimedia Appendix 1).

Gray Literature Search

Following the literature review, we performed a gray literature search on April 7, 2017 using the Google search engine (Multimedia Appendix 2). If the search yielded >100 hits, the reviewer read through all results until 10 pages (1000 hits) had passed without yielding any information about a new service or the end of the search was reached.

Key Informant Interviews

Upon completing the literature review and gray literature search, we generated a list of potential interview candidates from among the stakeholders identified in the acquired documents. Potential participants included researchers, health care providers (eg, physicians), and decision makers involved in the development or implementation of an eConsult service. To ensure sufficient breadth of participant experience, we used the maximum variation sampling [10], with relevant factors including the service's country of origin, technology platform, and host organization. Of note, we did not attempt to contact Canadian services for interviews, as our team had already developed partnerships with all services identified by the scan.

Potential participants were contacted by emails, which were written in English. For services based in countries with majority languages other than English, we generated brief descriptions of the project in their language using Google Translate. A member of our research team (JJ) conducted semistructured interviews by telephone between August 30, 2017 and November 14, 2017 using an interview guide structured around the RE-AIM framework, which assesses a project's ability to translate research into action using the 5 following categories: reach, effectiveness, adoption, implementation, and maintenance (Multimedia Appendix 3) [11]. The interviewer was a research coordinator with a master's degree and experience conducting previous qualitative studies; he had no prior relationship with any interview subjects. Interviews began with a brief discussion of the research project's objectives. All interviews were conducted in English and lasted 20-45 minutes. Interviews were audiorecorded and transcribed verbatim. Participants received

a copy of the interview transcript to review and correct if necessary [12].

Data Analysis

Transcripts were uploaded into NVivo version 11 (QSR International). Team members followed the thematic synthesis approach outlined by Thomas and Harden [13]. One member of the research team (JJ) reviewed the transcripts and developed an initial framework of descriptive and analytical themes. The remaining 6 team members independently reviewed the transcripts using the framework, meeting to discuss progress, identify any disconfirming data, and confirm whether data saturation had been reached. Emerging themes were agreed upon by consensus and amended as needed based on new data.

Ethics Approval

The Ottawa Health Science Network Research Ethics Board (20120894-01H) and the Bruyère Continuing Care Research Ethics Board (M16-12-052) provided ethics approval for this study.

Results

Service Details

A search of the MEDLINE database returned 262 cases, of which 115 were deemed sufficiently relevant to be reviewed by abstract. A search of the EMBASE database returned 441 cases, of which 172 were sufficiently relevant for abstract review. The results of both searches were combined, resulting in 206 citations after duplicates were removed. A review of these citations revealed 28 distinct eConsult services that met our definition of eConsult (ie, asynchronous platforms that allow PCPs and specialists to discuss a patient's care). Additional 25 services emerged from the gray literature search, resulting in 53 eConsult services from 17 regions (16 countries plus one international service). The United States had the highest number of identified services (n=28), followed by Canada (n=4), Brazil (n=3), and Spain (n=3). Figure 1 presents a map of all services.

We sent emails to representatives from 49 services (Canadian services, including our own, were excluded from interview recruitment to avoid bias). Representatives from 11 services responded to our emails and completed telephonic interviews. In 2 cases, we held joint interviews with 2 representatives from the service. In another case, 2 separate interviews were conducted about the same service because the initial respondent recommended that we interview another representative. One of the services we interviewed was excluded from our analysis because it was still in its preliminary stages and had not yet developed an eConsult platform. Our final dataset, thus, consisted of 11 interviews with 13 representatives from 10 eConsult services in 4 countries. Respondents held a number of roles, including researchers (n=3), PCPs (n=2), specialists (n=2), managers or directors (n=2), and chief executive or medical or information officers (n=4) and represented a range of service types, varying in size, technology leveraged, and funding model. Table 1 describes the service characteristics.

The thematic analysis of the interviews revealed 4 themes as follows: service structure, benefits, implementation challenges, and implementation enablers (Figure 2).

Service Structure

Respondents discussed a number of issues pertaining to the structure of their eConsult service, including its usage, platform, implementation, and payment.

Usage

Usage patterns varied considerably between services, which operated in a range of environments and at vastly different scales. For instance, the Bradford Teaching Hospitals eConsult service offers different single-specialty services, among the largest of which—renal medicine—handles roughly 30 cases a month answered by a single nephrologist, whereas the Veteran's Health Administration's New England region processed 90,600 cases in 2015 alone.

Figure 1. Map of services that were identified by the environmental scan (n=53) and participated in interviews (n=10).

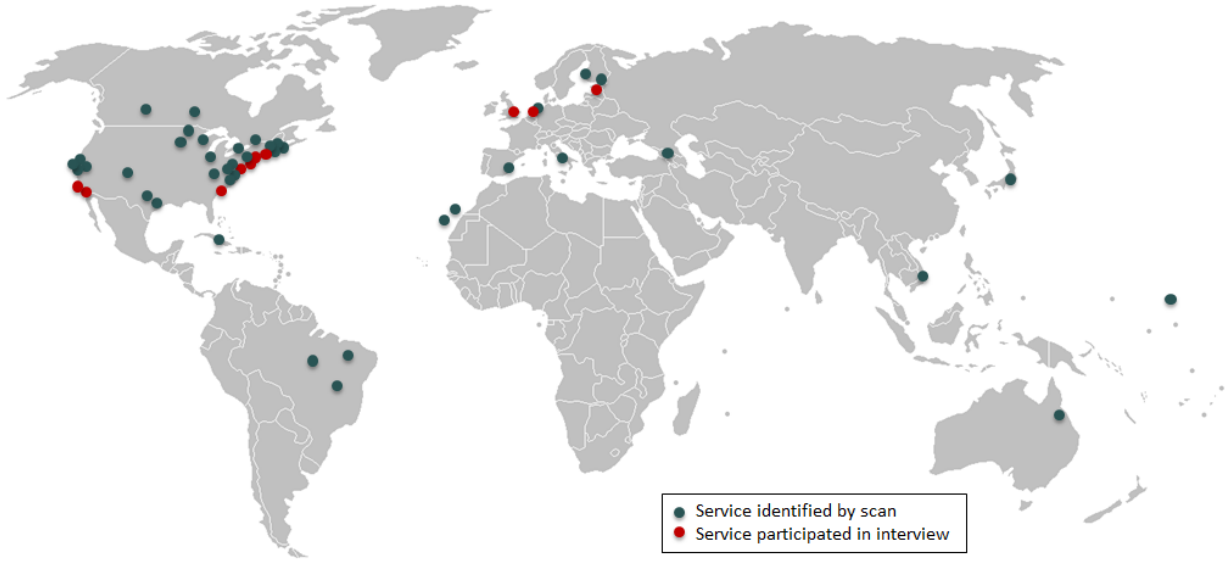


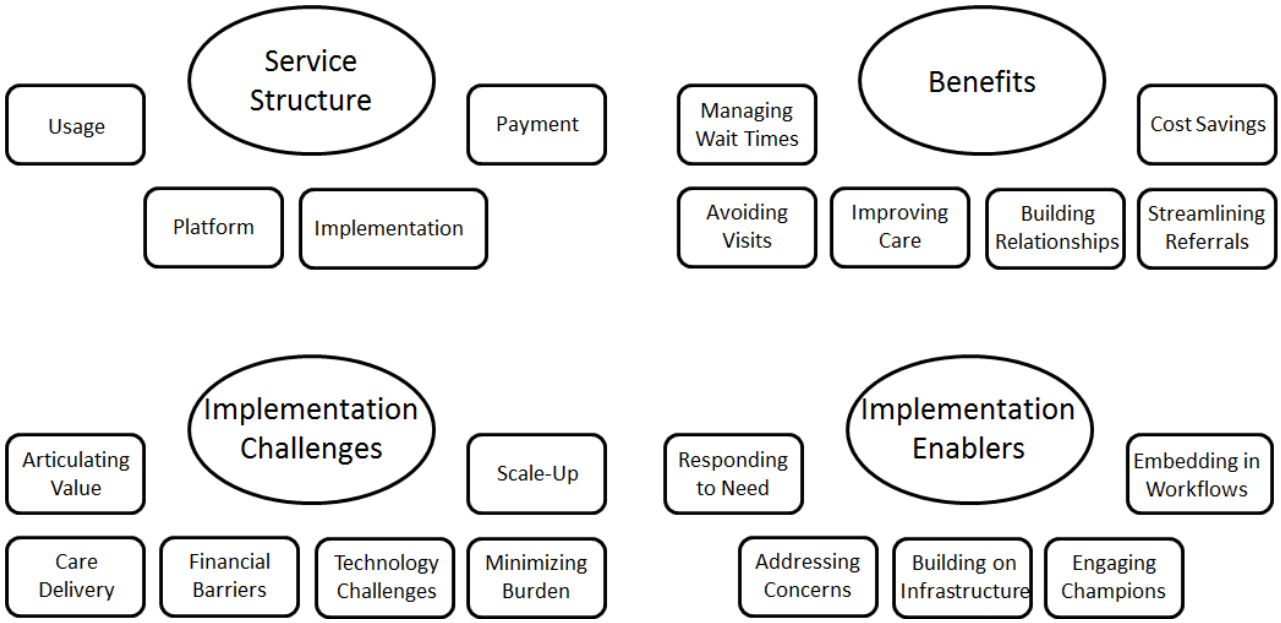
Table 1. The characteristics of services discussed in telephonic interviews.

Name	Country	Active since	Host organization	Tech platform	Payment model
Estonian Health Information System	Estonia	2011	Government	EMR[a]	Nonprofit
ZorgDomein	Netherlands	2001	Business	EMR	Profit
Bradford Teaching Hospitals	UK	2005	Hospital or clinic	EMR	Nonprofit
AristaMD	US	2014	Business	EMR	Profit
Los Angeles Dept Health Services	US	2012	Government	Web[b]	Nonprofit
NYC Health + Hospitals	US	2015	Hospital or clinic	EMR	Nonprofit
CHC Association of Connecticut	US	2017	Nonprofit	EMR	Nonprofit
Veteran's Health Administration	US	2011	Government	EMR	Nonprofit
Duke Institute for Health Innovation	US	2016	Research institute	EMR	Nonprofit
RubiconMD	US	2013	Business	Web	Profit

[a]EMR: electronic medical record.

[b]Web: browser-based Web application.

Figure 2. Map of themes and subthemes.



Platform

All respondents' services utilized 1 of 2 main platforms—those integrated into electronic medical records used by participating clinics, and those hosted on the Web and accessed through a Web browser. However, platforms varied considerably within these categories. In some cases, eConsult functioned as part of the referral process, with all referrals automatically made eligible for eConsult. For instance, in the Los Angeles Department of Health Services, "eConsult is the mandated way to request nonurgent, nonemergent outpatient specialty care services from us. There is no other pathway" (Respondent 11). Others, such as RubiconMD, offer "a Web-based and also mobile app-based eConsult platform" (Respondent 9) through which PCPs can submit eConsults if they so choose.

Implementation

Respondents' services were at various stages of implementation, with some well-established services having operated for years, whereas others were only recently launched and still in their pilot phases. Many respondents described the implementation as a gradual process that leveraged grassroots connections, beginning in one instance as "a bottom-up initiative between one family doctor and one hospital" (Respondent 1). Another respondent described the initial service he worked on as operating largely independently alongside a handful of sympathetic providers:

> *We deliberately went under the radar to start with because we thought there'd be a lot of red tape trying to get this approved. We just thought it was such an obvious thing to bring advantage to patients that we should generate some under-the-radar momentum and enthusiasm and run with that.* [Respondent 2]

Payment

Participating services included for-profit businesses, as well as nonprofit organizations affiliated with universities, hospitals, and regional or national governments. As such, payment mechanisms varied widely based on the objectives of the organization and the health system of the country in which it operated. In US-based systems, the payee was typically a patient's insurer or, for some populations (eg, safety-net services for low-income individuals, the Veteran's Health Administration) the state-funded Medicare or Medicaid. In countries with universal health care (eg, the United Kingdom and Estonia), payment came directly from the government. Some services remunerated PCPs and specialists for participating, whereas others—particularly those that had integrated eConsult into the referral process—considered it an extension of the provider's regular duties and provided no additional or alternate means of payment. For instance, in the Los Angeles Department of Health Services' system:

> *[PCPs] don't see any change in their revenue as a result of using the system or not. The incentive for them to use the system is that this is how they get referrals to their patients.[...]The same goes for the specialists, because in the safety-net systems every doctor just has a flat salary and the whole system is just a flat capitated system.* [Respondent 6]

Benefits

Participants described a number of benefits that eConsult provided—managing wait times, avoiding unnecessary visits, improving the quality of care, streamlining the referral process, building provider relationships, and cost savings.

Managing Wait Times

Many respondents cited rapid turnaround times as a major benefit of eConsult, noting that their service has helped manage wait times for patients seeking specialist advice, "by doing an eConsult you're getting all the patients immediate specialist impact by getting someone to weigh in on their care plan" (Respondent 4). Several respondents noted that eConsult provided much-needed relief in areas where wait times were substantial, "There was a pretty significant backlog of referrals that hadn't been managed at one of the health centers. And so they're using this pilot as an opportunity to clear out that backlog" (Respondent 7). Respondents stressed how patients benefit from better management of wait times, "It's also good for the patients as well to get that feedback quickly" (Respondent 2).

Avoiding Unnecessary Visits

Several respondents stated that their eConsult service "in many cases helps to avoid a referral" (Respondent 9). Respondents noted the benefit this has for patients, as many of them are able to receive care without the long waits and inconvenience associated with a specialist referral.

Improving Quality of Care

Respondents also discussed how eConsult services improve the quality of care patients receive; this improvement was multifaceted and extended beyond the speed of replies and capacity to avoid unnecessary specialist visits. As one respondent noted:

> *You can improve the quality of care, you can improve the speed of care, you can reduce the cost of care. There are so many aspects associated to teleconsultation.* [Respondent 3]

While promptness and efficiency emerged as key benefits, respondents argued that eConsult still had value in cases where a face-to-face consultation was required, as it allowed PCPs to better support patients prior to the specialist consultation. As one respondent described:

> *A third [of cases are] new work, a third avoid a live visit and a third don't avoid a live visit, but it may actually prepare patients and providers for the live visit better by having trialed a change in medicine before they see the specialist. Or allow the [PCP] to order certain tests that then would be available to the sub-specialist at the time of the visit.* [Respondent 8]

Streamlining the Referral Process

Another benefit of eConsult was its ability to "streamline the referral process" (Respondent 6). One respondent described her service as providing a kind of triage, allowing patients who can be treated at the primary care level to avoid unnecessary visits while freeing up space for those who require face-to-face specialist referrals:

> *For patients who have higher acuity issues that do need a face-to-face visit, you're able to identify those patients and expedite them. And because you're clearing out these lower acuity patients from the waitlist to see the specialist, you're seeing a huge opening of access to getting face-to-face [appointments]...by giving them earlier face-to-face care by the specialist, you're not seeing patients sitting for months and months on a waitlist, getting worse, and then having some acute event and ending up in the E.R.* [Respondent 4]

Another respondent noted that eConsult's inherent tracking of consultation requests improved accountability by "making sure that every referral gets a specialist's eyes on it and gets some follow-up" (Respondent 5).

Building Provider Relationships and Empowering Primary Care Providers

Several respondents mentioned that the interprovider connections fostered by eConsult can help build relationships between PCPs and specialists. In addition, eConsult can help empower PCPs by providing them with the necessary guidance to perform a broader scope of patient care. As one respondent noted, PCPs who use eConsult "feel that they can provide more [health care services] than expected of them initially" (Respondent 1).

Cost Savings

Finally, several respondents discussed eConsult's ability to save money for patients and the health care system. Respondents

noted that a case answered by eConsult costs substantially less than a face-to-face specialist visit:

> *Keeping the patient at the primary care, that's the least expensive setting to treat a patient in. [Payers] recognize immediate return on their investment just from avoiding the more expensive specialist visits. And the things that come along with the specialists visits that are often these extremely extensive workups that may or may not be necessary, right. [...] So you're seeing a reduction in things like E.R. visits and hospital admissions, that's where gigantic, really, savings come into play.* [Respondent 4]

Implementation Challenges

Respondents mentioned several challenges associated with implementing eConsult—articulating service value, ensuring care is effectively delivered, financial barriers, technological challenges, minimizing provider burden, and scale-up.

Articulating Service Value

When discussing implementation challenges, nearly all respondents mentioned that they found it difficult to convince stakeholders of eConsult's value. Often this challenge occurred at the management level, with respondents struggling to secure investment in the implementation from leaders who were skeptical of the service's efficacy, "the initial challenge was actually convincing people that providers would use this, if it was made available" (Respondent 9). Convincing providers to engage was also sometimes a challenge, though in their case, it was more a question of fighting inertia and getting practitioners to adjust to new methods of delivering care:

> *The greatest challenge was getting people to think about their work differently. Specialists with the viewpoint that "how can I possibly care for somebody that I haven't seen face-to-face personally and laid my own hands on them?" Getting them to think about delivering specialty care through this interaction with a primary care physician. Getting PCPs to think about this not as extra work, [but] as an actual patient-centric intervention, because you are setting up a communication with the specialist.* [Respondent 11]

Ensuring Care is Effectively Delivered

According to a few respondents, one of the main challenges with eConsult is ensuring that the service consistently delivers appropriate care. These services tended to be nonprofit organizations that dealt with vulnerable patients and faced limitations in staffing, which at times made it difficult to reach patients and follow up with the advice received through eConsult:

> *Since we're a safety-net system there are often concerns with having accurate contact information for patients. Some may change phone numbers, some may not have been comfortable giving us a phone number. [...]Capacity is really an issue for us.* [Respondent 5]

A respondent from another service noted the particular challenges associated with using eConsult for urgent cases:

> *If you need urgent specialty care you're still kind of stuck sitting sometimes in emergency room or begging the specialist, the office, to squeeze somebody in. And it's hard to get that kind of urgent access. [Respondent 6)*

Financial Barriers

A few respondents cited financial issues as a challenge to eConsult implementation; these included the logistics of paying providers, as well as securing sufficient funds to implement and run the service. Respondents spoke of the need for buy-in from decision makers capable of financing the service "through a pilot or for some seed money to get it off the ground" (Respondent 9), some of whom were reluctant to support new or unproven programs:

> *I think the biggest challenge for us has been the politics of some of this with the CEOs who look at this and say 'yeah, that's great. But how am I going to get paid? And how am I going to make money from this? Or how am I going to cover my costs?'"* [Respondent 7]

In addition, one respondent noted that their eConsult service lacked "formal reimbursement mechanisms," and that it was a challenge to develop "a payment mechanism to support the delivering of eConsult" (Respondent 10). This challenge extended to articulating the value eConsult delivered to patients without an existing business case model.

Technological Challenges

Several respondents described technical challenges in eConsult implementation. However, these issues were characterized not as serious issues but as inconveniences or growing pains associated with implementing any new system:

> *You're going to run into some things where the information isn't processing right or there's something screwy in the EHR or whatever. [...]It's just a matter of working through those issues.* [Respondent 7]

This ran counter to some expectations in implementing a technical innovation. One respondent noted that his team "anticipated incorrectly that the main challenge would be technical" (Respondent 10).

Minimizing Provider Burden

When discussing their eConsult services, several respondents emphasized the need to minimize the burden of usage it placed on PCPs and specialists. While respondents viewed eConsult as time-saving for the system overall, they noted that adopting the service meant fitting new tasks into extremely busy workflows, an action which some providers resisted:

> *Whenever you change something there's always new challenges. [...]PCPs have to make a larger investment in the conversation with the specialists to get their patient in for specialty care, [while specialists] need to have a more robust conversation*

> *with the PCPs in order to manage the patient. And so probably our biggest area of complaint or pushback has been the PCP is feeling like it's more work.* [Respondent 11]

Scale-Up

A few respondents articulated ongoing challenges with scale-up, as their initial services attempt to serve a broader scope of patients over a wider area. Respondents noted that at a larger scale, issues such as payment and service delivery must be more formalized, as structures that worked for a few hundred providers may no longer work with a user base in the thousands.

Implementation Enablers

Respondents described a number of factors that contributed to the success of their services—responding to an existing need, addressing providers' concerns and frustrations, building on existing infrastructure, engaging clinical champions, and embedding into provider workflows.

Responding to an Existing Need

The most commonly cited enabler for the successful implementation was answering a need that had been articulated by the target population; this need might stem from a policy initiative enacted by regional or national decision makers or from providers frustrated with the current state of affairs. As one respondent described:

> *We had very long wait times. Many of our specialties had specialty care wait times over 6 months. Some more than a year. There was...the black hole phenomenon where a request would come into us and it would disappear.* [Respondent 11]

A successful service will…

> *...build in the right cultural and financial system to make sure that incentives are aligned. So that PCPs have a reason to use it, specialists have a reason to be courteous and timely.* [Respondent 6]

Building on Existing Infrastructure

When designing an eConsult service, many respondents found it advantageous to leverage existing platforms. In many cases, this consisted of an electronic medical records, which had the benefit of already offering a secure digital link between providers and clinics. By harnessing the established infrastructure, respondents were able to build their services at a fraction of the time and cost it would have taken to develop a wholly independent system. One respondent, describing the creation of an eConsult service inside an established network, stated, "I was almost stunned at how straightforward it was" (Respondent 10).

Engaging Clinical Champions

Several respondents spoke to the importance of engaging clinical champions early in the implementation process. These individuals were PCPs or specialists who believed strongly in the service, used it often, and advocated on its behalf to their colleagues. As the primary end users of eConsult, health care providers are uniquely positioned to offer feedback on how the service works, and respondents stated that their advocacy lent momentum and legitimacy to the project. In the words of one respondent:

> *Having those clinical champions as true believers upfront has made all the difference in the world.* [Respondent 7]

Embedding Into Provider Workflows

Several respondents underscored the importance of developing a service that fits "[as] seamlessly as possible into the clinician's workflow. Because these guys are really strapped for time." Ease of use was critical to successful adoption, and respondents described taking pains to cut out any extraneous or cumbersome elements from the application:

> *Understanding the limitations that your teams have on a day-to-day basis and the bottlenecks that they experience has been really critical for us. [...]We had the time to really implement, see how things were going, find out that "x" component here was a few more clicks than it really needed to be, and that was a barrier for staff. And we could resolve that and improve that workflow.* [Respondent 5]

Addressing Providers' Concerns and Frustrations

To support buy-in from providers, several respondents made a point to seek user feedback regularly throughout the implementation process and address their concerns. Respondents stressed that to get physicians to consider using eConsult, it has to be, at least, as effective and easy to use as the traditional referral-consultation process:

> *The main selling point for the service has been the commonsense nature of it and the fact that it works well for [PCPs] and it works well for [specialists].* [Respondent 2]

Discussion

Principal Findings

This study found that eConsult services are being implemented in countries around the world. Services can take a number of different forms, with variations in scope, technology platform, financial structure, and engagement strategy. They did not come predominantly from any one sector, emerging as private companies, research pilots, government initiatives, and extensions of existing hospitals or health care clinics. Despite these differences, respondents frequently described facing similar barriers in their implementation and cited common factors that enabled the successful implementation and growth of their services. Gaining interest from stakeholders, ensuring the service effectively meets its stated aims, and securing financial support were among the most frequently cited barriers, while engaging clinical champions, building on existing infrastructure, and addressing an existing need emerged as the main enablers of success.

Limitations

This study has several limitations. Of 53 services identified by the environmental scan, only 11 participated in interviews (10 of which were included). Services from the United States are

disproportionately represented, making generalization to other countries more difficult; this limitation is exacerbated by our ability to conduct interviews in only 2 languages (English and French). Although the effort was made to contact all services regardless of their location, our lack of fluency in other languages likely hindered our ability to recruit participants. In addition, all health care providers who participated in this study were physicians. As such, the views of other eConsult users (eg, nurse practitioners) may not have been reflected.

Comparison With Prior Work

Among enablers, addressing an existing need was often described as a particularly important step. All services in this study emerged to address a common problem of poor access to specialist care, with individual approaches tailored to address each service's target population. This approach reflects our own experience with the Champlain BASE eConsult service. Our team created eConsult as a direct response to excessive wait times for specialist care, which remains a significant and ongoing problem in Canada. A 2016 survey by the Commonwealth Fund assessed 11 countries on measures of the health care quality, including access to care. Canada ranked last on wait times for specialist care, with 56% of patients waiting ≥4 weeks for an appointment versus an average of 36% [14]. The severity of this issue drove the Champlain BASE eConsult service's implementation in our region. Likewise, a number of respondents in this study built their own services around the needs of their communities. For instance, in the Commonwealth Fund survey cited above, the United States fared relatively well on the metric of specialist wait times—ranking third out of 11 participants—but faced a number of substantial barriers related to equity and cost of care [14]. As such, several of the United States-based services in this study developed their programs with a lens toward improving equity. Notably, several were "safety net services" specifically designed to help vulnerable individuals who lacked private insurance.

Encouragingly, eConsult is a flexible and multifaceted solution and has shown itself to be well-positioned to address the wide range of access issues presented by communities in different countries. Respondents witnessed a wide range of benefits of their eConsult services, including their ability to avoid unnecessary specialist visits, improve the overall quality of care, reduce costs, and improve communication between providers. These assertions are supported by the literature, which has reported many of the same benefits for eConsult services [7,8]. A systematic review conducted in 2015 identified 27 peer-reviewed papers discussing eConsult services and found high levels of provider satisfaction (70%-95%), quick response times (<3 days in most cases), and avoidance of unnecessary referrals [7]. A systematic review by our team found similar results, as well as some evidence of reduced costs [8].

Future of eConsult

The breadth of eConsult services now operating worldwide suggests a promising future for this model of health service delivery. In many cases, regional health authorities have integrated eConsult into the fabric of the health system, making it a mandatory component of the referral-consultation process. Other systems, including Champlain BASE, are supplemental and voluntary, relying on the provider and patient interest to drive engagement. While barriers to the eConsult's expansion exist and must be addressed [15], the overall picture is encouraging, as evidenced by the experiences highlighted in this study. Furthermore, our efforts at the expansion have been highly successful; the service is currently expanding province-wide, and the College of Family Physicians of Canada recently released a statement identifying eConsult as a standard of practice.

The growing focus on eConsult as a method of improving patients' access to care can be seen as an extension of the Patient's Medical Home, a model of health service delivery that emphasizes that each patient should have a dedicated family practice that serves as "the central hub for the timely provision and coordination of a comprehensive menu of health and medical services patients need" [16].

The goal of the Patient's Medical Home fits naturally into eConsult, as such services allow PCPs to take a more central role in their patients' care. By using eConsult, PCPs are often able to gain the guidance they need to treat patients themselves when they would otherwise have referred them, and its capacity for direct interprovider communication improves care coordination and reduces the risk of cases being forgotten or recommendations lost.

Conclusions

eConsult services have emerged in a variety of countries and health system contexts worldwide. Structure, platform, and delivery model varied, but the services consistently demonstrated improved access and high levels of satisfaction. Respondents encountered several barriers to implementation but were able to overcome them by addressing an existing need and working with engaged clinician leaders. Lessons learned from this group will be helpful for those looking to implement an eConsult service in their own jurisdictions.

Acknowledgments

We would like to thank Canada Health Infoway for their support, and the individuals who participated in our interviews. Funding was provided by the Ontario Ministry of Health and Long-Term Care, the Champlain Local Health Integration Network, and the Canadian Institutes of Health Research. The funders played no part in the study design, collection, analysis, or interpretation of the data, in the writing of the report, or in the decision to submit the paper for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy matrix for literature review.

[PDF File (Adobe PDF File), 21KB - jmir_v20i12e11112_app1.pdf]

Multimedia Appendix 2

Strategy used for gray literature search.

[PDF File (Adobe PDF File), 19KB - jmir_v20i12e11112_app2.pdf]

Multimedia Appendix 3

Semistructured interview guide.

[PDF File (Adobe PDF File), 23KB - jmir_v20i12e11112_app3.pdf]

References

1. Barua B, Esmail N. Fraser Institute. Vancouver, British Columbia: Fraser Institute; 2013 Oct. Waiting you turn: Wait times for health care in Canada, 2013 Report URL: https://www.fraserinstitute.org/sites/default/files/waiting-your-turn-2013.pdf [accessed 2018-09-17] [WebCite Cache ID 72VClVX7b]
2. Day B. The Consequences of Waiting. In: Globerman S, editor. Reducing Wait Times for Health Care - What Canada Can Learn From Theory and International Experience. Vancouver, British Columbia: Fraser Institute; 2015.
3. Barua B, Esmail N, Jackson T. Fraser Institute. Vancouver, British Columbia: Fraser Institute; 2014 May. The Effect of Wait Times on Mortality in Canada URL: https://www.fraserinstitute.org/sites/default/files/effect-of-wait-times-on-mortality-in-canada.pdf [accessed 2018-09-17] [WebCite Cache ID 72VCvAk7x]
4. Gray BH, Bowden T, Johansen I, Koch S. Electronic health records: an international perspective on "meaningful use". Issue Brief (Commonw Fund) 2011 Nov;28:1-18. [Medline: 22164356]
5. Straus SG, Chen AH, Yee H, Kushel MB, Bell DS. Implementation of an electronic referral system for outpatient specialty care. AMIA Annu Symp Proc 2011;2011:1337-1346 [FREE Full text] [Medline: 22195195]
6. Chen AH, Murphy EJ, Yee HF. eReferral--a new model for integrated care. N Engl J Med 2013 Jun 27;368(26):2450-2453. [doi: 10.1056/NEJMp1215594] [Medline: 23802515]
7. Vimalananda VG, Gupte G, Seraj SM, Orlander J, Berlowitz D, Fincke BG, et al. Electronic consultations (e-consults) to improve access to specialty care: a systematic review and narrative synthesis. J Telemed Telecare 2015 Sep;21(6):323-330 [FREE Full text] [doi: 10.1177/1357633X15582108] [Medline: 25995331]
8. Liddy C, Drosinis P, Keely E. Electronic consultation systems: worldwide prevalence and their impact on patient care-a systematic review. Fam Pract 2016 Jun;33(3):274-285. [doi: 10.1093/fampra/cmw024] [Medline: 27075028]
9. Liddy C, Hogel M, Blazkho V, Keely E. The current state of electronic consultation and electronic referral systems in Canada: an environmental scan. Stud Health Technol Inform 2015;209:75-83. [Medline: 25980708]
10. Patton M. Purposeful Sampling. In: Qualitative evaluation and research methods. Beverly Hills, California: Sage Publications; 1990.
11. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
12. Silverman D. Doing qualitative research: A practical handbook. In: Doing qualitative research: A practical handbook. Thousand Oaks, California: Sage Publications; 2013.
13. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
14. Canadian Institute for Health Information. Ottawa, Ontario: CIHI; 2016. How Canada Compares: Results From The Commonwealth Fund's 2016 International Health Policy Survey of Adults in 11 Countries URL: https://www.cihi.ca/sites/default/files/document/text-alternative-version-2016-cmwf-en-web.pdf [accessed 2018-09-17] [WebCite Cache ID 72VD8bs1Z]
15. Liddy C, Joschko J, Keely E. Policy Innovation is Needed to Match Health Care Delivery Reform: The Story of the Champlain BASE eConsult Service. HRO-ORS 2015 Dec 10;3(2):1-11. [doi: 10.13162/hro-ors.v3i2.2747]
16. The College of Family Physicians of Canada. 2011 Sep. A Vision for Canada: Family Practice - The Patient's Medical Home URL: https://www.cfpc.ca/uploadedFiles/Resources/Resource_Items/PMH_A_Vision_for_Canada.pdf [accessed 2018-09-17] [WebCite Cache ID 72VDEG31u]

Abbreviations

BASE: Building Access to Specialists through eConsultation
eConsult: electronic consultation

eReferral: electronic referral
PCP: primary care provider

Edited by G Eysenbach; submitted 22.05.18; peer-reviewed by L Oppenheimer, G Mazowita, J Pecina; comments to author 11.07.18; revised version received 18.07.18; accepted 18.07.18; published 21.12.18.

Please cite as:
Joschko J, Keely E, Grant R, Moroz I, Graveline M, Drimer N, Liddy C
Electronic Consultation Services Worldwide: Environmental Scan
J Med Internet Res 2018;20(12):e11112
URL: http://www.jmir.org/2018/12/e11112/
doi:10.2196/11112
PMID:30578187

Viewpoint

Primary Care Patient Records in the United Kingdom: Past, Present, and Future Research Priorities

Brian McMillan[1], BSSc (Hons), MBChB, PhD, MRCGP; Robert Eastham[2], BA (Hons), MBChB, PGCert Ed, MRCGP; Benjamin Brown[1,3], MSc, MPH, PhD, MRCGP; Richard Fitton[4], MBBS, DCH, DRCOG, MRCGP; David Dickinson[5], BEd (Hons)

[1]Centre for Primary Care and Health Services Research, Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

[2]Whitehall Surgery, Wortley Beck Health Centre, Lower Wortley, Leeds, United Kingdom

[3]Centre for Health Informatics, Division of Informatics, Imaging & Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

[4]West Pennine Local Medical Committee, Barley Clough Medical Centre, Nugget Street, Oldham, United Kingdom

[5]Unlike Minds, Gorton Monastery, Manchester, United Kingdom

Corresponding Author:
Brian McMillan, BSSc (Hons), MBChB, PhD, MRCGP
Centre for Primary Care and Health Services Research, Division of Population Health, Health Services Research and Primary Care
School of Health Sciences, Faculty of Biology, Medicine and Health
University of Manchester
Williamson Building
Oxford Road
Manchester, M13 9PL
United Kingdom
Phone: 44 0161 2757662
Email: brian.mcmillan@manchester.ac.uk

Abstract

This paper briefly outlines the history of the medical record and the factors contributing to the adoption of computerized records in primary care in the United Kingdom. It discusses how both paper-based and electronic health records have traditionally been used in the past and goes on to examine how enabling patients to access their own primary care record online is changing the form and function of the patient record. In addition, it looks at the evidence for the benefits of Web-based access and discusses some of the challenges faced in this transition. Finally, some suggestions are made regarding the future of the patient record and research questions that need to be addressed to help deepen our understanding of how they can be used more beneficially by both patients and clinicians.

(J Med Internet Res 2018;20(12):e11293) doi:10.2196/11293

KEYWORDS

primary care; access to records; medical records; computerized records

A Brief History of the Medical Record

The history of medical records can be dated back as far as the Edwin Smith papyrus of 1600 BC, which describes 48 surgical case histories and was most likely written as an Egyptian surgical manual [1]. Later examples include the case histories of Hippocrates from around 400 BC [2] and medieval Islamic texts from around AD 925, which were largely adapted from Graeco-Roman case histories [3]. Throughout the centuries, medical records were mainly used for teaching purposes [4], and the popularity of cadaveric dissection in the 17th century focused on the use of case histories for the teaching of anatomy [5]. By the 1700s, the keeping of case history books by physicians was becoming more commonplace [6], and medical centers were keeping increasingly detailed patient records toward the end of that century and into the 1800s [7,8]. In the late 1800s, attempts were made to control the content and quality of hospital records for insurance and medicolegal purposes [7], but it was common at this time for physicians to keep their private notes separately to aid patient care [4].

In United Kingdom, Lloyd George's National Insurance Act of 1911 made it compulsory for employed men aged 16-70 years to take out health insurance, and for general practitioners (GPs) providing their care to keep a written record of these patients [9]. While the content and layout of the record were not stipulated, their size was determined by the tin storage boxes provided by the government at that time [10]. These metal boxes were later replaced by envelopes, but the size of the primary care record persisted after the introduction of the National Health Service (NHS) in 1948 [10]. Early criticisms of the format of general practice records focused on the inconvenience caused by the small size of the envelopes, and the absence of a separate problems list [10]. To overcome these problems, there were calls for primary care surgeries to change to records in an A4 format in the 1960s and 1970s, but these failed to materialize [10]. Such concerns were soon to be made redundant by the introduction of computerized records systems [9].

Transition to Electronic Records

The history of computerized records in general practice can be traced back to Exeter in 1970 when John Preece became the first GP to use a computer in the consulting room [11]. The first government-sponsored electronic records system involved a small pilot by the Department of Health in Exeter in 1972 [9]. Ten years later, the government-sponsored "Micros for GPs" involving 150 UK practices, laying the foundations for further innovations [9]. In 1987, 2 private companies began offering computer systems to general practices free of charge with a plan to offer anonymized data to pharmaceutical companies to recoup their initial investment [11]. These schemes were hugely popular with GPs and this, coupled with remuneration changes in 1990, resulted in an exponential growth in the number of GP practices using computerized systems [9]. While <5% of GP practices used electronic records in the early 1980s, this increased to 80% in 1992 as government incentives continued [9] and by 1996, 96% of general practices used computerized record systems [11].

Evolving Functions of the Electronic Record

While the functions of the paper-based patient record expanded slowly over the centuries, the computerization of medical records in primary care has opened up a wealth of additional functionality. The functions of the electronic patient record can be roughly categorized into clinical, administrative, and statistical, although there is some degree of overlap. The electronic record continues to be used primarily as a clinician's aide memoir, enabling primary care staff to see what was discussed at previous appointments or refer to a list of patients' current and previous medical problems. Clinical tasks, such as prescribing, have become easier, safer, and more cost-efficient as electronic record systems can flag allergies, contraindications, potential drug interactions, and suggest lower cost-generic alternatives. Some electronic record systems link to knowledge databases, such as the National Institute of Health and Care Excellence Clinical Knowledge Summaries, or provide handy links to patient information leaflets such as those hosted on "patient.info." Computerized records make it easier to ensure patients are followed up in a timely manner through the use of a "recall" function. Clinical audits can be carried out at the push of a button, enabling clinicians to ascertain how patient care can be improved, or identify patients who are slipping through the net.

In addition, administrative tasks are now vastly less labor-intensive. Keeping an up-to-date list of patients containing accurate demographic and clinical information no longer requires meters of filing cabinet; letters to patients and other specialties can be prepopulated with important information from a patient's record; and patient record transfers between GP surgeries is now increasingly an electronic process. Moreover, electronic record systems are used in the financial management of practices, for purposes such as securing reimbursement, budget planning, and reducing costs. Furthermore, the electronic patient record system can be used to enable secure communication between members of staff, reducing the risk of tasks being left undone and with the added benefit of an audit trail.

Computerized primary care records also provide a wealth of statistical information. The UK government has long seen the potential value of collecting such information [10], and there have been ill-fated attempts to monetize this information in the past by private companies [11]. The early GP computer enthusiasts designed computer systems to collect epidemiological data, and this tradition has continued to this day. Research using the Clinical Practice Research Datalink, which holds data on over 11.3 million patients from 674 UK practices [12], has resulted in a multitude of improvements in patient care and over 1800 scientific publications [13]. There is a growing interest in using machine learning approaches to define disease phenotypes in electronic primary care health records [14] while others are using statistical techniques used in astrophysics to develop predictive models of disease from the Clinical Practice Research Datalink [15].

In addition, the patient record can now be used by clinicians to send referrals directly to secondary care. Standardizing information flow between referrer and service provider is becoming an increasingly important function of clinical systems. A 2016 audit of suspected cancer referrals in Leeds found that only 48% were completed with the minimum required clinical information; this can lead to a delay in investigation and diagnosis. By leveraging existing functionality within SystmOne, the "DART" project to streamline the referrals process led to 100% of forms completed correctly within 3 months of introduction [16].

Projects such as "DART" illustrate how clinical systems have the potential to both improve patient safety and free-up much needed clinical resources. However, some initiatives to improve patient outcomes by harnessing the functionality within clinical systems may conversely have a detrimental impact on GP workload. The 2016 King's Fund report aimed at "Understanding pressures in general practice" [17] cited the potential for new preventive services to impact the GP workload negatively. Preventive services (such as monitoring of chronic disease) have largely been made possible by recent advances

in clinical systems. However, by linking chronic disease management functions to Quality and Outcomes Framework targets, there is an inevitable pressure for a huge amount of information to be manually read-coded within the record. Failure to do so can have a direct impact on practice income. Mindful of these tensions, it would seem imperative that future initiatives to use clinical systems to improve patient outcomes must take great care not to impact a clinician's workload adversely.

Enabling Patients' Access to Their Own Records

Throughout history, the medical record has traditionally primarily served clinicians and served patients only indirectly. The idea of enabling patients to have full access to their medical record, however, is not entirely new. For example, in 1973, Shenkin and Warner noted,

> *Dissatisfaction with the functioning of the medical care system has become widespread. Four serious problems are maintaining high quality of care, establishing mutually satisfactory physician-patient relations, ensuring continuity and avoiding excessive bureaucracy. We believe these problems could be alleviated, in part, if patients were given copies of all their medical records.* [18, p 688]

Early proponents of granting patients open access to their primary care record included GPs from Balsall Health Centre in Birmingham who started enabling patients to access their full primary care record in 1977 [19], and GPs from Wells Park Road Practice in London who enabled full access from 1983 [20]. Reviews of the impact of promoting such access have shown beneficial effects and minimal risks [21].

The introduction of the Data Protection Act in 1998 gave patients the legal right to access their health records [22], setting the scene for changes to come. While the patient records aspect of the NHS Connecting for Health 2004/2005 business plan focused mainly on providing a single electronic record for health professionals across hospitals, primary care, and community services, it introduced a very limited degree of interactivity through the "chose and book" service [23]. At the same time, however, private companies were developing services that would enable patients to access their own electronic primary care record securely. In 2003, a private company started installing kiosks in GP surgeries that enabled patients to use fingerprint and pin authentication to gain access to their full GP electronic record [24]. By 2006, around 5000 patients had accessed their records in this way, and it was also possible to gain Web-based record access from home [24]. In 2007, the NHS introduced HealthSpace, a Web-based personal electronic health record, which enabled people to enter their health information and gain secure access to the summary care information in their GP record [25].

In 2010, the Department of Health outlined their vision of an information revolution incorporating Web-based access, giving people more control over their health care and improving choice [26]. The same year, the Royal College of General Practitioners published guidelines on enabling patients to access their electronic health records [27] and later published a more detailed "Road Map" on this topic [28]. Despite the British Medical Association's concerns [29], the idea of Web-based patient access was now firmly on the UK government's agenda, and in 2014, the National Information Board published a framework for action incorporating a vision stating,

> *In 2015, all citizens will have online access to their GP records and will be able to view copies of that data through apps and digital platforms of their choice...it is essential that citizens have access to all their data in health and care, and the ability to 'write' into it so that their own preferences and data from other relevant sources, like wearable devices, can be included... This framework prioritises comprehensive access—with the ability for individuals to add to their own records—by 2018.* [30, p 21]

Providing patients with the ability to write in their own health record will facilitate the collection of Patient-Reported Outcome Measures as advocated by Gensheimer et al [31].

The Impact of Web-Based Access to Records

In 2012, to ascertain the impact of enabling patients to access their primary care record online, the Department of Health commissioned a systematic review of the evidence, supported by the Royal College of General Practitioners [32,33]; the review identified 17 randomized controlled trials, cohort, or cluster studies and summarized both the benefits and challenges of providing patients Web-based access to their record.

Potential Benefits of Web-Based Access

Providing patients with Web-based access to their record has been shown to benefit both patients and clinicians. Web-based access enables patients to book appointments online, request repeat prescriptions, and view test results, letters, problems lists, and free-text GP entries [34], although there are wide variations in the degree of access provided by GP surgeries [35]. Patients who use Web-based access report higher levels of satisfaction [36] and improved communication with health care professionals [32]. Benefits to patients include being able to use the Web-based record as an aide memoir and help them prepare for their next appointment [35,37]. Patients like the convenience of Web-based access, stating that it saves time and money, and reduces the number of telephone calls and appointments required [32,35]. In addition, Web-based access can be empowering and increase patients' feelings of autonomy, with one study noting that 77%-87% of patients with Web-based access feel more in control of their care [38]. Other benefits include enabling patients to share their records with family members or other health care providers, or to appoint a proxy to access their record [33]. Web-based access benefits both patients and clinicians in other ways such as improving self-care, increasing the uptake of preventive services, and enabling patients to spot medication errors and have them corrected [32]. The use of Web-based Patient-Reported Outcome Measures built into the patient record increases patients' confidence in managing their condition and has been shown to reduce remission rates for conditions such

as inflammatory bowel disease [39]. One study found that 70% of clinicians reported Web-based access improved trust, strengthened relationships, and enhanced decision making [38], while another found it reduced the annual number of visits and telephone calls [40].

Challenges and Potential Negative Consequences

Despite the many benefits of enabling patients to access their record online, there are also a number of associated challenges. Clinicians have been especially resistant to opening up patient records for Web-based access owing to concerns that it will lead to an increased workload, cause unnecessary anxiety among patients, increase the likelihood of litigation, or challenge the current primary care business model [33]. Other concerns relate to security and confidentiality, equality issues (eg, literacy and internet access), risk of coercion, and information technology system compatibility [28]. The evidence regarding the impact of Web-based access on the clinicians' workload is currently mixed, but there is inevitably an increase in the workload in the early transitional stage, including activities such as staff training [32]. As the patient record was not initially designed to be viewed by patients, the manner in which clinicians write in the notes will have to change if they are to be easily understood by a lay audience. One study, for example, noted that up to 36% of clinicians changed the record content to allow for Web-based access, and up to 21% reported spending more time writing notes [38]. Despite clinicians' concerns regarding Web-based access causing anxiety among patients, leading to an increased risk of litigation, or data security breaches, a review of the studies, to date, has found little evidence these concerns are realized [33]. There is some evidence, however, that Web-based access could potentially lead to increases in health inequalities as those using Web-based access are more likely to be white, female, and middle class [32]. Although one might expect Web-based access to increase patient activation and, thus, improve health outcomes, less activated patients may be less likely to take advantage of Web-based access [41], thus potentially exacerbating health inequalities. Disappointingly, reviews of the literature, to date, reveal a lack of evidence for the impact of Web-based access on health outcomes [32,33], although an up-to-date systematic review is under way [42].

Future Directions

We are still some way from realizing the National Information Board's vision of all UK citizens having read and write access to their full primary care record through a variety of digital platforms that enable them to upload data from wearable devices. Enabling such read and write access could help GPs improve their understanding of the effect of disease and treatment on the everyday lives of patients [39]. The majority of GP practices offering Web-based access do so in a limited way, and although there are some notable exceptions [43], most do not allow access to the clinicians' free-text entries [44]. As De Lusignan et al noted, there is a need for further research to determine "how the medical record might be redesigned to guide and teach patients in a way that promotes self-management and ultimately improves health" [33] (p 7). Such research should be multidisciplinary, drawing upon expertise from fields beyond medicine such as health psychology and human-computer interaction. We need to engage with health economists to ascertain the full economic potential of Web-based access and the impact it may have on the primary care business model. Although some studies using self-report measures exist [35,37,45], further research is also needed to examine how patients actually interact with their Web-based record and the functionality they would like to see. The impact of Web-based access on the patient-clinician relationship and the power dynamic is also worthy of further investigation, especially with regards to the impact of enabling access to the full free-text record. All of these issues underlie what must be our prime concern, and something for which the evidence is still limited, that is, how we can harness the potential of Web-based access to improve health outcomes. Patients' expectations regarding access to their health information are changing, and the newly introduced General Data Protection Regulations [46] will undoubtedly shift the conversation further toward full unrestricted Web-based access. Clinicians will need to change how they view the patient record and learn to work with systems providers and patients to help instigate changes that will lead to improved health outcomes and increased savings for the NHS.

Acknowledgments

BM is employed as a National Institute for Health Research (NIHR) Clinical Lecturer (CL-2016-06-006). The views expressed are those of the authors and not necessarily those of the National Health Services, the NIHR, or the Department of Health.

Conflicts of Interest

None declared.

References

1. Al-Awqati Q. How to write a case report: lessons from 1600 B.C. Kidney Int 2006 Jun;69(12):2113-2114 [FREE Full text] [doi: 10.1038/sj.ki.5001592] [Medline: 16761021]
2. Hippocrates, Adams F. The Genuine Works of Hippocrates. London: The Sydenham Society; 1891.
3. Álvarez Millán C. Graeco-Roman case histories and their influence on Medieval Islamic clinical accounts. Soc Hist Med 1999 Apr;12(1):19-43. [Medline: 11623808]
4. Gillum RF. From papyrus to the electronic tablet: a brief history of the clinical medical record with lessons for the digital age. Am J Med 2013 Oct;126(10):853-857. [doi: 10.1016/j.amjmed.2013.03.024] [Medline: 24054954]

5. Gonzalez-Crussi F. A Short History of Medicine. New York: Modern Library; 2007.
6. Fraser S, Ayres C. The doctor's casebook--a process of discovery. Scott Med J 2009 Aug;54(3):36-41. [doi: 10.1258/RSMSMJ.54.3.36] [Medline: 19725281]
7. Siegler EL. The evolving medical record. Ann Intern Med 2010 Nov 16;153(10):671-677. [doi: 10.7326/0003-4819-153-10-201011160-00012] [Medline: 21079225]
8. Hess V. Formalizing observation: The emergence of the modern patient record exemplified by Berlin and Paris medicine, 1725-1830. Medizinhist J 2010;45(3-4):293-340. [Medline: 21328920]
9. Millman A, Lee N, Brooke A. ABC of medical computing. Computers in general practice--I. BMJ 1995 Sep 23;311(7008):800-802 [FREE Full text] [Medline: 7580446]
10. Tait I. History of our records. Br Med J (Clin Res Ed) 1981 Feb 28;282(6265):702-704 [FREE Full text] [Medline: 6781623]
11. Benson T. Why general practitioners use computers and hospital doctors do not--Part 1: incentives. BMJ 2002 Nov 09;325(7372):1086-1089 [FREE Full text] [Medline: 12424171]
12. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol 2015 Jun;44(3):827-836 [FREE Full text] [doi: 10.1093/ije/dyv098] [Medline: 26050254]
13. National Institute for Health Research. Welcome to Clinical Practice Research Datalink 2018 URL: https://www.cprd.com/intro.asp [accessed 2018-06-13] [WebCite Cache ID 709KsnzuK]
14. Zhou S, Fernandez-Gutierrez F, Kennedy J, Cooksey R, Atkinson M, Denaxas S, UK Biobank Follow-up Outcomes Group, et al. Defining Disease Phenotypes in Primary Care Electronic Health Records by a Machine Learning Approach: A Case Study in Identifying Rheumatoid Arthritis. PLoS One 2016;11(5):e0154515 [FREE Full text] [doi: 10.1371/journal.pone.0154515] [Medline: 27135409]
15. The ASTRODEM project aims to create a predictive model which will help general practitioners (GPs) identify patients at high risk of dementia. 2018 URL: https://www.bsms.ac.uk/research/primary-care-and-population-health/health-informatics/astrodem/index.aspx [accessed 2018-06-13] [WebCite Cache ID 709L6zLkc]
16. Eastham R, Duffy S, Foster C, Beaumont A. DART: a new look at the 2-week wait suspected cancer referral process. Br J Gen Pract 2017 Dec;67(665):560. [doi: 10.3399/bjgp17X693713] [Medline: 29192106]
17. Baird B, Charles A, Honeyman M, Maguire D, Das P. Understanding pressures in general practice. King's Fund. London: The King's Fund; 2016. URL: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Understanding-GP-pressures-Kings-Fund-May-2016.pdf [accessed 2018-06-13] [WebCite Cache ID 73p6VOoK3]
18. Shenkin BN, Warner DC. Sounding board. Giving the patient his medical record: a proposal to improve the system. N Engl J Med 1973 Sep 27;289(13):688-692. [doi: 10.1056/NEJM197309272891311] [Medline: 4727972]
19. Bird AP, Walji MT. Our patients have access to their medical records. Br Med J (Clin Res Ed) 1986 Mar 01;292(6520):595-596 [FREE Full text] [Medline: 3081186]
20. Baldry M, Cheal C, Fisher B, Gillett M, Huet V. Giving patients their own records in general practice: experience of patients and staff. Br Med J (Clin Res Ed) 1986 Mar 1;292(6520):596-598 [FREE Full text] [Medline: 3081187]
21. Ross SE, Lin C. The effects of promoting patient access to medical records: a review. J Am Med Inform Assoc 2003 Apr;10(2):129-138 [FREE Full text] [Medline: 12595402]
22. Office of Public Sector Information. Data Protection Act, 1998 (Public General Acts - Elizabeth II). London: Stationery Office Books; 1998.
23. National Health Service. NHS Connecting for health. Business Plan 2005/2006. 2005. URL: http://webarchive.nationalarchives.gov.uk/20111115164529/http:/www.connectingforhealth.nhs.uk/resources/archive/nhs_cfh_business_plan.pdf [accessed 2018-06-13] [WebCite Cache ID 709Umt76a]
24. Fisher B, Fitton R, Poirier C, Stables D. Patient record access--the time has come!. Br J Gen Pract 2007 Jun;57(539):507-511 [FREE Full text] [Medline: 17550687]
25. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. BMJ 2010;341:c5814 [FREE Full text] [Medline: 21081595]
26. Department of Health. Liberating the NHS: An Information Revolution. A consultation on proposals. 2010. URL: http://webarchive.nationalarchives.gov.uk/20130104222405/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_120598.pdf [accessed 2018-06-13] [WebCite Cache ID 709Vpvhib]
27. Royal College of General Practitioners. Enabling Patients to Access Electronic Health Records. 2010. URL: http://www.rcgp.org.uk/-/media/Files/Informatics/Health-Informatics-Enabling-Patient-Access.ashx?la=en [accessed 2018-06-13] [WebCite Cache ID 73p7zK9kJ]
28. Royal COGP. Patient Online: The Road Map. London: Royal College of General Practitioners; 2013. URL: https://www.researchgate.net/publication/262014118_RCGP_Road_Map_Web_Version_28_Feb_2013 [accessed 2018-06-13] [WebCite Cache ID 73p8q9H1d]
29. Cross M. BMA warns against letting patients have access to their electronic records. BMJ 2011 Jan 12;342:d206. [Medline: 21228010]

30. National Information Board. Personalised Health and Care. Leeds: HM Government; 2014. URL: https://www.digitalhealth.net/includes/images/news0254/PDF/0172_NHS_England_NIB_Report_WITH_ADDITIONAL_MATERIAL_S8.pdf [accessed 2018-06-13] [WebCite Cache ID 73qfaqPNT]
31. Gensheimer SG, Wu AW, Snyder CF, PRO-EHR Users' Guide Steering Group, PRO-EHR Users' Guide Working Group. Oh, the Places We'll Go: Patient-Reported Outcomes and Electronic Health Records. Patient 2018 Dec;11(6):591-598. [doi: 10.1007/s40271-018-0321-9] [Medline: 29968179]
32. Mold F, de Lusignan S, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic review in primary care. Br J Gen Pract 2015 Mar;65(632):e141-e151 [FREE Full text] [doi: 10.3399/bjgp15X683941] [Medline: 25733435]
33. de Lusignan S, Mold F, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic interpretative review. BMJ Open 2014;4(9):e006021 [FREE Full text] [doi: 10.1136/bmjopen-2014-006021] [Medline: 25200561]
34. NHS England. About Patient Online 2017 URL: https://www.england.nhs.uk/patient-online/about-the-prog/ [accessed 2018-04-25] [WebCite Cache ID 709WahhWV]
35. Shah SGS, Fitton R, Hannan A, Fisher B, Young T, Barnett J. Accessing personal medical records online: a means to what ends? Int J Med Inform 2015 Feb;84(2):111-118 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.10.005] [Medline: 25453275]
36. Matheny ME, Gandhi TK, Orav EJ, Ladak-Merchant Z, Bates DW, Kuperman GJ, et al. Impact of an automated test results management system on patients' satisfaction about test result communication. Arch Intern Med 2007 Nov 12;167(20):2233-2239. [doi: 10.1001/archinte.167.20.2233] [Medline: 17998497]
37. Fisher B, Bhavnani V, Winfield M. How patients use access to their full health records: a qualitative study of patients in general practice. J R Soc Med 2009 Dec;102(12):539-544 [FREE Full text] [doi: 10.1258/jrsm.2009.090328] [Medline: 19966130]
38. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. Ann Intern Med 2012 Oct 2;157(7):461-470 [FREE Full text] [doi: 10.7326/0003-4819-157-7-201210020-00002] [Medline: 23027317]
39. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ 2015 Feb 10;350:g7818. [doi: 10.1136/bmj.g7818] [Medline: 25670183]
40. Kummervold PE, Trondsen M, Andreassen H, Gammon D, Hjortdahl P. Patient-physician interaction over the internet. Tidsskr Nor Laegeforen 2004 Oct 21;124(20):2633-2636 [FREE Full text] [Medline: 15534640]
41. Hibbard J, Helen G. Supporting people to manage their health: An introduction to patient activation. London: The King's Fund; 2014 May. URL: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/supporting-people-manage-health-patient-activation-may14.pdf [accessed 2018-06-13] [WebCite Cache ID 73qgGwahI]
42. Ammenwerth E, Lannig S, Hörbst A, Muller G, Schnell-Inderst P. Adult patient access to electronic health records. Cochrane Database of Syst Rev 2017 Jun(6):1-14 [FREE Full text]
43. Hannan A, Webber F. Towards a partnership of trust. Stud Health Technol Inform 2007;127:108-116. [Medline: 17901604]
44. McDermott A. Patient Online: extended session from the NUG Conference. 2015 Sep Presented at: EMIS National Users Group; 23-25 September 2015; East Midlands Conference Centre URL: https://www.emisnug.org.uk/patient-online-extended-session-nug-conference-2015
45. Bhavnani V, Fisher B, Winfield M, Seed P. How patients use access to their electronic GP record--a quantitative study. Fam Pract 2011 Apr;28(2):188-194 [FREE Full text] [doi: 10.1093/fampra/cmq092] [Medline: 21084568]
46. Guide to the General Data Protection Regulation (GDPR). London: Information Commissioner's Office; 2017. URL: https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/ [accessed 2018-06-13] [WebCite Cache ID 709XqzP6y]

Abbreviations

GP: general practitioner
NHS: National Health Service

Edited by G Eysenbach; submitted 27.06.18; peer-reviewed by G Berntsen, A Errazuriz; comments to author 29.08.18; accepted 04.09.18; published 19.12.18.

Please cite as:
McMillan B, Eastham R, Brown B, Fitton R, Dickinson D
Primary Care Patient Records in the United Kingdom: Past, Present, and Future Research Priorities
J Med Internet Res 2018;20(12):e11293
URL: http://www.jmir.org/2018/12/e11293/
doi:10.2196/11293
PMID:30567695

Review

Accessibility and Applicability of Currently Available e-Mental Health Programs for Depression for People With Poststroke Aphasia: Scoping Review

Stephanie Jane Clunne[1], BSpPath (Hons); Brooke Jade Ryan[1], BSpPath (Hons), PhD; Annie Jane Hill[1], BSpPath, PhD; Caitlin Brandenburg[1], BSpPath (Hons), PhD; Ian Kneebone[2], BA (Hons), MPsych, PsychD

[1]School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia

[2]Discipline of Clinical Psychology, Graduate School of Health, University of Technology, Sydney, Ultimo, Australia

Corresponding Author:
Brooke Jade Ryan, BSpPath (Hons), PhD
School of Health and Rehabilitation Sciences
The University of Queensland
84A Services Road, St Lucia
Brisbane, QLD 4067
Australia
Phone: 61 336 51380
Email: brooke.ryan@uq.edu.au

Abstract

Background: Depression affects approximately 60% of people with aphasia 1 year post stroke and is associated with disability, lower quality of life, and mortality. Web-delivered mental health (e-mental health) programs are effective, convenient, and cost-effective for the general population and thus are increasingly used in the management of depression. However, it is unknown if such services are applicable and communicatively accessible to people with poststroke aphasia.

Objective: The aim of this study was to identify freely available e-mental health programs for depression and determine their applicability and accessibility for people with poststroke aphasia.

Methods: A Web-based search was conducted to identify and review freely available e-mental health programs for depression. These programs were then evaluated in terms of their (1) general features via a general evaluation tool, (2) communicative accessibility for people with aphasia via an aphasia-specific communicative accessibility evaluation tool, and (3) empirical evidence for the general population and stroke survivors with and without aphasia. The program that met the most general evaluation criteria and aphasia-specific communicative accessibility evaluation criteria was then trialed by a small subgroup of people with poststroke aphasia.

Results: A total of 8 programs were identified. Of these, 4 had published evidence in support of their efficacy for use within the general population. However, no empirical evidence was identified that specifically supported any programs' use for stroke survivors with or without aphasia. One evidence-based program scored at least 80% (16/19 and 16/20, respectively) on both the general and aphasia-specific communicative accessibility evaluation tools and was subject to a preliminary trial by 3 people with poststroke aphasia. During this trial, participants were either unable to independently use the program or gave it low usability scores on a post-trial satisfaction survey. On this basis, further evaluation was considered unwarranted.

Conclusions: Despite fulfilling majority of the general evaluation and aphasia-specific evaluation criteria, the highest rated program was still found to be unsuitable for people with poststroke aphasia. Thus, e-mental health programs require substantial redevelopment if they are likely to be useful to people with poststroke aphasia.

(J Med Internet Res 2018;20(12):e291) doi:10.2196/jmir.9864

KEYWORDS

aphasia; stroke; depression; mental health; internet; technology; access to health care

Introduction

Background

Aphasia is a language disorder that can impact a person's ability to understand and produce spoken language, read, write, calculate, and use gestures [1]. It is an acquired neurological condition that results from brain damage and occurs in approximately 30% of first-time strokes [2,3]. The communication changes experienced by individuals with poststroke aphasia may cause social exclusion, diminished social networks, activity limitations, reduced life participation, and lower quality of life [4-6]. Depression affects about one-third of stroke survivors without aphasia [7,8] and approximately 60% of stroke survivors with aphasia 1 year post stroke [3]. Negative emotional outcomes after stroke are associated with disability, lower quality of life, and mortality [8]. Hence, poststroke depression is a pressing clinical issue for those with poststroke aphasia.

Talk-based psychological interventions, such as cognitive behavioral therapy (CBT), are widely recommended in the treatment of depression [9-11]. However, a lack of suitably trained professionals, barriers related to cost, distance to care, transport, physical disability, time constraints, stigma, and impaired mobility make psychological interventions relatively inaccessible to many in need [12-16]. Significant communication disability may also prevent people with poststroke aphasia from participating in traditional talk-based psychological treatments [17]. For example, a recent study found that speech pathologists in the United Kingdom often perceived mental health professionals as being underskilled in working with people with poststroke aphasia and that this was a major barrier to referring aphasia patients onto mental health professionals [17]. It is acknowledged that during face-to-face communication, people with aphasia can make use of gesture, facial expression, sign, tone of voice, etc, to help understand others and to get their own message across. Although these communication supports are not offered by e-mental health interventions, there are still many examples of people with poststroke aphasia successfully accessing Web- and computer-based programs and interventions [18-20].

e-mental health is a digital form of mental health care and may offer a solution to the accessibility issues of face-to-face therapy commonly encountered by the general population. e-mental health services provide treatment and assistance to people suffering from mental illness via digital platforms such as computers and Web-based programs [21]. The term *e-mental health* encompasses both e-mental health *literacy*, which is the provision of information pertaining to the nature and treatment of mental health illnesses, and e-mental health *programs*, which are structured self-help programs designed to treat or prevent mental health disorders via an interactive interface [22].

In the context of e-mental health, accessibility refers to the ease with which patients can utilize the health care service in proportion to their needs, as well as the usability of the actual technology through which that service is provided [23]. The increased accessibility and convenience offered by e-mental health programs may enable patients to surpass the barriers associated with limited services, transportation, time, cost, and stigma [13,21]. This increased accessibility may be especially beneficial for patients with poststroke aphasia because of the high incidence of both mood disorders [8] and physical impairment [24] after stroke.

There is a large body of research that supports the use of e-mental health programs for depression; however, this research is generally limited to patients with depression who are otherwise healthy. Although 1 study reported preliminary evidence for computerized CBT (cCBT) in reducing depressive symptoms in people with traumatic brain injury [25] (a population that may also present with acquired communication deficits including aphasia [26]), patients with "insufficient English language skills" were excluded. Thus, such findings cannot be generalized to people with poststroke aphasia. Furthermore, a recent feasability randomized control trial (RCT) of a cCBT intervention for stroke survivors concluded that guided cCBT could potentially increase the accessibility of psychological support for stroke survivors [27]. However, the participants did not specifically have poststroke aphasia, and the communicative needs of stroke survivors with aphasia differ from those without. To the best of the authors' knowledge, these are the only 2 studies that have investigated e-mental health programs directly in people with an acquired brain injury. Therefore, it is not yet known whether such services, and the digital technologies through which they are delivered, are communicatively accessible to people with poststroke aphasia.

Objectives

Previously, a scoping review identified and evaluated currently available e-mental health interventions for depression [28]. This review also acknowledged a lack of e-mental health programs for special populations and recommended that future studies investigate the accessibility needs of such populations so that adequate treatment can be made available to them. To the authors' knowledge, no study has explored e-mental health treatment in terms of its suitability for people with depression and concurrent poststroke aphasia. This is the aim of this study.

Specific objectives were to (1) evaluate the general features of each program, (2) review the published evidence of each program for the general population and for stroke survivors with and without aphasia, (3) evaluate each program's communicative accessibility for people with poststroke aphasia, and (4) determine which e-mental health program(s) may be most suitable for people with poststroke aphasia. It should also be noted that this study did not aim to evaluate how psychotherapeutic concepts, such as CBT principles and abstract concepts, were presented to users in the context of a broader psychotherapeutic community.

Methods

Scoping Review

In all, 2 previous scoping reviews that evaluated available e-mental health interventions for depression [28] and anxiety [29] within the general population were used as the basis of the methodology for this review. The authors aimed to simulate a Web-based search that would likely be carried out by a person

with poststroke aphasia seeking free Web-based treatment for depression.

Search Strategy

A Web-based search for e-mental health interventions for depression was conducted in July 2017. Consistent with previous scoping reviews [28,29], the search engine Google was used for the Web-based search. Prior research has found that people often use search engines, particularly Google, when seeking Web-based health information [30,31]. Hence, people with poststroke aphasia are more likely to find publicly accessible e-mental health programs via search engines rather than through academic resources such as journal databases. The Web-based search consisted of 2 stages: (1) a general Web-based search for e-mental health treatment for depression and (2) a Web-based search for e-mental health depression treatments for stroke survivors with and without aphasia. A total of 12 general search terms were used in the first stage of the Web-based search; they consisted of simple, lay keywords and did not include the words "stroke" or "aphasia." As the most recent E-mental Health Strategy for Australia specifies the Government's investment in Web-based CBT programs [21], many of these search terms relate to CBT. During stage 2 of the Web-based search, the authors collaborated with an academic advisory group, consisting of clinicians and academics who work with people with poststroke aphasia, to generate a set of search terms they thought a person with poststroke aphasia might use if searching for Web-based treatment for depression. These search terms were then combined with the general search terms in stage 1 that yielded the most results. This resulted in 6 aphasia-specific search terms that were used in stage 2 of the Web-based search. All 18 search terms are included in Textbox 1.

Program Screening

Consistent with a previous scoping review [29], the first 25 hyperlinks generated by each search term were screened. This methodology was replicated because it has been found that 75% of users never scroll past the first page of search results [32]. This resulted in 450 hyperlinks being screened. All 450 hyperlinked websites were recorded in an Excel document and, as done in a previous scoping review [29], categorized as (1) websites with e-mental health programs; (2) websites linking to websites with e-mental health programs; and (3) websites with irrelevant content. Irrelevant content included advertisements, scholarly articles, blogs, websites for face-to-face psychology clinics, chatrooms, and support forums. All websites classified as "irrelevant content" and all duplicates identified within the first 2 categories were removed. The remaining hyperlinked websites were then screened using the following inclusion criteria: (1) designed for depressive symptoms in adults; (2) publicly accessible to the general population via the internet; (3) is a structured, self-management program; and (4) free for Australian residents.

Textbox 1. Search terms used in stage 1 and stage 2 of the Web-based search.

Stage 1—general search terms

- Internet therapy for depression
- Internet treatment for depression
- Internet help for depression
- Internet cognitive behavioral therapy for depression
- Web therapy for depression
- Web treatment for depression
- Web help for depression
- Web cognitive behavioral therapy for depression
- Online therapy for depression
- Online treatment for depression
- Online help for depression
- Online cognitive behavioral therapy for depression

Stage 2—aphasia-specific search terms

- Online therapy for depression aphasia
- Online therapy for depression after stroke
- Online treatment for depression aphasia
- Online treatment for depression after stroke
- Online cognitive behavioral therapy for depression aphasia
- Online cognitive behavioral therapy for depression after stroke

Websites were excluded if they (1) provided e-mental health literacy only; (2) offered purely Web-based counseling; (3) were designed for specific populations other than stroke survivors with or without aphasia (ie, adults who stutter); (4) were designed exclusively for adolescents and/or young adults (ie, 25 years or younger); or (5) were not available in English. As some programs required users to complete a depression symptom questionnaire to determine whether the program was suitable for the user before creating an account, the authors contacted all program developers asking for research access to the program. Programs were excluded if research access was not granted.

Data Extraction and Evaluation of Programs

In all, 3 separate tools were used to collect and evaluate the data from each program: a data extraction form, a general evaluation tool, and then an aphasia-specific evaluation tool. The authors used the data extraction form to extract relevant information and data from each program, which then underwent evaluation using the general evaluation and the aphasia-specific evaluation tools. The categories within both the general and aphasia-specific evaluation tools contained a set of closed choice ("yes" or "no") questions. A score of 1 was awarded if the question was answered with a "yes"; a score of 0 was awarded if the question was answered with a "no" or could not be evaluated (ie, due to restricted access to the program). Consistent with previous scoping reviews [28,29], the total numerical score for each captured program was converted into a percentage, with higher percentages representing higher levels of criteria fulfillment.

Data Extraction

With permission from the authors, the data extraction form used in a previous scoping review [29] was adapted for use in this study (see Multimedia Appendix 1). For each e-mental health program, data were extracted for the following categories: website characteristics (ie, origin, organizational affiliation, general accessibility and credibility), program characteristics (ie, intervention focus, design, and delivery), intervention characteristics (ie, therapeutic approach and intervention features), and empirical evidence for program efficacy within the general population as well as for stroke survivors with and without aphasia.

To determine empirical evidence for each program, the authors scrutinized each website for relevant information and searched the program's name in each of the following databases: PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane Library, and Web of Science. If these methods failed to identify published evidence, the authors contacted the program's developers to enquire about program efficacy. Data extraction was completed by the first author during July 2017.

General Evaluation Tool

The general evaluation tool was used to evaluate each e-mental health program in terms of its general features (ie, website, program and intervention characteristics) and empirical evidence (see Multimedia Appendix 2 for completed form). It was based on the program evaluation criteria used in a previous review [29]. With permission from the authors, the original evaluation tool was adapted by adding in the following questions:

1. Was a mobile app version of the program available?
2. Did the program send completion reminders?
3. Were text-entry fields present?

Aphasia-Specific Evaluation Tool

The aphasia-specific evaluation tool was developed by the authors to assess each program in terms of its communicative accessibility for people with poststroke aphasia (see Multimedia Appendix 3 for the completed form). The evaluation criteria within this tool were based on existing aphasia-friendly guidelines for printed materials and accessibility features of related products [33-35], as well as usability considerations for older people (ie, simple menu hierarchies and text that contrasts the color of the background [36]). The aphasia-specific evaluation criteria contain 7 main categories: vocabulary and syntax, screen clarity, formatting, graphics, navigation, interface design, and media type. Vocabulary and syntax were evaluated by determining the readability level of the text, according to the Flesch-Kincaid reading grade levels [37]. This was done electronically by copying and pasting all text from the first module or session of each program into a Microsoft Word document. One program (myCompass; The Black Dog Institute) did not require modules to be completed in a specific order. Therefore, readability was based on the *Tackling Unhelpful Thinking* module, as it was described to be useful for anyone with mild to moderate depression. Text written at level 5 readability or lower was considered to be appropriate for people with poststroke aphasia, as per the Stroke Association's Accessible Information Guidelines [38]. An example of text written at or below a level 5 readability level when presented as a 3-lined paragraph is as follows: *A stroke can cause aphasia. People with aphasia are still smart. People with aphasia can still solve problems*. Font size was determined by selecting text on the webpage and using the right click *inspect* function. A ruler was used to physically measure the amount of white space between lines of text as presented on a 15.6-inch (40-cm) laptop screen at 100% zoom. Spacing of 4 mm or more was considered adequate for people with aphasia, as it is equivalent to the amount of white space measured between the lines of Times New Roman typeface set at 1.5 line spacing. The presence or absence of aphasia-friendly design characteristics, such as the use of bullet points and numbering to establish key points, the use of headings to make important information stand out, and the use of bolding to highlight important information, was also determined.

Accessibility Test

Participants

Participants were recruited through a speech pathology intervention clinic operating at The University of Queensland. All members attending a weekly aphasia group were invited to participate in the study. People with aphasia who were aged 18 years and older, had sufficient knowledge of English language to participate without a translator, and had adequate vision for reading were invited to participate. The exclusion criteria were as follows: presence of a concomitant progressive neurological

condition (eg, dementia) or a concurrent medical condition impacting their mental health (eg, cancer) as confirmed by self-report.

Materials

Computer Use Survey

A computer use survey was used to determine participants' level of computer use before and after aphasia, as well as their reasons for using a computer. This survey was created for a separate study that explored computer use by people with aphasia [39]. With permission from its publishers, this survey was adapted by asking participants if they currently or previously used a computer for the treatment of mental health difficulties and which programs they used. Furthermore, questions about computerized speech therapy programs and participants' likes and dislikes about using a computer in general were removed.

Observation Form

An observation form was developed to rate each participant's level of independence using the selected e-mental health program. Independence was rated for 6 main categories: logging in to a premade account, navigating the program, reading and understanding text, completing interactive activities, completing exercises, and finishing the session. The checklist used a 5-point scale where 1=not at all independent, 2=minimally independent, 3=moderately independent, 4=mostly independent, and 5=totally independent.

Satisfaction Survey

A satisfaction survey was developed to evaluate participants' satisfaction with using the e-mental health program, including its accessibility and ease of use. The survey consisted of 16 questions and statements for which participants answered using a 5-point rating scale (1=no, definitely not; 2=no, I don't think so; 3=neutral; 4=yes, I think so; 5=yes, definitely so). There were also opportunities for participants and their family members to add comments about the program.

Procedures

Ethical approval was granted by The University of Queensland Human Research Ethics Committee. All participants provided written informed consent before participation. Participants completed the computer use survey with assistance from their family member, carer, or research assistant, as needed. Individually, each participant then completed 1 module of the selected e-mental health program on a desktop computer. The research assistant observed each participant trialing the e-mental health program while completing the observation form. The research assistant also provided support to the participants by reading aloud informative text, reading aloud instructions, showing participants where to click/type, re-explaining instructions, controlling the mouse, and/or typing for the participants, as required. After trialing the e-mental health program, each participant completed the satisfaction survey with assistance from their family member, carer or research assistant, as needed. Each trial session was video-recorded and rewatched by the research assistant to confirm observations that had been made.

Results

Scoping Review

Program Selection

In total, 43 websites with e-mental health programs, 30 websites linking to websites with e-mental health programs, and 377 irrelevant websites were identified. After duplicate and irrelevant websites were removed, 41 websites remained, which yielded 44 individual programs. Of those programs, 8 programs met the inclusion and exclusion criteria. Figure 1 depicts the flow diagram for the program selection process, including reasons for exclusion and the final 8 programs reviewed. The authors were not granted full access to one of the included programs (the Wellbeing Course; MindSpot), and therefore evaluated this program using a "demo" version of the course that is available for practitioners.

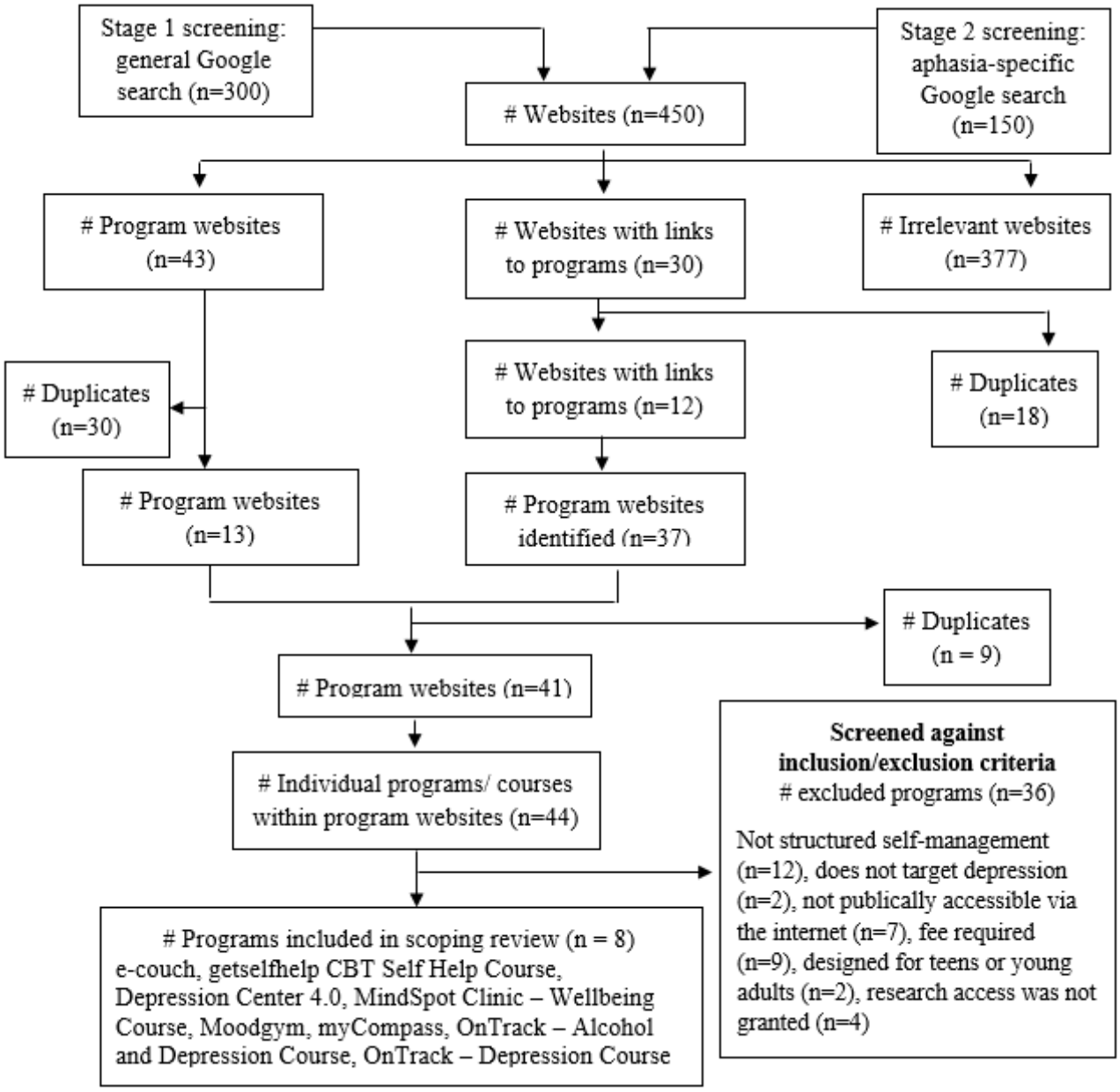
Data Extraction

An overview of the programs' website characteristics, program characteristics, and intervention characteristics can be found in Multimedia Appendices 4-6.

Two programs purely targeted depressive symptoms (OnTrack—Depression; Queensland University of Technology, and Depression Center 4.0; Evolution Health). The program OnTrack—Alcohol and Depression (OnTrack—AD; Queensland University of Technology) focused on depression with comorbid alcohol problems. The remaining 5 programs were reported to be designed for people with either depression with or without other conditions, including anxiety, anger, worry, stress and low mood, as well as emotional difficulties related to divorce, separation, bereavement, and loss. One program (myCompass; Black Dog Institute) specified user suitability criteria, while the remaining programs gave general information about whom the program would be appropriate for. All programs specified that they were designed for adult users, with 2 programs defining a specific age range for users; one being between 26 and 64 years (the Wellbeing Course) and the other being between 18 and 75 years (myCompass). In its program suitability criteria, myCompass specified that users should be able to "read English with ease." Furthermore, 3 programs had the option to send completion reminders to users via email or text message (myCompass, OnTrack—AD, and OnTrack—Depression).

All but 2 programs provided unguided therapy. One of the guided programs provided users with free support from a trained therapist via telephone or email (the Wellbeing Course). The user could choose to receive this support weekly or could choose to contact the therapist whenever he or she wished. It was not clear whether the therapist support was optional or whether the user had to contact the therapist at some point, and an attempt to contact the program's developers was unsuccessful in obtaining this information. The other guided program (Depression Center 4.0) included a *Questions to the Expert* section, where users could submit questions to be answered by a clinical psychologist. This program also contained forums that were moderated by health educators.

Figure 1. Flow diagram of program selection.



All programs claimed to employ CBT, an approach that helps people to identify and challenge negative thought patterns and learn practical self-help strategies [9]. In addition, 3 programs also used interpersonal therapy (e-couch by e-hub Health; Moodgym by e-hub Health; and myCompass), a talk-based, time-limited treatment approach that targets symptom resolution, increased social support, and improved interpersonal functioning [40]. Furthermore, 2 programs employed problem-solving therapy (e-couch and myCompass), a treatment approach that focuses on constructive problem-solving skills and attitudes to help individuals cope more effectively with life's stressors [41]. Positive psychology, which focuses on one's positive experiences and qualities rather than on situations and qualities that cause suffering, was also used by 1 program (myCompass) [42], whereas 2 programs (OnTrack—AD and OnTrack—Depression) included meditational mindfulness, an approach that uses aspects of meditation to facilitate self-regulation of attention and help users adopt an accepting outlook on their experiences [43].

Empirical Evidence

Empirical Evidence for Program Efficacy Within the General Population

An initial search through all the programs' websites for research evidence identified 2 published articles reporting on program efficacy for depression for Moodgym [44] and myCompass [45]. A database search in PubMed, CINAHL, Cochrane Library, and Web of Science yielded 9 more published research articles investigating e-couch [46-48] and the Wellbeing Course [49-54]. The authors contacted the developers of the remaining programs to enquire about program efficacy. In response, the authors received unpublished data investigating the efficacy of 2 programs (OnTrack—AD and OnTrack—Depression; personal communication by J. Connolly, August 2017). Only the highest level of evidence for each program was reviewed according to the National Health and Medical Research Council's evidence hierarchy [55]. For example, if an RCT was identified for a

program, lower level evidence (eg, a case-control study) for that same program was not reviewed. There was, however, an exception to this, whereby an RCT evaluating Moodgym was reviewed despite there being a meta-analysis available for this program, as the RCT included participants with stroke [56]. Table 1 outlines the highest level of published evidence for each program.

Empirical Evidence for Program Efficacy After Stroke and/or for People With Poststroke Aphasia

No studies were found that evaluated efficacy of use specifically for stroke survivors with or without aphasia. Due to a technical error in the research portal, 1 RCT included in Moodgym's meta-analysis unintendedly recruited a number of participants who were later found to have a brain injury or stroke or who were already receiving CBT treatment for bipolar disorder [56]. This RCT used an attention-control (e-mental health literacy package) and reported improvements in depression in both the intervention and attention-controlled groups but no significant differences between groups [56]. It is unknown how many stroke survivors participated, as they could not be differentiated from those with brain injury or those receiving CBT for bipolar disorder (personal communication by J. Schneider, September 2017). One of the RCTs for e-couch included participants with a previous doctor's diagnosis of heart disease, stroke, or hypertension [46]. This RCT reported improvements in depressive symptoms after completing the e-couch program [46]. A total of 15 out of 562 participants were stroke survivors, but there were no data regarding the inclusion of people with aphasia (personal communication by N Glozier, September 2017). These were the only studies that included participants with known long-term health conditions. Neither of these studies completed subgroup analysis.

Results From the General Program and Aphasia-Specific Communicative Accessibility Evaluation Tools

The overall scores and corresponding percentages from the general evaluation and aphasia-specific evaluations are outlined in Table 2. General program evaluation scores for each criterion within the general evaluation tool and aphasia-specific communicative accessibility evaluation tool are available in Multimedia Appendices 2 and 3, respectively. General program evaluation scores ranged from 26% (5/19; getselfhelp; Carol Vivyan) to 84% (16/19; OnTrack—AD, Moodgym, and myCompass; mean: 71% [SD 18.1]). Aphasia-specific communicative accessibility evaluation scores ranged from 55% (11/20; getselfhelp) to 85% (17/20; OnTrack—AD; mean: 72% [SD 9.3]). As Moodgym was the only evidence-based program to score at least 80% on both the general evaluation and aphasia-specific communicative accessibility evaluation tools, it was selected to be trialed by people with poststroke aphasia in the next stage of the study.

Results of Accessibility Test

Participant Demographics

A total of 3 participants with a diagnosis of poststroke aphasia (2 females, 1 male) were recruited. Participants' demographics, including what and how often they used a computer for pre- and poststroke, are included in Table 3.

Observation Form

The level of independence each participant demonstrated while completing the first module of Moodgym, as observed by the research assistant using the observation form, is available in Table 4.

Satisfaction Survey

Participants' satisfaction ratings of Moodgym, as well as their comments about the program, are provided in Table 5.

Table 1. Summary of highest level of published evidence for each program.

Program and level of evidence	Control group	Sample size	Findings	Supports use in general population	Included people with stroke
Moodgym (e-hub health)					
I [44] (meta-analysis)	N/A[a]	11 studies	Small effect size for improving symptoms of depression. Nonsignificant effect size when adjusted for potential publication bias.	✓[b]	N/A
II [56] (RCT[c]—included in above meta-analysis)	AC[d](online mental health information package)	340 and 219 participants completed the PHQ-9[d] at 6 and 12 weeks follow-up.	No significant difference between Moodgym and AC in terms of psychological outcomes or service use, although improvement to subthreshold levels of depression seen in nearly half the participants in both groups at 6-week follow-up.	X[e]	✓
E-couch (e-hub health)					
II [46] (RCT)	AC (online health information package)	487 participants completed the postintervention assessment.	Small but robust improvement in depression symptoms in treatment group relative to AC post intervention.	✓	✓
II [47] (RCT)	Control website with delayed access to e-couch	209 and 176 participants completed the 6 and 12-month follow-up assessment.	E-couch was effective relative to control at post intervention but not at 6-month follow-up.	✓	X
II [48] (RCT)	Moodgym	549 and 336 participants completed the postintervention and follow-up assessments.	Significant reduction in depression symptoms at post intervention and 6-month follow-up for both e-couch's CBT[f] and IPT[g] modules and both were noninferior to Moodgym.	✓	X
myCompass (Black Dog Institute)					
II [45] (RCT)	AC and WLC[h]	449 and 350 participants completed the postintervention and 3-month follow-up.	Reduction in depression symptoms relative to both control conditions post intervention. Participants in AC group showed gradual reductions in depression symptoms during postintervention stage and scores did not differ from the myCompass group at follow-up.	✓	X
The Wellbeing Course (MindSpot)					
II [49] (RCT)	Mood Course[i]	229 participants completed the 24-month follow-up assessment.	Consistent reductions in MDD[j] symptoms across conditions post intervention and 24-month follow-up.	✓	X
II [50] (RCT)	Wellbeing course with and without automated emails compared with WLC[k]	219 and 199 participants completed the postintervention and 3-month follow-up assessments.	Reductions in symptoms of anxiety and depression relative to WLC at post intervention and 3-month follow-up.	✓	X
II [51] (RCT)	WLC	77 participants in total, 38 of which had depression.	Reduced depression symptoms post intervention and maintained at 3-month follow-up.	✓	X

Program and level of evidence	Control group	Sample size	Findings	Supports use in general population	Included people with stroke
II [52] (RCT)	Social Confidence Course[i]	172 and 170 participants completed the postintervention and 24-month follow-up assessment. 87 of these participants had depression symptoms	Consistent reduction in comorbid depression[l] symptoms across conditions postintervention and at 24-month follow-up.	✓	X
II [54] (RCT)	The Panic Course[i]	122 and 111 participants completed the postintervention and 24-month follow-up assessment. 38 of these participants met the diagnostic criteria for MDD.	Consistent reduction in comorbid depression[l] symptoms across conditions over 24-month follow-up.	✓	X
II [53] (RCT)	The Worry Course[i]	282 and 260 pts completed the postintervention and 24-month follow-up assessment. 157 participants had depression symptoms	Consistent reduction in comorbid depression[l] symptoms across conditions post intervention and at 3-month follow-up. Treatment group's depression symptoms slightly improved relative to AC from 3- to 12-month follow-up.	✓	X
OnTrack—Alcohol and Depression (Queensland University of Technology)[m]	N/A	N/A	N/A	N/A	N/A
OnTrack—Depression (Queensland University of Technology)[m]	N/A	N/A	N/A	N/A	N/A
Depression Center 4.0 (Evolution Health)[m]	N/A	N/A	N/A	N/A	N/A
Getselfhelp (Carol Vivyan)[m]	N/A	N/A	N/A	N/A	N/A

[a]N/A: not applicable.
[b]✓: yes.
[c]RCT: randomized controlled trial.
[d]AC: attention-control.
[e]Patient Health Questionnaire-9.
[f]X: no.
[g]CBT: cognitive behavioral therapy.
[h]IPT: interpersonal therapy.
[i]Developed specifically for the study.
[j]MDD: major depressive disorder.
[k]WLC: waitlist-control.
[l]Depression as a secondary outcome.
[m]No published evidence.

Table 2. Program evaluation scores and percentages.

Program	General evaluation score (N=19), n (%)[a]	Aphasia-specific evaluation score (N=20), n (%)[a]
OnTrack—Alcohol and Depression	16 (84)	17 (85)
Moodgym	16 (84)	16 (80)
myCompass	16 (84)	13 (65)
OnTrack—Depression	15 (79)	16 (80)
Wellbeing Course (Demo version)	14 (74)	15 (75)
e-couch	13 (68)	15 (75)
Depression Center 4.0	13 (68)	13 (65)
Getselfhelp	5 (26)	11 (55)

[a]Percentages were rounded up/down to the nearest whole number.

Table 3. Participant demographic data.

Demographic characteristics	Participant 1	Participant 2	Participant 3
Age in years	77	57	46
Time since stroke	3 years, 2 months	6 years, 4 months	6 years, 4 months
Severity of aphasia	Mild	Moderate	Severe
Highest level of education	University	University and other	University
Previous computer use/current computer use for:			
Work	Weekly/N/A	Daily/N/A	Daily/N/A
Writing letters	Fortnightly/Fortnightly	Weekly/Never	Daily/Never
Household budgeting/financing	Weekly/Weekly	Weekly/Never	Daily/Never
Photograph management	Never/Weekly	Monthly/Rarely	Daily/Monthly
Home movie creation	Never/ Never	Never/Never	Monthly/Never
PowerPoint creation	Never/Never	Weekly/Never	Weekly/Never
Banking	Weekly/Weekly	Daily/Never	Daily/Never
Email	Daily/Daily	Daily/Monthly	Never/Never
Social media	Never/Weekly	Weekly/Fortnightly	Daily/Daily
Skype	Never/Never	Monthly/Rarely	Never/Never
General interest/web searching	Never/Weekly	Daily/Daily	Daily/Weekly
Shopping	Never/Weekly	Monthly/Never	Daily/Never
Entertainment	Monthly/Fortnightly	Weekly/Weekly	Daily/Daily
Therapy—speech, language	Never/Daily	Never/Weekly	Never/Never
Therapy—for MH difficulties	Never/Never	Never/Never	Never/Never
Other	Nil/Nil	Nil/Nil	Nil/Nil
Type of computer/s currently used	Desktop, tablet, smartphone	Tablet	Tablet, smartphone
Needs help using a computer for:	Setting up, getting into programs, using the computer, turning computer off	Setting up	Setting up, getting into programs, using the computer, turning computer off

[a]N/A: not applicable.

Table 4. Participants' levels of independence as assessed via the observation tool.

Task and rating	Participant 1	Participant 2	Participant 3
Log into premade account			
Enter log-in details	Minimally independent	Minimally independent	Not at all independent
Click on the log-in tab in upper right hand corner	Totally independent	Totally independent	Not at all independent
Navigate program			
Access the "Feeling Module"	Mostly independent	Mostly independent	Not at all independent
Use scroll bar/arrows to view all text on the page	Mostly independent	Totally independent	Not at all independent
Use side arrows to click onto next page	Totally independent	Mostly independent	Not at all independent
Read and understand text			
Read informative text	Mostly independent	Not at all independent	Not at all independent
Read and correctly follow instructions	Minimally independent	Not at all independent	Not at all independent
Complete interactive activities			
Click on the image /tab/link to access indicated information	Minimally independent	Minimally independent	Not at all independent
Complete exercises			
Select yes/no answers during tasks/quizzes	Mostly independent	Minimally independent	Not at all independent
Click on 'submit' to submit answers	Totally independent	Minimally Independent	Not at all independent
Answer open-ended questions via text-entry field	Totally independent	Not at all independent	Not at all independent
Finish session			
Log out of Moodgym	Minimally independent	Minimally independent	Not at all independent
Exit out of Moodgym	Minimally independent	Totally independent	Minimally independent
Research assistant's comments	Independently read informative text, but reading was slow and effortful.	All text read aloud by research assistant; difficulties completing yes/ no quizzes	All text read aloud by research assistant; mouse controlled by research assistant

Table 5. Results of satisfaction survey.

Question/statement in satisfaction survey	Participant 1	Participant 2	Participant 3
1. Was it easy to login?	No—I don't think so	Yes—I think so	Yes—I think so
2. Did Moodgym look appealing?	No—I don't think so	Neutral	Neutral
3. Was the information worded in a way that was easy to understand?	No—I don't think so	Neutral	No—I don't think so
4. Were the instructions worded in a way that was easy to understand?	No—I don't think so	Yes—I think so	Yes—I think so
5. Were the words and pictures clear on the screen?	No—I don't think so	Yes—I think so	No—I don't think so
6. Was the text style easy to read?	Neutral	Yes—I think so	Yes—I think so
7. Was the text size easy to read?	No—I don't think so	Yes—I think so	Neutral
8. Was there enough white space on each page?	Neutral	Neutral	Yes—I think so
9. Was it easy to find important information?	Neutral	Yes—I think so	Neutral
10. Did the pictures help you to understand the information?	No—I don't think so	Yes—I think so	Yes—I think so
11. Was Moodgym simple to use?	No—definitely not	Neutral	No—I don't think so
12. Could you use Moodgym without help?	No—definitely not	No—definitely not	No—definitely not
13. Moodgym looked like it was developed for someone with aphasia to use	No—definitely not	No—I don't think so	No—I don't think so
14. Did you enjoy using Moodgym?	No—I don't think so	Yes—I think so	No—definitely not
15. Overall, were you satisfied with Moodgym?	No—I don't think so	Yes—I think so	No—definitely not
16. Overall, was Moodgym easy to use?	No—I don't think so	Neutral	No—definitely not
Comments made	"Very complex language"	"Once it was read out and explained, it was easy"	"Hard to understand"

Discussion

Principal Findings

This scoping review identified 8 e-mental health programs for depression that are freely available to Australian residents. Of these, 4 programs (Moodgym, e-couch, myCompass, and the Wellbeing Course) have been shown to reduce symptoms of depression within the general population. No empirical evidence was identified that specifically supported any programs' use for stroke survivors with or without aphasia. However, it was found that e-couch and Moodgym had been evaluated using RCT methodology that included a small number of participants with stroke. Findings from a general and an aphasia-specific evaluation indicated that Moodgym was the only evidence-based program to score at least 80% (16/19 and 16/20) on both evaluations, suggesting that it may be suitable for people with poststroke aphasia. However, when trialing Moodgym, 2 of the 3 participants with aphasia required assistance for more than half of the skills assessed on the observation form, and the other participant who demonstrated higher levels of independence still gave Moodgym low satisfaction and usability scores on the posttrial satisfaction survey. Therefore, despite fulfilling majority of the general evaluation and aphasia-specific evaluation criteria, Moodgym was still found to be unsuitable for people with poststroke aphasia. This also suggests that the aphasia-specific criteria used to initially assess communicative accessibility may not have been sensitive enough to detect suitability for people with poststroke aphasia and thus should be revised if it is to be used in any future research.

Consistent with previous findings in the literature, this study found that the programs' content was usually generic for all users and that the programs themselves tended to be based on CBT, either alone or in combination with other approaches [28,29]. This may be explained by the fact that CBT is the most researched psychotherapy and one of the first forms of psychotherapy to be established as evidence-based [57]. Furthermore, because of its structured protocols and ability to be manualized, CBT lends itself well to self-help e-mental health platforms [58]. This study also builds on the findings of the previous scoping review on e-mental health programs for depression [28]. First, only a few programs were found to send out completion reminders via email and/or text message. This strategy has been found to increase adherence to medical treatment [59] and may be especially useful for e-mental health programs that tend to have high dropout rates [60]. Second, no program was found to have a companion app able to be purchased at an official Australian Apple or Android app store at the time of this review. Previous research investigating mobile computing technology and aphasia [33] suggests that an advantage of mobile computer apps is their ability to be adapted for people with disabilities. For example, buttons on touch screens can be customized, unlike the physical mouse and keyboard of desktop computers, and apps and touch interfaces can be changed to suit users' vision and mobility needs, as there is no set button size [33]. This may be particularly relevant to stroke survivors, who often face upper extremity motor impairment post stroke [61]. Furthermore, mobile computer devices such as the Apple iPad may offer more accessibility options such as predictive text, switch control, and VoiceOver.

While this review did not identify any new e-mental health programs for depression released since the previous scoping review [28], it did identify new empirical evidence for some of the programs. Namely, this study identified a meta-analysis supporting the efficacy of Moodgym [44] and RCTs supporting the use of myCompass [45] and the Wellbeing Course [49,52-54] by the general population. Unfortunately, this study found that there is currently a lack of high-quality research investigating the efficacy of e-mental health programs for stroke survivors with and without aphasia. Furthermore, the studies that did include stroke participants failed to specify the presence or absence of people with poststroke aphasia within the included stroke samples. This supports recent findings that people with poststroke aphasia are often excluded or inadequately reported on as a subsample within stroke populations in mental health research [62]. As suggested by Baker and colleagues [62], it is important that mental health studies include a minimum dataset for people with poststroke aphasia included in a study (ie, report number of people with poststroke aphasia and severity and nature of communication difficulties) and conduct subgroup analyses. Failure to do so will hinder the progression of research for people with poststroke aphasia in this area. This is of particular importance as the prevalence of psychological conditions among people with stroke (with and without aphasia) is likely to continue to increase alongside an aging population [63,64] and accompanied increased incidence of stroke [65].

People with poststroke aphasia have previously been identified as being victims of *digital exclusion* as a result of their language deficits, age- and stroke-related changes, and lack of premorbid computer and internet skills often attributed to their generally older age [33,66]. However, consistent with previous findings [39], the results of the computer use survey indicated a high level of computer usage by the participants with aphasia before and after their stroke. This aligns with other research, which suggests that tomorrow's elders with disabilities will generally have more access to and increased proficiency with wireless technologies than their predecessors [67]. Thus, it can be assumed that as technology continues to advance to meet the growing needs of consumers, and as the number of older adults with digital literacy increases, it is likely that less people with poststroke aphasia will face digital exclusion than in previous years [67].

The communicative accessibility of the reviewed programs for people with poststroke aphasia was initially evaluated against an aphasia-specific evaluation tool. Aphasia-specific evaluation scores ranged from 55% (11/20) to 85% (17/20), with only 3 programs scoring 80% or more (Moodgym, OnTrack—AD, and OnTrack—Depression). Hence, these findings suggest that while all programs incorporated design features consistent with published aphasia-friendly guidelines and other recommendations, the communicative accessibility of e-mental health programs could potentially be improved to render these programs more appropriate to people with poststroke aphasia. Findings indicated that many programs might be improved by greater use of visuographic supports. For instance, pictures, which directly support the text, have been found to enhance reading comprehension in people with aphasia [68,69]. For example, it was noted that Moodgym's "yes" or "no" checkbox quizzes, which did not contain pictorial support, were particularly difficult for 1 participant with poststroke aphasia to complete. Unreliable "yes" or "no" responses is a common symptom of aphasia [70]. Therefore, such exercises could be made more aphasia-friendly by including pictorial supports such as a picture of a tick next to the "yes" option and a picture of a cross next to the "no" option for each item. It should also be noted that poststroke aphasia can differentially affect one's language modalities (speaking, understanding, reading, and writing, etc) [71], resulting in people with poststroke aphasia having different language "profiles," which may make it easier or harder for them to access e-mental health support. Therefore, incorporating a wider variety of media types, including videos, animation, graphics, and audio, would likely increase the accessibility of the programs to meet the differing communicative needs of people with poststroke aphasia.

As assessed via the aphasia-specific communicative accessibility evaluation tool, this study found that none of the evaluated programs had an average readability level of 5 or lower, which is the level recommended for people with aphasia [38]. As reported by the participants with aphasia, the language used in Moodgym, which had a readability of 7.9, was complex and hard to understand, and 2 participants required all text to be read aloud to them. Furthermore, prior literature has found that reliable Web-based health information is often presented at reading levels too high for the general population as it is [72-76]. Therefore, reducing the text readability level of e-mental health programs should increase their accessibility not only to people with poststroke aphasia but also to the general population as a whole.

While appropriate readability levels of written materials are important, there are other formatting characteristics that can facilitate reading comprehension in people with poststroke aphasia. These include use of large font (ie, 14-point or larger) [77], generous spacing between lines to maximize white space [77], the use of headings and bolding to make important information stand out [34], and the use of bullet points and numbers to clearly establish key points [34]. All reviewed programs were found to use headings, bolding, bullet points and numbering, although to differing degrees. However, some programs could enhance their accessibility by using more white space between lines of text (ie, 1.5 spacing) to make content appear more appealing and readable [77] and by using bolding to highlight key information [34].

Limitations and Direction for Further Research

Despite the authors' efforts to be systematic and inclusive in their search for e-mental health programs for depression, it is possible that some programs were missed. The dynamic nature of Web content also means that the identified programs may eventually be changed or discontinued, and new programs will likely be developed in the future. Furthermore, this study excluded paid programs (n=9) and programs for which the authors were not granted research access (n=4). Inclusion of these programs may have yielded different results and thus resulted in different recommendations.

The aphasia-specific search terms used in the search strategy were informed by suggestions from clinicians and academics

that work with people with poststroke aphasia, rather than people with poststroke aphasia themselves. Future studies investigating how people with poststroke aphasia use the internet to find e-mental health programs, and if their search terms are successful in locating the same programs identified in this review, may also be of benefit.

This small trial of people with poststroke aphasia was helpful in determining whether the highest rated e-mental health program was communicatively accessible to people with poststroke aphasia, or not. However, a larger trial of people with poststroke aphasia is paramount in determining which aspects of current e-mental health programs are most accessible, and which aspects need improvement. The use of think-aloud studies, whereby participants' experiences of the program are captured using multimodal communication, may be one way to identify accessibility facilitators and barriers of such e-mental health programs. Furthermore, if e-mental health programs are to be redesigned or future programs developed specifically for people with poststroke aphasia, end users (ie, people with aphasia) should participate in the design and development process to ensure usability within the target population. It is acknowledged that participatory design approaches rely heavily on effective communication between the participants and designers and thus may be challenging for people with poststroke aphasia [78]. However, there are numerous examples of people with poststroke aphasia successfully partaking in participatory design studies to enhance the usability of aphasia-specific programs [19,78-82].

Conclusions

E-mental health services offer convenient and cost-effective interventions that have an ability to reach a more diverse population than traditional face-to-face psychological interventions [21]. Thus, the next decade will likely see mental health services progress toward a digital medium, which will present numerous opportunities for both clinical practice and e-mental health research [83]. It is important that people with poststroke aphasia, a population with an increased risk of developing depression, are considered in future e-mental health research. Failure to do so may mean that people with poststroke aphasia and mild to moderate depression may not be offered e-mental health treatment, or the options available to them may remain inaccessible, and therefore potentially nonbeneficial to them.

Acknowledgments

We would like to acknowledge the program developers and study authors who provided us with additional information and details about their programs/studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data extraction form.
[PDF File (Adobe PDF File), 35 KB - jmir_v20i12e291_app1.pdf]

Multimedia Appendix 2
Completed general evaluation form.
[PDF File (Adobe PDF File), 44 KB - jmir_v20i12e291_app2.pdf]

Multimedia Appendix 3
Completed aphasia-specific evaluation form.
[PDF File (Adobe PDF File), 48 KB - jmir_v20i12e291_app3.pdf]

Multimedia Appendix 4
Website characteristics of included programs.
[PDF File (Adobe PDF File), 35 KB - jmir_v20i12e291_app4.pdf]

Multimedia Appendix 5
Program characteristics of included programs.
[PDF File (Adobe PDF File), 32 KB - jmir_v20i12e291_app5.pdf]

Multimedia Appendix 6
Intervention characteristics of included programs.
[PDF File (Adobe PDF File), 203 KB - jmir_v20i12e291_app6.pdf]

References

1. Cruice M, Worrall L, Hickson L. Health-related quality of life in people with aphasia: implications for fluency disorders quality of life research. J Fluency Disord 2010 Sep;35(3):173-189. [doi: 10.1016/j.jfludis.2010.05.008] [Medline: 20831966]
2. Flowers HL, Silver FL, Fang J, Rochon E, Martino R. The incidence, co-occurrence, and predictors of dysphagia, dysarthria, and aphasia after first-ever acute ischemic stroke. J Commun Disord 2013;46(3):238-248. [doi: 10.1016/j.jcomdis.2013.04.001] [Medline: 23642855]
3. Kauhanen ML, Korpelainen JT, Hiltunen P, Määttä R, Mononen H, Brusin E, et al. Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. Cerebrovasc Dis 2000;10(6):455-461. [doi: 10.1159/000016107] [Medline: 11070376]
4. Dorze GL, Brassard C. A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. Aphasiology 1995 May;9(3):239-255. [doi: 10.1080/02687039508248198]
5. Hilari K. The impact of stroke: are people with aphasia different to those without? Disabil Rehabil 2011;33(3):211-218. [doi: 10.3109/09638288.2010.508829] [Medline: 20712416]
6. Parr S. Living with severe aphasia: tracking social exclusion. Aphasiology 2007 Jan;21(1):98-123. [doi: 10.1080/02687030600798337]
7. Aström M, Adolfsson R, Asplund K. Major depression in stroke patients. A 3-year longitudinal study. Stroke 1993 Jul;24(7):976-982. [Medline: 8322398]
8. Ayerbe L, Ayis S, Wolfe CD, Rudd AG. Natural history, predictors and outcomes of depression after stroke: systematic review and meta-analysis. Br J Psychiatry 2013 Jan;202(1):14-21 [FREE Full text] [doi: 10.1192/bjp.bp.111.107664] [Medline: 23284148]
9. National Collaborating Centre for Mental Health. Depression: The Treatment and Management of Depression in Adults (Updated Edition). United Kingdom: Royal College of Psychiatrists; 2010.
10. Malhi GS, Bassett D, Boyce P, Bryant R, Fitzgerald PB, Fritz K, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry 2015 Dec;49(12):1087-1206. [doi: 10.1177/0004867415617657] [Medline: 26643054]
11. Australian Government, Department of Health. 2015. Australian Government response to contributing lives, thriving communities: review of mental health programmes and services URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/0DBEF2D78F7CB9E7CA257F07001ACC6D/$File/response.pdf [WebCite Cache ID 6uAZ1Gktm]
12. Hickie IB, McGorry PD. Increased access to evidence-based primary mental health care: will the implementation match the rhetoric? Med J Aust 2007 Jul 16;187(2):100-103. [Medline: 17635093]
13. Orman J, O'Dea B, Shand F, Berk M, Proudfoot J, Christensen H. e-Mental health for mood and anxiety disorders in general practice. Aust Fam Physician 2014 Dec;43(12):832-837. [Medline: 25705730]
14. Schierenbeck I, Johansson P, Andersson L, van Rooyen D. Barriers to accessing and receiving mental health care in Eastern Cape, South Africa. Health Hum Rights 2013 Dec 12;15(2):110-123. [Medline: 24421159]
15. Kazdin AE, Rabbitt SM. Novel models for delivering mental health services and reducing the burdens of mental illness. Clin Psychol Sci 2013 Jan 23;1(2):170-191. [doi: 10.1177/2167702612463566]
16. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
17. Northcott S, Simpson A, Moss B, Ahmed N, Hilari K. How do speech-and-language therapists address the psychosocial well-being of people with aphasia? Results of a UK online survey. Int J Lang Commun Disord 2017 Dec;52(3):356-373. [doi: 10.1111/1460-6984.12278] [Medline: 27593505]
18. Zheng C, Lynch L, Taylor N. Effect of computer therapy in aphasia: a systematic review. Aphasiology 2015 Jan 03;30:1-34. [doi: 10.1080/02687038.2014.996521]
19. Marshall J, Booth T, Devane N, Galliers J, Greenwood H, Hilari K, et al. Evaluating the benefits of aphasia intervention delivered in virtual reality: results of a quasi-randomised study. PLoS One 2016;11(8):e0160381 [FREE Full text] [doi: 10.1371/journal.pone.0160381] [Medline: 27518188]
20. Fotiadou D, Northcott S, Chatzidaki A, Hilari K. Aphasia blog talk: how does stroke and aphasia affect a person's social relationships? Aphasiology 2014 Jun 19;28(11):1281-1300. [doi: 10.1080/02687038.2014.928664]
21. Australian Government, Department of Health and Aging. 2012. E-Mental Health Strategy for Australia Report URL: https://www.health.gov.au/internet/main/publishing.nsf/content/7C7B0BFEB985D0EBCA257BF0001BB0A6/$File/emstrat.pdf [WebCite Cache ID 710zySwQJ]
22. Beyond Blue. 2013. Improved access - e-mental health programs URL: https://www.beyondblue.org.au/docs/default-source/policy-submissions/bw0161.pdf?sfvrsn=2 [WebCite Cache ID 6uAZqn6zN]
23. Levesque J, Harris MF, Russell G. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. Int J Equity Health 2013 Mar 11;12:18 [FREE Full text] [doi: 10.1186/1475-9276-12-18] [Medline: 23496984]
24. Danielsson A, Willén C, Sunnerhagen KS. Physical activity, ambulation, and motor impairment late after stroke. Stroke Res Treat 2012;2012:818513 [FREE Full text] [doi: 10.1155/2012/818513] [Medline: 21941689]

25. Topolovec-Vranic J, Cullen N, Michalak A, Ouchterlony D, Bhalerao S, Masanic C, et al. Evaluation of an online cognitive behavioural therapy program by patients with traumatic brain injury and depression. Brain Inj 2010;24(5):762-772. [doi: 10.3109/02699051003709599] [Medline: 20370383]
26. American Speech–Language–Hearing Association. Traumatic Brain Injury (TBI) URL: http://www.asha.org/public/speech/disorders/TBI/ [WebCite Cache ID 6uAccY8vO]
27. Simblett SK, Yates M, Wagner AP, Watson P, Gracey F, Ring H, et al. Computerized cognitive behavioral therapy to treat emotional distress after stroke: a feasibility randomized controlled trial. JMIR Ment Health 2017 May 31;4(2):e16 [FREE Full text] [doi: 10.2196/mental.6022] [Medline: 28566265]
28. Renton T, Tang H, Ennis N, Cusimano MD, Bhalerao S, Schweizer TA, et al. Web-based intervention programs for depression: a scoping review and evaluation. J Med Internet Res 2014;16(9):e209 [FREE Full text] [doi: 10.2196/jmir.3147] [Medline: 25249003]
29. Ashford MT, Olander EK, Ayers S. Finding web-based anxiety interventions on the world wide web: a scoping review. JMIR Ment Health 2016 Jun 01;3(2):e14 [FREE Full text] [doi: 10.2196/mental.5349] [Medline: 27251763]
30. De Choudhury M, Morris M, White R. Seeking and sharing health information online: Comparing search engines and social media. 2014 Presented at: Proceedings of the SIGCHI Conference on human factors in computing systems; 2014; Toronto, Canada.
31. Pang PC, Verspoor K, Chang S, Pearce J. Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study. Health Technol 2015 Feb 10;5(1):45-55. [doi: 10.1007/s12553-015-0096-0]
32. Siu E. Search Engine Journal. 2012. 24 Eye-Popping SEO Statistics URL: https://www.searchenginejournal.com/24-eye-popping-seo-statistics/42665/ [accessed 2017-10-13] [WebCite Cache ID 6uAcyDLJ5]
33. Brandenburg C, Worrall L, Rodriguez AD, Copland D. Mobile computing technology and aphasia: an integrated review of accessibility and potential uses. Aphasiology 2013 Apr;27(4):444-461. [doi: 10.1080/02687038.2013.772293]
34. Egan J, Worrall L, Oxenham D. Accessible internet training package helps people with aphasia cross the digital divide. Aphasiology 2004 Mar;18(3):265-280. [doi: 10.1080/02687030344000562]
35. Shen N, Levitan M, Johnson A, Bender JL, Hamilton-Page M, Jadad AA, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
36. Gao Q, Ebert D, Chen X, Ding Y. Design of a mobile social community platform for older Chinese people in urban areas. Hum Factors Man 2012 Jun 27;25(1):66-89. [doi: 10.1002/hfm.20523]
37. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [Medline: 18867058]
38. Stroke Association. 2012. Accessible Information Guidelines URL: https://www.stroke.org.uk/sites/default/files/accessible_information_guidelines.pdf1_.pdf [WebCite Cache ID 6uAd5KPXw]
39. Finch E, Hill AJ. Computer use by people with aphasia: a survey investigation. Brain Impair 2014 Sep 18;15(02):107-119. [doi: 10.1017/BrImp.2014.17]
40. Markowitz JC, Weissman MM. Interpersonal psychotherapy: principles and applications. World Psychiatry 2004 Oct;3(3):136-139 [FREE Full text] [Medline: 16633477]
41. Dobson K. Handbook of cognitive-behavioral therapies. 3rd edition. New York: Guilford Press; 2010.
42. Seligman ME, Csikszentmihalyi M. Positive psychology: an introduction. Am Psychol 2000;55(1):5-14. [doi: 10.1037/0003-066X.55.1.5]
43. Bishop S, Lau M, Shapiro S. Mindfulness: a proposed operational definition. Clin Psychol-Sci Pr 2006;11(3):230-241. [doi: 10.1093/clipsy.bph077]
44. Twomey C, O'Reilly G. Effectiveness of a freely available computerised cognitive behavioural therapy programme (MoodGYM) for depression: Meta-analysis. Aust N Z J Psychiatry 2017 Mar;51(3):260-269. [doi: 10.1177/0004867416656258] [Medline: 27384752]
45. Proudfoot J, Clarke J, Birch M, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
46. Glozier N, Christensen H, Naismith S, Cockayne N, Donkin L, Neal B, et al. Internet-delivered cognitive behavioural therapy for adults with mild to moderate depression and high cardiovascular disease risks: a randomised attention-controlled trial. PLoS One 2013;8(3):e59139 [FREE Full text] [doi: 10.1371/journal.pone.0059139] [Medline: 23555624]
47. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. PLoS One 2012;7(12):e53244 [FREE Full text] [doi: 10.1371/journal.pone.0053244] [Medline: 23285271]
48. Donker T, Bennett K, Bennett A, Mackinnon A, van Straten A, Cuijpers P, et al. Internet-delivered interpersonal psychotherapy versus internet-delivered cognitive behavioral therapy for adults with depressive symptoms: randomized controlled noninferiority trial. J Med Internet Res 2013;15(5):e82 [FREE Full text] [doi: 10.2196/jmir.2307] [Medline: 23669884]
49. Titov N, Dear BF, Staples LG, Terides MD, Karin E, Sheehan J, et al. Disorder-specific versus transdiagnostic and clinician-guided versus self-guided treatment for major depressive disorder and comorbid anxiety disorders: a randomized

controlled trial. J Anxiety Disord 2015 Oct;35:88-102 [FREE Full text] [doi: 10.1016/j.janxdis.2015.08.002] [Medline: 26422822]
50. Titov N, Dear BF, Johnston L, Lorian C, Zou J, Wootton B, et al. Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: randomised controlled trial. PLoS One 2013;8(7):e62873 [FREE Full text] [doi: 10.1371/journal.pone.0062873] [Medline: 23843932]
51. Titov N, Dear BF, Schwencke G, Andrews G, Johnston L, Craske MG, et al. Transdiagnostic internet treatment for anxiety and depression: a randomised controlled trial. Behav Res Ther 2011 Aug;49(8):441-452. [doi: 10.1016/j.brat.2011.03.007] [Medline: 21679925]
52. Dear BF, Staples LG, Terides MD, Fogliati VJ, Sheehan J, Johnston L, et al. Transdiagnostic versus disorder-specific and clinician-guided versus self-guided internet-delivered treatment for Social Anxiety Disorder and comorbid disorders: a randomized controlled trial. J Anxiety Disord 2016 Aug;42:30-44 [FREE Full text] [doi: 10.1016/j.janxdis.2016.05.004] [Medline: 27261562]
53. Dear BF, Staples LG, Terides MD, Karin E, Zou J, Johnston L, et al. Transdiagnostic versus disorder-specific and clinician-guided versus self-guided internet-delivered treatment for generalized anxiety disorder and comorbid disorders: a randomized controlled trial. J Anxiety Disord 2015 Dec;36:63-77 [FREE Full text] [doi: 10.1016/j.janxdis.2015.09.003] [Medline: 26460536]
54. Fogliati VJ, Dear BF, Staples LG, Terides MD, Sheehan J, Johnston L, et al. Disorder-specific versus transdiagnostic and clinician-guided versus self-guided internet-delivered treatment for panic disorder and comorbid disorders: a randomized controlled trial. J Anxiety Disord 2016 Apr;39:88-102 [FREE Full text] [doi: 10.1016/j.janxdis.2016.03.005] [Medline: 27003376]
55. National Health Medical Research Council. How to review the evidence: Systematic identification and review of the scientific literature. Canberra: National Health and Medical Research Council; 2000.
56. Phillips R, Schneider J, Molosankwe I, Leese M, Foroushani PS, Grime P, et al. Randomized controlled trial of computerized cognitive behavioural therapy for depressive symptoms: effectiveness and costs of a workplace intervention. Psychol Med 2014 Mar;44(4):741-752 [FREE Full text] [doi: 10.1017/S0033291713001323] [Medline: 23795621]
57. David D, Cristea I, Hofmann SG. Why cognitive behavioral therapy is the current gold standard of psychotherapy. Front Psychiatry 2018;9:4 [FREE Full text] [doi: 10.3389/fpsyt.2018.00004] [Medline: 29434552]
58. Gaudiano BA. Cognitive-behavioural therapies: achievements and challenges. Evid Based Ment Health 2008 Feb;11(1):5-7 [FREE Full text] [doi: 10.1136/ebmh.11.1.5] [Medline: 18223042]
59. van Dulmen S, Sluijs E, van Dijk L, de Ridder D, Heerdink R, Bensing J. Patient adherence to medical treatment: a review of reviews. BMC Health Serv Res 2007 Apr 17;7:55 [FREE Full text] [doi: 10.1186/1472-6963-7-55] [Medline: 17439645]
60. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
61. Hatem SM, Saussez G, Della Faille M, Prist V, Zhang X, Dispa D, et al. Rehabilitation of motor function after stroke: a multiple systematic review focused on techniques to stimulate upper extremity recovery. Front Hum Neurosci 2016;10:442 [FREE Full text] [doi: 10.3389/fnhum.2016.00442] [Medline: 27679565]
62. Baker C, Worrall L, Rose M, Hudson K, Ryan B, O'Byrne L. A systematic review of rehabilitation interventions to prevent and treat depression in post-stroke aphasia. Disabil Rehabil 2018 Aug;40(16):1870-1892. [doi: 10.1080/09638288.2017.1315181] [Medline: 28420284]
63. Australian Institute of Health and Welfare. 2017. Older Australians at a glance URL: https://www.aihw.gov.au/reports/older-people/older-australia-at-a-glance/contents/demographics-of-older-australians [accessed 2017-10-13] [WebCite Cache ID 6uAdlKvFB]
64. He W, Goodkind D, Kowal P. U.S. Census Bureau. Washington, DC: U.S. Government Publishing Office; 2016. An Aging World URL: https://www.census.gov/content/dam/Census/library/publications/2016/demo/p95-16-1.pdf [accessed 2018-10-15] [WebCite Cache ID 73BEC9FTi]
65. Deloitte Access Economics. Deloitte. 2013. The economic impact of stroke in Australia: Scoping cost effective prevention URL: https://www2.deloitte.com/au/en/pages/economics/articles/economic-impact-stroke-australia.html [WebCite Cache ID 6uAdyPwGJ]
66. Menger F, Morris J, Salis C. Aphasia in an internet age: wider perspectives on digital inclusion. Aphasiology 2015 Nov 07;30(2-3):112-132. [doi: 10.1080/02687038.2015.1109050]
67. Morris J, Mueller J, Jones M. Tomorrow's elders with disabilities: what the wireless industry needs to know. JED 2009 Nov 11;21(2-3):131-146. [doi: 10.1080/09544820903303431]
68. Rose TA, Worrall LE, Hickson LM, Hoffmann TC. Aphasia friendly written health information: content and design characteristics. Int J Speech Lang Pathol 2011 Aug;13(4):335-347. [doi: 10.3109/17549507.2011.560396] [Medline: 21682542]
69. Dietz A, Hux K, McKelvey ML, Beukelman DR, Weissling K. Reading comprehension by people with chronic aphasia: a comparison of three levels of visuographic contextual support. Aphasiology 2009 Jul;23(7-8):1053-1064. [doi: 10.1080/02687030802635832]

70. American Speech-Language-Hearing Association. Aphasia URL: https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589934663&section=Signs_and_Symptoms [WebCite Cache ID 70HL0Tp2a]
71. Pedersen PM, Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Aphasia in acute stroke: incidence, determinants, and recovery. Ann Neurol 1995 Oct;38(4):659-666. [doi: 10.1002/ana.410380416] [Medline: 7574464]
72. Bantan N, Abenhaim HA. Vaginal births after caesarean: what does Google think about it? Women Birth 2015 Mar;28(1):21-24. [doi: 10.1016/j.wombi.2014.10.004] [Medline: 25458612]
73. Gosselin MM, Mulcahey MK, Feller E, Hulstyn MJ. Examining internet resources on gender differences in ACL injuries: what patients are reading. Knee 2013 Jun;20(3):196-202. [doi: 10.1016/j.knee.2012.11.006] [Medline: 23266137]
74. Mercer MB, Agatisa PK, Farrell RM. What patients are reading about noninvasive prenatal testing: an evaluation of Internet content and implications for patient-centered care. Prenat Diagn 2014 Oct;34(10):986-993. [doi: 10.1002/pd.4410] [Medline: 24823864]
75. Sacks S, Abenhaim HA. How evidence-based is the information on the internet about nausea and vomiting of pregnancy? J Obstet Gynaecol Can 2013 Aug;35(8):697-703. [Medline: 24007704]
76. Tavare AN, Alsafi A, Hamady MS. Analysis of the quality of information obtained about uterine artery embolization from the internet. Cardiovasc Intervent Radiol 2012 Dec;35(6):1355-1362. [doi: 10.1007/s00270-012-0345-9] [Medline: 22274841]
77. Rose TA, Worrall LE, Hickson LM, Hoffmann TC. Guiding principles for printed education materials: design preferences of people with aphasia. Int J Speech Lang Pathol 2012 Feb;14(1):11-23. [doi: 10.3109/17549507.2011.631583] [Medline: 22136650]
78. Galliers J, Wilson S, Roper A. Words are not enoughmpowering people with aphasia in the design process. 2012 Presented at: Proceedings of the 12th Participatory Design Conference; 2012; Roskilde, Denmark p. 51-60.
79. Koppenol T, Al Mahmud A, Martens JB. When Words Fall Short: Helping People with Aphasia to Express. 2010 Presented at: International Conference on Computers for Handicapped Persons; 2010; Vienna, Austria p. 14-16.
80. Simic T, Leonard C, Laird L, Cupit J, Höbler F, Rochon E. A usability study of internet-based therapy for naming deficits in aphasia. Am J Speech Lang Pathol 2016 Nov 01;25(4):642-653. [doi: 10.1044/2016_AJSLP-15-0030] [Medline: 27784033]
81. Buhr HR, Hoepner JK, Miller H, Johnson C. AphasiaWeb: development and evaluation of an aphasia-friendly social networking application. Aphasiology 2016 Sep 22;31(9):999-1020. [doi: 10.1080/02687038.2016.1232361]
82. Hill AJ, Breslin HM. Refining an asynchronous telerehabilitation platform for speech-language pathology: engaging end-users in the process. Front Hum Neurosci 2016;10:640 [FREE Full text] [doi: 10.3389/fnhum.2016.00640] [Medline: 28066211]
83. Jorm AF, Morgan AJ, Malhi GS. The future of e-mental health. Aust N Z J Psychiatry 2013 Feb;47(2):104-106. [doi: 10.1177/0004867412474076] [Medline: 23382504]

Abbreviations

CBT: cognitive behavioral therapy
cCBT: computerized cognitive behavioral therapy
CINAHL: Cumulative Index to Nursing and Allied Health Literature
IPT: interpersonal therapy
RCT: randomized controlled trial
OnTrack—AD: OnTrack—Alcohol and Depression

Edited by G Eysenbach; submitted 19.01.18; peer-reviewed by S Northcott, M Danovska; comments to author 30.03.18; revised version received 19.07.18; accepted 21.08.18; published 04.12.18.

Please cite as:
Clunne SJ, Ryan BJ, Hill AJ, Brandenburg C, Kneebone I
Accessibility and Applicability of Currently Available e-Mental Health Programs for Depression for People With Poststroke Aphasia: Scoping Review
J Med Internet Res 2018;20(12):e291
URL: https://www.jmir.org/2018/12/e291/
doi:10.2196/jmir.9864
PMID:30514696

Original Paper

Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset

Amit Baumel[1], PhD; John M Kane[2], MD

[1]Department of Community Mental Health, University of Haifa, Haifa, Israel

[2]Zucker Hillside Hospital, Glen Oaks, NY, United States

Corresponding Author:
Amit Baumel, PhD
Department of Community Mental Health
University of Haifa
Abba Khoushy Ave 199
Haifa, 3498838
Israel
Phone: 972 48288496
Email: abaumel@univ.haifa.ac.il

Abstract

Background: The literature suggests that the product design of self-guided electronic health (eHealth) interventions impacts user engagement. Traditional trial settings, however, do not enable the examination of these relationships in real-world use.

Objective: This study aimed to examine whether the qualities of product design, research evidence, and publicly available data predict real-world user engagement with mobile and Web-based self-guided eHealth interventions.

Methods: This analysis included self-guided mobile and Web-based eHealth interventions available to the public—with their qualities assessed using the Enlight suite of scales. Scales included Usability, Visual Design, User Engagement, Content, Therapeutic Persuasiveness, Therapeutic Alliance, Credibility, and Research Evidence. Behavioral data on real-world usage were obtained from a panel that provides aggregated nonpersonal information on user engagement with websites and mobile apps, based on a time window of 18 months that was set between November 1, 2016 and April 30, 2018. Real-world user engagement variables included average usage time (for both mobile apps and websites) and mobile app user retention 30 days after download.

Results: The analysis included 52 mobile apps (downloads median 38,600; interquartile range [IQR] 116,000) and 32 websites (monthly unique visitors median 5689; IQR 30,038). Results point to moderate correlations between Therapeutic Persuasiveness, Therapeutic Alliance, and the 3 user engagement variables (.31≤*r*s≤.51; *P*s≤.03). Visual Design, User Engagement, and Content demonstrated similar degrees of correlation with mobile app engagement variables (.25≤*r*s≤.49; *P*s≤.04) but not with average usage time of Web-based interventions. Positive correlations were also found between the number of reviews on Google Play and average app usage time (r=.58; P<.001) and user retention after 30 days (r=.23; P=.049). Although several product quality ratings were positively correlated with research evidence, the latter was not significantly correlated with real-world user engagement. Hierarchical stepwise regression analysis revealed that either Therapeutic Persuasiveness or Therapeutic Alliance explained 15% to 26% of user engagement variance. Data on Google Play (number of reviews) explained 15% of the variance of mobile app usage time above Enlight ratings; however, publicly available data did not significantly contribute to explaining the variance of the other 2 user-engagement variables.

Conclusions: Results indicate that the qualities of product design predict real-world user engagement with eHealth interventions. The use of real-world behavioral datasets is a novel way to learn about user behaviors, creating new avenues for eHealth intervention research.

(J Med Internet Res 2018;20(12):e11491) doi:10.2196/11491

KEYWORDS

eHealth; mHealth; user engagement; user experience; therapeutic alliance; persuasive design; behavior change

Introduction

Background

Self-guided electronic health (eHealth) interventions have the potential to increase access to evidence-based care, while reducing the costs associated with service uptake [1,2]. The impact of these interventions, however, is limited by the ability to engage users in therapeutic activities and to support users' adherence to the therapeutic process [3,4]. As eHealth interventions require individuals to engage with self-care outside of traditional health care settings [5-7], individuals' engagement must compete with other events in their daily lives and fluctuating motivation to engage in effortful behavior [8]. As a result, user engagement with mobile apps and websites across the behavior change spectrum is low in the absence of human support [9].

There is a body of literature examining how intervention design may facilitate engagement and behavior change [10-14]. For example, systematic reviews have found relationships between the incorporation of therapeutic persuasiveness (ie, persuasive design and behavior change techniques) into the eHealth intervention and user adherence [15] and the intervention's efficacy [16]. Studies have also shown that the incorporation of conversational agents within self-guided eHealth interventions impacts user engagement [17,18], suggesting that conversational agents enhance the relational factors within the program—factors that are part of the therapeutic alliance fostered between the user and the program [19-21]. Although these studies provide convincing evidence linking product design to user engagement, the understanding of these relationships is still limited by our ability to examine user engagement in the real world, while comparing large numbers of products within the same study.

Comparing a Large Number of Products

Certain methodologies enable the cost-effective comparison of a large number of interventions. The Multiphase Optimization Strategy, for example, offers a paradigm to incorporate a randomized experimentation in a way that directly compares many different intervention components to identify the group of active components leading to the best outcome [22]. When it comes to product designs that are largely different in their functionalities, however, to control all the active components, the different functions would have to be available within the utilized digital platform—a process that would be highly expensive given the average development price of a single health app [23].

A novel way to compare large numbers of digital products is to utilize datasets that record user behaviors on a large number of websites and mobile apps. This approach can enable the documentation of variance in user engagement across a wide range of product designs using the same analytical framework. The big data commonly generated and stored by digital platforms can be used to learn about user behaviors in order to refine conceptual models or theories, and to understand processes related to eHealth interventions [24]. Additionally, using commonly generated data enables us to record real-world user behaviors outside of study settings that involve interactions with research staff—interactions that might impact user engagement [25,26].

Identifying the Different Quality Aspects of Product Design

To utilize commonly generated data and to understand their relationship with eHealth product design, there is also a need to identify the different aspects of product design in a reliable way. The literature suggests several approaches to evaluating eHealth product design (see the study by BinDhim et al [27] for a review). These include using a predefined list of what the app should contain, assessing the inclusion of evidence-based content, assessing the usability of predefined app functions, using consumer reviews and ratings, and utilizing criteria-based rating scales. Criteria-based rating scales rest on a heuristic evaluation approach—a method that has been broadly researched and used for assessing eHealth and technology products [28-30]. Heuristics are broad principles of product design that can be inspected by evaluators before empirical testing. The advantage of heuristic evaluation is that it enables the cost-efficient identification of design problems without the need for a predefined list of features that may exclude the users' experience of utilizing the product [29,31,32]. Criteria-based rating scales are based on core concepts, each comprising different heuristics, which are used by trained raters to objectively examine and score the quality of eHealth programs [12,33-35]. In the absence of such clearly defined rating systems, scoring tends to be less reliable [36].

Enlight is a suite of criteria-based rating scales that was used in this study. It covers 11 different quality constructs and checklists (eg, Usability and Credibility) that were produced by trained raters. Enlight is the first suite of measures to incorporate behavior change, persuasive design, and therapeutic alliance concepts—concepts that have been found to affect a program's therapeutic potential [15,16]. As the tool shows high inter-rater reliability scores at the construct level, it enables us to use it to examine the relationships between different aspects of product design and metrics of user engagement.

Within this context, Baumel and Yom-Tov conducted a preliminary investigation, examining the correlations between 6 quality aspects of product design and real-world user engagement with 30 Web-based programs [37]. Real-world user engagement was based on a proprietary dataset of anonymized logs from consenting users of a Microsoft Internet Explorer add-on. The quality scores were generated using Enlight [21]. In the preliminary study, it was found that product quality ratings predicted which Web-based interventions were more engaging and in particular that Therapeutic Persuasiveness was the most robust predictor of user adherence (ie, duration of use, number of unique sessions; $40 \leq rs \leq .58$; $Ps \leq .005$) [37].

Although these findings were novel in terms of the methods applied, they were limited in several aspects. First, the dataset only included Web-based interventions, whereas the literature suggests that there has been a massive increase in mobile phone ownership and mobile health app usage in recent years [38]. Second, as the analysis was based on a small sample of interventions, the question arises as to whether the same pattern of results will be found with a larger dataset, enabling

identification of correlations with other aspects of product design. Third, the study did not incorporate important metrics that relate to research evidence, a product's credibility, and publicly available data on user acceptance (eg, user ratings on app stores). Examining the relationship between research evidence and real-world uptake is key in light of the notion that efficacy trials largely emphasize internal validity over real-world issues, such as the technological environment, implementation, and sustainability, and thus may not provide the needed validation [26]. Furthermore, to the best of our knowledge, there has been no study linking the credibility of the source that developed the program or user ratings on app stores *and* real-world user engagement with eHealth programs.

This Study

The aims of this study were, therefore, to fill this gap in the literature by (1) examining whether different qualities of product design predict real-world user engagement with both mobile- and Web-based self-guided eHealth interventions; (2) exploring the associations between scale items, data, and real-world user engagement; (3) examining whether research evidence and product credibility metrics are associated with real-world user engagement; (4) examining the associations between publicly available data on program acceptance (eg, star ratings on app stores) and real-world usage and whether these data enhance the prediction of user engagement above expert ratings; and (5) establishing Enlight's validity in predicting user engagement with eHealth interventions based on a large and independent dataset of user behaviors.

Methods

Selection of Interventions

We screened for eligibility all eHealth programs that were assessed based on Enlight suite of scales between September 2016 and December 2017—during the scale development phase and afterward as part of a nonprofit project aimed at evaluating the quality of eHealth programs [39]. The clinical aims of the selected programs were broad, spanning the behavioral health domain. These programs could be grouped into those targeting mental health (eg, depression, anxiety, and well-being) and those targeting health-related behaviors (eg, diet, physical activity, and smoking and alcohol cessation).

The sources of the eHealth programs that were screened for eligibility are presented in Table 1. A total of 84 programs were randomly selected and rated between September and December 2016, following a systematic identification process, conducted as part of Enlight tool's development [21]. In this process, we created a list of mobile apps and Web-based programs through keyword searches in Google Play and Google search engine (this systematic process and a complete list of keywords is presented in Multimedia Appendix 1). Following the tool's development phase, we used a similar systematic identification process—that now also included paid programs—to rate additional programs between January and December 2017. All programs found in the top 10 Web or mobile app search results (and that had not been rated before) were rated, reaching a number of additional 50 eHealth programs. Finally, 8 eHealth programs were identified based on recommendations from eHealth researchers and product developers (eg, the Digital Behavioral Health Workgroup at Northwell Health). This selection process yielded a list of 142 programs, among which, 21 programs had both mobile and Web-based versions. As the behavioral data on program usage and our inclusion criteria in this study relate separately to websites and mobile apps, we eventually screened for eligibility 71 Web-based programs and 92 mobile apps.

Inclusion and Exclusion Criteria

To be included in this analysis, interventions had to be (1) self-guided and (2) cost-free. Apps that were free to install, yet had a trial period (ie, paid with free trial) were also included; however, we examined their impact on results using a sensitivity analysis, as will be further described in the data analysis section. Exclusion criteria included (1) programs that were only trackers, as their quality ratings do not rely on the full list of product ratings covered in Enlight; (2) websites that were not focused mostly on the intervention itself (eg, websites with many blog articles that did not require user log-in)—as the data on the platform are provided in an aggregated way that does not distinguish between the website's different Web pages; (3) apps that did not have an Android version, as the behavioral dataset did not include information on iOS apps; and (4) programs without usage data in our behavioral dataset (due to, for example, small number of users).

Measures

Enlight Quality Ratings

Enlight is a comprehensive suite of criteria-based measurements—completed by trained raters who review the eHealth program—that was developed by a multidisciplinary team through a rigorous process of content analysis, grouping, and classification of 476 identified criteria items [21,40]. The tool covers 6 different product quality domains: Usability, Visual Design, User Engagement, Content, Therapeutic Persuasiveness (ie, persuasive design and incorporation of behavior change techniques), and Therapeutic Alliance (see Multimedia Appendix 2 for a detailed description of the scales used and operational definitions of all items). Each quality domain score ranges from 1 to 5 and is based on averaging the criteria ratings produced by the raters on a Likert scale (ranging from 1 to 5). For example, Therapeutic Persuasiveness is calculated by averaging the raters' scores of the following criteria items: call to action (eg, goal settings, prompting goals, and encouragement to complete goals), load reduction of activities (eg, set graded tasks), therapeutic rationale and pathway (eg, reasoned action and provide information about behavioral health link), rewards (eg, contingent rewards), real data-driven/adaptive content (eg, self-monitoring of behavior), ongoing feedback, and expectations and relevance.

Table 1. Sources of previously rated eHealth programs that were screened for eligibility in this study.

Source	eHealth programs (N=142), n (%)	Web-based program (N=71), n (%)	Mobile app (N=92), n (%)	Programs with 2 delivery mediums[a] (N=21), n (%)
Systematic review: programs randomly selected and rated during Enlight's development phase	84 (59.2)	43[b] (61)	49[b] (53)	8 (38)
Additional programs found in the top 10 mobile or Web search results of the systematic identification process	50 (35.2)	25 (35)	38 (41)	13 (62)
Additional programs that were identified based on personal recommendations	8 (5.6)	3 (4)	5 (5.4)	0 (0)

[a]The 2 delivery mediums are (1) Web-based programs and (2) mobile apps.

[b]In the original study, we examined 42 mobile apps and 42 Web-based programs. Eventually, 7 websites and 1 mobile app had a similar version in the other delivery medium (mobile or website).

Enlight covers 2 additional measures that were included in this analysis. Credibility consists of a checklist calculated by aggregating the scores received in each of its respective categorical items (owners' credibility, maintenance, strong advisory support, third party endorsement, and evidence of successful implementation). The checklist differs from the product quality assessments in 2 ways: (1) most credibility items cannot be rated before product deployment (as they rely on data that are collected afterward) and (2) the criteria included in the checklist are not expected to directly impact the end users' experience of the product's efficacy (however, they may expose the user to acknowledged risks or benefits). We also included an evidence-based program scale (ie, research evidence) that assesses the quality of the empirical research supporting the program within the current zeitgeist on a scale from 1 (very poor) to 5 (very good) [21].

A total of 3 individuals with clinical experience (eg, clinical psychologists) and with at least 1 year experience working in the eHealth domain (eg, content writing of eHealth programs and health technology coaching) performed the ratings, with 2 of them independently rating each program. Their training included a review of Enlight items, individual ratings of 7 eHealth programs that included group-solicited feedback on ratings, and then testing the ratings based on 5 additional programs (for a detailed review, see the study by Baumel et al [21]). In this study (post-training), the Enlight quality scales exhibited excellent inter-rater reliability scores between the 2 raters (intraclass correlations=.74 to .98, median .86).

Behavioral Data on User Engagement in the Real World

Information on user behaviors was obtained from SimilarWeb's Pro panel data [41]. The panel provides aggregated nonpersonal information on user engagement with websites and mobile apps all over the world to enable Web and mobile app traffic research and analytics. It is based on several sources of anonymized usage data, such as that obtained from consenting users of mobile apps or browser add-ons (ie, products). A dedicated product team at SimilarWeb is responsible for building and partnering with hundreds of high-value consumer products that make up the panel. According to SimilarWeb, the products are used across diverse audiences, without cluttering the user with advertisements. Although benefiting from the products, users contribute to the panel, as they enable to document their online or mobile apps' usage activities seamlessly and anonymously [41]. The data are not used by SimilarWeb or provided to any third parties for the purposes of marketing, advertising, or targeting any individual subjects. The data-gathering procedures comply with data privacy laws, including the way the data are collected, anonymized, stored, secured, and used. These procedures are updated regularly based on evolving data privacy legislation and requirements, such as the European Union's General Data Protection Regulation [42].

The study was approved by University of Haifa Institutional Review Board. The measures were set to include data gathered over an 18-month period from November 1, 2016 to April 30, 2018. To examine user engagement with mobile apps, we used 2 measures retrieved from SimilarWeb Pro: (1) average app usage time (of all users) and (2) the percentage of users who downloaded the app and were still using it 30 days later. As this information is provided separately per country, we examined the country with the most app downloads. User engagement with the Web-based interventions was calculated based on average monthly visits to the website multiplied by the average visit duration.

To test the reliability of the data for the remaining websites, we calculated the Spearman correlation between Web usage time obtained from the SimilarWeb platform and the usage time calculated for websites in our previous work with Microsoft Research [37]. A strong positive correlation (r=.69; P<.001) was found between the 2 datasets for the 17 websites that had data on both platforms. In light of the difference between the 2 datasets (our previous work was focused only on Explorer browser users), we also examined the Spearman correlation between website global popularity ranks (ie, a rank that reflects the extent to which a website is utilized all over the world) on SimilarWeb platform and Alexa [43]—an independent source of information on user traffic. A very strong Spearman correlation was found (n=28; r=.78; P<.001). Relating to the validity of mobile app usage data, an Oath researcher [44] (RW) examined 30 randomly selected mobile apps with data on SimilarWeb and Oath's own independent records of mobile app usage data. The researcher examined the correlation between the average number of user sessions per day in the 2 datasets, finding a very strong Spearman correlation (n=30; r=.77; P<.001). These findings suggested that there was sufficient convergent validity, which can be claimed if the correlation

coefficient is above .50, although a value above .70 is usually recommended [45].

Finally, we documented publicly available data that relate to a program's acceptability. The reported number of installs, average star ratings, and number of reviews were obtained from the Google Play store. Websites' total monthly visits were obtained from the publicly available version of SimilarWeb.

Data Analysis

Medians and interquartile ranges (IQRs) were used to present the distribution of the variables. Pearson correlations were used to examine the relationships between the continuous variables. As Pearson correlations and linear regressions assume a normal distribution of the variables, we examined the skewness and kurtosis of the variables within acceptable limits of ±2 and performed a logarithmic transformation (base 10) of these variables [46,47]. This transformation was eventually applied to the following variables: Credibility, Research Evidence, number of reviews on Google Play, mobile app usage time, and website's total monthly visits.

A sensitivity analysis was conducted to examine whether the pattern of correlations found differed depending on (1) the clinical aim of the intervention (ie, mental health or health-related behaviors) and (2) whether paid programs with a free trial were included or excluded. The differences between correlations of these groups were calculated using Fisher Z-transformation. Independent sample *t* tests with Bonferroni correction were performed to examine the difference between categorical items on the behavioral measures.

Finally, a series of hierarchical stepwise linear regressions was applied to examine the ability to predict user engagement in the real world independent of empirical testing. In the first step, we examined the percentage of variance explained by Enlight quality ratings. In the second step, we examined whether empirical data on real-world usage that were available at no cost (eg, Google Play for mobile apps: number of downloads, star reviews, and number of reviews; SimilarWeb for websites: monthly visits) significantly increased the explained variance. In each of these steps, a stepwise approach was applied to avoid adding predicting variables that did not contribute significantly to the overall model.

Results

The analysis included user behavior data involving 52 mobile apps and 32 websites (see Figure 1 for the flow diagram of program selection and Multimedia Appendix 3 for a full list of included programs). Overall, 9 programs had usage data for both Web- and mobile-app versions, reaching a total of 75 different programs with data. Within the 18-month time window of the analysis, the 52 mobile apps had a median of 38,600 downloads (IQR 116,000), and the 32 Web-based interventions had a median of 5689 monthly unique visitors (IQR 30,038). The median monthly usage time was 5.11 min (IQR 20.51) for mobile apps and 9.0 min (IQR 7.37) for websites. The medians and IQRs of Enlight product quality ratings were as follows: Usability median 3.66 (IQR 1.66); Visual Design median 3.00 (IQR 1.00); User Engagement median 3.20 (IQR 1.10); Content median 3.50 (IQR 1.00); Therapeutic Persuasiveness median 2.71 (IQR 1.00); and Therapeutic Alliance median 2.66 (IQR 1.33).

Pearson correlations between Enlight product quality ratings, Credibility, and Research Evidence are presented in Table 2. User Engagement, Content, Therapeutic Persuasiveness, and Therapeutic Alliance were positively correlated with Credibility (.26≤*r*s≤.51, *P*s≤.01) and Research Evidence (.21≤*r*s≤.39, *P*s≤.04). That is, the interventions with higher scores in these quality domains had higher Credibility and Research Evidence ratings.

Pearson correlations between Enlight quality ratings and items, Credibility and Research Evidence, publicly available data on program acceptance, and the 3 behavioral variables are presented in Table 3. Results point to moderate positive correlations between Therapeutic Persuasiveness, Therapeutic Alliance, and the 3 behavioral variables (.31≤*r*s≤.51, *P*s≤.03). Visual Design, User Engagement, and Content had a similar degree of correlations with the user engagement variables of mobile apps (.25≤ *r*s≤.49, *P*s≤.04), but not Web-based interventions.

Altogether, a similar pattern of results was found between the criteria items of domains with significant correlations and the respective behavioral variables. It is worth noting that 3 criteria items of the Therapeutic Persuasiveness domain—rewards, data-driven/adaptive, and ongoing feedback—showed significant correlations with mobile app user retention after 30 days; however, these items did not correlate significantly with mobile app usage time. In terms of the Enlight checklists, a program's Credibility had positive correlations with the usage time of apps (*r*=.34; *P*=.006) and websites (*r*=.30; *P*=.04), but no significant correlations were found between Research Evidence and the 3 behavioral variables. Finally, results point to positive correlations between the number of reviews on Google Play and average app usage time (*r*=.58; *P*<.001) and user retention after 30 days (*r*=.23; *P*=.049). The number of installs and average star reviews had significant positive correlations with mobile app usage time (*r*s=.25 and .36, respectively; *P*s=.04 and .005, respectively).

As the Credibility checklist covers different independent categorized items, a series of independent sample *t* tests with Bonferroni correction was performed to examine the difference between Credibility items (Owners' Credibility, Maintenance, Strong Advisory Support, Evidence of Successful Implementation) in terms of the usage time of mobile and Web-based interventions. (Third Party Endorsement was not included in this test as a relevant party had endorsed only a few programs.) A significant difference was found in the average usage time of mobile apps, favoring programs that had Evidence of Successful Implementation (ie, a high number of downloads or positive reviews) t_{50}=3.88, *P*<.001, Cohen *d*=1.07, and in maintenance, favoring programs that had been updated within the previous 6 months, t_{50}=2.63, *P*=.28, Cohen *d*=0.73. As both variables relied on the data documented on Google Play, we added them in the second step of the regression analysis described below.

Figure 1. Flow diagram of program selection.

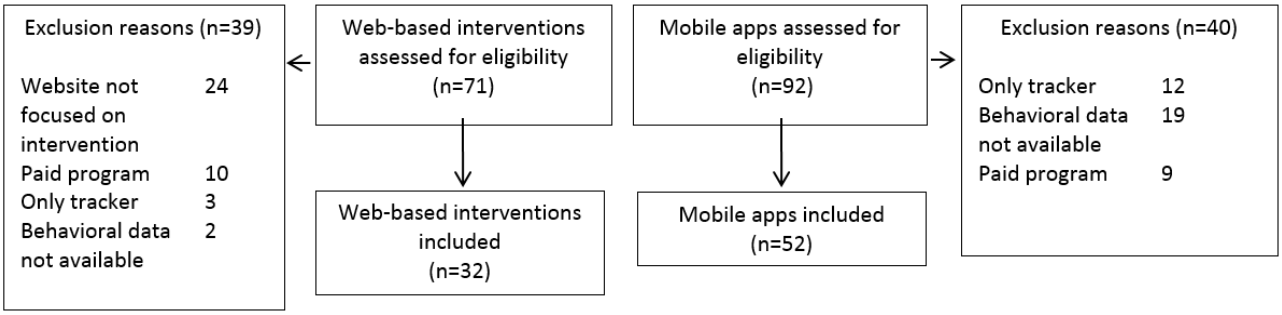


Table 2. Pearson correlations between Enlight product quality ratings, Credibility, and Research Evidence metrics (n=75).

Scale	Credibility		Research evidence	
	r [a]	*P* value	*r*	*P* value
Usability	.02	.45	.00	.50
Visual Design	.08	.24	−.13	.14
User Engagement	.39[b]	<.001	.21[b]	.03
Content	.51[b]	<.001	.39[b]	<.001
Therapeutic Persuasiveness	.26[b]	.01	.21[b]	.04
Therapeutic Alliance	.39[b]	<.001	.26[b]	.01

[a]Using Fisher Z-transformation, no significant differences in Pearson correlation values were found between programs targeting mental health (n=51) and those targeting health-related behaviors (n=24).

[b]Indicates significant correlations.

Table 3. Pearson correlations between Enlight scales, publicly available data on program acceptance, and real-world user engagement with eHealth interventions.

Scale	Mobile app interventions (n=52)				Web interventions (n=32)	
	Usage time		User 30-day retention		Usage time	
	r [a]	*P* value	*r*	*P* value	*r*	*P* value
Usability	.15	.14	−.07	.32	.01	.48
Ease of use	.21	.07	.05	.36	.10	.29
Learnability	.07	.31	−.13	.18	−.14	.22
Navigation	.13	.18	−.11	.22	.03	.43
Visual Design	.31[b]	.01	.25[b]	.04	.09	.32
Aesthetics	.28[b]	.02	.28[b]	.02	.12	.26
Layout	.32[b]	.01	.19	.09	.14	.22
Size	.24[b]	.047	.18	.10	−.00	.49
User Engagement	.49[b]	<.001	.39[b]	.002	.21	.12
Captivating	.51[b]	<.001	.31[b]	.01	.38[b]	.02
Content presentation	.53[b]	<.001	.29[b]	.02	.22	.12
Interactive	.13	.19	.25[b]	.04	.28	.06
Not irritating	.23	.07	.17	.15	−.11	.31
Targeted/tailored/personalized	.41[b]	.001	.38[b]	.003	.01	.48
Content	.45[b]	<.001	.33[b]	.009	.16	.19
Evidence-based content	.43[b]	.001	.31[b]	.01	.15	.20
Information provision quality	.35[b]	.006	.34[b]	.006	−.01	.48
Complete and concise	.39[b]	.002	.35[b]	.005	.27	.07
Clarity about program’s purpose	.43[b]	.001	.10	.23	.09	.30
Therapeutic Persuasiveness	.36[b]	.004	.39[b]	.002	.33[b]	.03
Call to action	.34[b]	.007	.49[b]	<.001	.29	.055
Load reduction of activities	.44[b]	.001	.37[b]	.004	.37[b]	.02
Therapeutic rationale/pathway	.47[b]	<.001	.36[b]	.004	.21	.12
Rewards	.14	.16	.26[b]	.03	.22	.12
Data-driven/adaptive	.11	.21	.24[b]	.046	.23	.10
Ongoing feedback	.14	.16	.33[b]	.01	.14	.23
Expectations and relevance	.31[b]	.01	.03	.41	.19	.15
Therapeutic Alliance	.51[b]	<.001	.31[b]	.01	.36[b]	.02
Basic acceptance and support	.45[b]	<.001	.40[b]	.002	.42[b]	.009
Positive therapeutic expectations	.45[b]	<.001	.33[b]	.009	.22	.12
Relatability	.40[b]	.002	.09	.26	.32[b]	.04
Enlight postempirical measures						
Credibility	.34[b]	.006	−.03	.42	.31[b]	.04
Research evidence	.08	.28	−.13	.18	.19	.15

Scale	Mobile app interventions (n=52)				Web interventions (n=32)	
	Usage time		User 30-day retention		Usage time	
	r [a]	*P* value	*r*	*P* value	*r*	*P* value
Google Store data						
Number of installs	.25[b]	.04	.10	.25	N/A[c]	N/A
Average star reviews	.36[b]	.005	.18	.11	N/A	N/A
Number of reviews	.58[b]	<.001	.23[b]	.049	N/A	N/A
Web traffic data						
Total monthly visits	N/A	N/A	N/A	N/A	.19	.15

[a]Using Fisher Z-transformation, no significant differences in Pearson correlation values were found between programs targeting mental health (mobile apps n=38 and websites n=22) and those targeting health-related behaviors (mobile apps n=14 and websites n=10) or between the full samples and the samples after subtracting paid programs with free trials (n=7).

[b]Indicates significant correlations.

[c]N/A: not applicable.

Table 4. Hierarchical stepwise regressions for predictors of user engagement with self-guided mobile and Web-based interventions.

Variable	B[a]	SE B	Beta	*P* value	R^2 change	*P* value
Mobile app usage time (n=52)						
Step 1						
Therapeutic Alliance	.47	.11	.51	<.001	.26	<.001
Step 2						
Therapeutic Alliance	.29	.11	.31	.02	.15	.001
Number of reviews	.29	.08	.44	.001	N/A[b]	N/A
Mobile app 30-day retention (n=52)						
Step 1						
Therapeutic Persuasiveness	.01	.01	.39	.004	.15	.004
Step 2[c]	N/A	N/A	N/A	N/A	N/A	N/A
Web intervention total usage time (n=32)						
Step 1						
Therapeutic Alliance	2.87	1.34	.36	.04	.13	.04
Step 2[c]	N/A	N/A	N/A	N/A	N/A	N/A

[a]B: unstandardized regression coefficient.

[b]N/A: not applicable.

[c]No new variables were entered into the equation.

Table 4 presents the hierarchical stepwise regressions performed to examine the predictors and percentages of real-world user engagement variance that were explained by Enlight product quality ratings and whether publicly available data on program acceptance added to the explained variance. The analysis showed that the variance of user engagement explained by Enlight ratings ranged between 13% and 26%. In each of the first steps, following the entrance of one Enlight quality rating—either Therapeutic Persuasiveness or Therapeutic Alliance—no other metric was found to significantly contribute to the model. The analysis also showed that publicly available data (number of reviews) explained 15% of the variance of mobile app usage time above Enlight ratings; publicly available data did not significantly explain the variance of mobile app user retention or Web-based intervention usage time.

Discussion

Principal Findings

This study presents novel findings about the relationship between product design and user engagement with self-guided eHealth interventions in the real world. We found that product quality in terms of Visual Design, User Engagement, Content, Therapeutic Persuasiveness, and Therapeutic Alliance was positively correlated with real-world usage of mobile apps. Therapeutic Persuasiveness and Therapeutic Alliance were also

positively correlated with real-world usage of Web-based programs. Similar to previous findings, Visual Design, User Engagement, and Content were not found to have a significant correlation with usage time of Web-based interventions [37]. Although the domains' criteria items had similar patterns of correlation with the behavioral variables, 3 Therapeutic Persuasiveness items—rewards, data-driven/adaptive, and ongoing feedback —showed significant correlations with mobile app user retention after 30 days but no significant correlations with mobile app usage time. It might be that these variables, which examine the ongoing reciprocal interaction between the software and the user, are sensitive to the time window of use rather than to the accumulated usage time. However, future research should first be conducted to examine whether this pattern of results is replicated.

Congruent with our previous examination [37], Usability was not associated with the behavioral variables of user engagement. It is important to note that as our analysis was based on between-program evaluation, it does not mean that improving program's usability will not enhance user engagement with the program. Alternatively, we suggest that using the Usability score to compare *different* programs might not capture which programs are more engaging to users. Therefore, Usability should be considered to be a barrier rather than a facilitator of user engagement [48,49]. A future direction could be to define ranges—in terms of Usability scores—that identify the point at which a program is usable enough to be evaluated based on other metrics.

Our analysis of Google Play data revealed positive correlations between the number of reviews, mobile app usage time, and user retention; number of installs and average star reviews were also positively correlated with mobile app usage time. These findings present a link between data available to the public on app stores and an app's overall tendency to engage users. Regression results, however, suggested that these data do not always enhance our understanding of user engagement after accounting for the quality of product design—which could be determined before real-world use of the program. Hence, it could be informative to examine how expert-based rating tools such as Enlight can be used by developers during a program's development phase to guide the process of product design before empirical testing.

Our analysis also showed that programs with better design quality had better research evidence. At the same time, research evidence did not predict user engagement in real-world use. This finding is congruent with Fleming et al's study that examined published reports on the real-world uptake of eHealth programs [50]. The researchers found that indications of completion or sustained use (ie, completion of all modules or the last assessment or continuing to use apps after 6 weeks or more) varied from 0.5% to 28.6%, concluding that actual usage may vary from that reported in trials. Accordingly, our analysis implies that when research evidence supports a certain program, it does not necessarily mean that users will engage with the program in the real world. It is important to note that there may be many reasons for what could be referred to as "trial versus real-world gap," including the impact of the study setting on user engagement [25] and populations being targeted for the study that differ from those using the program in the real world [26]. Future research should to be conducted to empirically examine this phenomenon and the factors influencing it.

It is also important to note that the analysis did not reveal an association between the Credibility items—developer's credibility (eg, academic institute) and strong advisory support group—and user engagement variables. One reason that could explain this finding is the high costs associated with app development [23], which may create barriers for teams that are more focused on answering academic research questions based on grant funding. However, this explanation should be further examined to draw firm conclusions.

Finally, from a methodological perspective, the notion that stands behind the development of criteria-based rating tools is that in the absence of proper training, proper expertise, and proper use of developed scales, inter-rater agreement will be low, and therefore, the examination will not be reliable. For example, in a study to examine inter-rater reliability of mobile health (mHealth) app rating measures, Powell et al gathered a panel of 6 reviewers to review the top 10 depression apps and top 10 smoking cessation apps from the Apple iTunes App Store and evaluate them based on several measures. The authors found a wide variation in the inter-rater reliability of measures, with some measures being more robust across categories of apps than others [36]. In recent years, several studies have demonstrated that it is feasible to achieve sufficient inter-rater reliability with criteria-based rating scales by using trained raters [35,51]. To perform the ratings for this study and previous evaluations using the Enlight tool, raters had to complete a certain level of training to provide reliable ratings. This notion has been acknowledged for decades within the psychological assessment field ([52,53]; George et al, unpublished data, 1985) we hope that as the eHealth evaluation field moves forward, more attention will be paid to the proper use of methods to train and examine evaluators' work.

Limitations

This study has several limitations. First, the findings are not based on an experimental procedure that compared different designs of the *same* intervention. However, it would not be possible to utilize an experimental procedure to compare many aspects of product design at the same time. Consequently, if an association is not found between a quality domain and program usage, it does not mean that improving the program in this domain will not influence usage. Instead, it means that when comparing *different* programs, some aspects of quality are more important than others in predicting usage time. Second, this study examined user engagement, which is not the same as efficacy. Data suggest that there is a strong relationship between engagement and efficacy [54-56]; however, more does not always equal better [57]. A future research direction would be to measure efficacy using a large sample of programs "in the real world" and to examine the correlations between product design and efficacy. This testing should take into account fundamental questions related to participant consent and how to measure intervention outcomes.

Conclusions

Results indicate that Enlight is a valid tool for predicting real-world user engagement with eHealth interventions, based on expert evaluations that were conducted before empirical testing with end users. The link between expert reviews and user behaviors in the real world supports the importance of rating tools that may enable trained experts to (1) guide the design of evidence-based interventions before testing with end users and (2) provide important details about the product's potential to enable end users to make educated decisions when searching for self-guided interventions. Such details are presented in MindTools.io [39], a nonprofit website that publishes in-depth app reviews using Enlight rating scales.

Finally, the use of real-world behavioral datasets, which are garnered from a massive number of users, is a novel way to learn about user behaviors, creating new avenues of research and advancing our understanding of eHealth interventions. More studies are needed to shed light on the relationships between real-world uptake and data that emerge from other sources of information.

Acknowledgments

AB and JMK's work on eHealth rating scales and the development of MindTools.io is supported by the Donald & Barbara Zucker Family Foundation. The authors would like to thank Dr Ran Wolff for his support in examining the validity of the behavioral user panel they were using.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of the systematic identification process.
[PDF File (Adobe PDF File), 60 KB - jmir_v20i12e11491_app1.pdf]

Multimedia Appendix 2
Description of scales and scale items.
[PDF File (Adobe PDF File), 144 KB - jmir_v20i12e11491_app2.pdf]

Multimedia Appendix 3
Included programs.
[PDF File (Adobe PDF File), 63 KB - jmir_v20i12e11491_app3.pdf]

References

1. Kazdin AE. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017 Dec;88:7-18. [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
2. Baumel A, Baker J, Birnbaum ML, Christensen H, De Choudhury M, Mohr DC, et al. Summary of key issues raised in the Technology for Early Awareness of Addiction and Mental Illness (TEAAM-I) Meeting. Psychiatr Serv 2018 May 01;69(5):590-592. [doi: 10.1176/appi.ps.201700270] [Medline: 29334875]
3. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
4. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
5. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
6. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
7. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015 Aug 19;24(5):321-332. [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
8. Baumeister RF, Vohs KD. Self-Regulation, ego-depletion, and motivation. Soc Personal Psychol Compass 2007;1(1):115-128. [doi: 10.1111/j.1751-9004.2007.00001]
9. 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]

10. 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]
11. Kok G, Schaalma H, Ruiter RA, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]
12. Provost M, Koompalum D, Dong D, Martin BC. The initial development of the WebMedQual scale: domain assessment of the construct of quality of health web sites. Int J Med Inform 2006 Jan;75(1):42-57. [doi: 10.1016/j.ijmedinf.2005.07.034] [Medline: 16169770]
13. Muench F, Baumel A. More than a text message: dismantling digital triggers to curate behavior change in patient-centered health interventions. J Med Internet Res 2017 Dec 26;19(5):e147 [FREE Full text] [doi: 10.2196/jmir.7463] [Medline: 28550001]
14. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
15. 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] [Medline: 23151820]
16. 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;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
17. Bickmore T, Schulman D, Yin L. Maintaining engagement in long-term interventions with relational agents. Appl Artif Intell 2010 Jul 01;24(6):648-666 [FREE Full text] [doi: 10.1080/08839514.2010.492259] [Medline: 21318052]
18. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Interv 2017;10:39-46. [doi: 10.1016/j.invent.2017.10.002] [Medline: 30135751]
19. Cavanagh K, Millings A. (Inter) personal computing: the role of the therapeutic relationship in e-mental health. J Contemp Psychother 2013;43(4):197-206. [doi: 10.1007/s10879-013-9242-z]
20. Holter MT, Johansen A, Brendryen H. How a fully automated eHealth program simulates three therapeutic processes: a case study. J Med Internet Res 2016 Jun 28;18(6):e176 [FREE Full text] [doi: 10.2196/jmir.5415] [Medline: 27354373]
21. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. J Med Internet Res 2017 Mar 21;19(3):e82 [FREE Full text] [doi: 10.2196/jmir.7270] [Medline: 28325712]
22. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
23. Research2Guidance. It costs $425,000 on average to develop an mHealth app – Download Free Whitepaper URL: https://research2guidance.com/it-costs-425000-dollars-on-average-to-develop-an-mhealth-app-download-free-whitepaper/ [accessed 2018-06-24] [WebCite Cache ID 70Pgz4yuP]
24. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler CK, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. J Behav Med 2016;40(1):85-98. [doi: 10.1007/s10865-016-9818-7]
25. Ebert DD, Baumeister H. Internet-based self-help interventions for depression in routine care. JAMA Psychiatry 2017 Aug 01;74(8):852-853. [doi: 10.1001/jamapsychiatry.2017.1394] [Medline: 28636717]
26. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. Psychiatr Serv 2017 May 01;68(5):427-429. [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
27. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
28. Madan A, Dubey S. Usability evaluation methods: a literature review. Int J Eng Sci 2012;4(2):590-599 [FREE Full text]
29. Kientz JA, Choe EK, Birch B, Maharaj R, Fonville A, Glasson C, et al. Heuristic evaluation of persuasive health technologies. 2010 Presented at: Proceedings of the 1st ACM International Health Informatics Symposium; November 11-12, 2010; Arlington, VA, USA.
30. Monkman H, Griffith J, Kushniruk AW. Evidence-based heuristics for evaluating demands on eHealth literacy and usability in a mobile consumer health application. Stud Health Technol Inform 2015;216:358-362. [doi: 10.3233/978-1-61499-564-7-358] [Medline: 26262071]
31. Nielsen J. Enhancing the explanatory power of usability heuristics. 1994 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 24 - 28, 1994; Boston, MA, USA.
32. Karat CM. A Comparison of user interface evaluation methods. In: Nielsen J, Mack RL, editors. Usability inspection methods. New York: John Wiley & Sons; 1994:203-233.

33. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [Medline: 10396471]
34. Olsina L, Rossi G. Measuring web application quality with WebQEM. IEEE Multimed 2002;9(4):20-29. [doi: 10.1109/MMUL.2002.1041945]
35. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
36. Powell AC, Torous J, Chan S, Raynor GS, Shwarts E, Shanahan M, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. JMIR Mhealth Uhealth 2016;4(1):e15 [FREE Full text] [doi: 10.2196/mhealth.5176] [Medline: 26863986]
37. Baumel A, Yom-Tov E. Predicting user adherence to behavioral eHealth interventions in the real world: examining which aspects of intervention design matter most. Transl Behav Med 2018 Sep 08;8(5):793-798. [doi: 10.1093/tbm/ibx037] [Medline: 29471424]
38. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
39. MindTools.io. URL: https://mindtools.io/ [accessed 2018-06-24] [WebCite Cache ID 70Pg0KTXJ]
40. Baumel A, Birnbaum ML, Sucala M. A systematic review and taxonomy of published quality criteria related to the evaluation of user-facing eHealth programs. J Med Syst 2017 Aug;41(8):128. [doi: 10.1007/s10916-017-0776-6] [Medline: 28735372]
41. SimilarWeb. URL: https://www.similarweb.com [accessed 2018-06-18] [WebCite Cache ID 70GZ6eAt5]
42. European Commission. Data protection: Rules for the protection of personal data inside and outside the EU URL: https://ec.europa.eu/info/law/law-topic/data-protection_en [accessed 2018-06-18] [WebCite Cache ID 70GY1yLYp]
43. Alexa. URL: http://www.alexa.com/ [accessed 2018-01-04] [WebCite Cache ID 6nGSLql6u]
44. Oath. URL: https://www.oath.com/ [accessed 2018-10-14] [WebCite Cache ID 73A1knb5w]
45. Carlson KD, Herdman AO. Understanding the impact of convergent validity on research results. Organ Res Methods 2012;15(1):17-32. [doi: 10.1177/1094428110392383]
46. Gravetter F, Wallnau L. Statistics for the Behavioral Sciences. Belnmont, CA, USA: Cengage Learning; 2012.
47. George D, Mallery P. IBM SPSS Statistics 23 step by step: A simple guide and reference. New-York, NY, USA: Routledge; 2016.
48. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Assoc Inf Sci Technol 2008;59(6):938-955. [doi: 10.1002/asi.20801]
49. Atkinson M, Kydd C. Individual characteristics associated with World Wide Web use: an empirical study of playfulness and motivation. ACM SIGMIS Database 1997;28(2):53-62. [doi: 10.1145/264701.264705]
50. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
51. Dobbins M, Watson S, Read K, Graham K, Nooraie RY, Levinson AJ. A tool that assesses the evidence, transparency, and usability of online health information: development and reliability assessment. JMIR Aging 2018;1(1):e3. [doi: 10.2196/aging.9216]
52. Lobbestael J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). Clin Psychol Psychother 2011;18(1):75-79. [doi: 10.1002/cpp.693] [Medline: 20309842]
53. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997 Jul;36(7):980-988. [doi: 10.1097/00004583-199707000-00021] [Medline: 9204677]
54. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17. [doi: 10.2196/jmir.1367] [Medline: 21320854]
55. Danaher BG, Smolkowski K, Seeley JR, Severson HH. Mediators of a successful web-based smokeless tobacco cessation program. Addiction 2008 Oct;103(10):1706-1712 [FREE Full text] [doi: 10.1111/j.1360-0443.2008.02295.x] [Medline: 18715238]
56. Makai P, Perry M, Robben SH, Schers HJ, Heinen MM, Olde Rikkert MG, et al. Evaluation of an eHealth intervention in chronic care for frail older people: why adherence is the first target. J Med Internet Res 2014;16(6):e156 [FREE Full text] [doi: 10.2196/jmir.3057] [Medline: 24966146]
57. Christensen H, Griffiths K, Groves C, Korten A. Free range users and one hit wonders: community users of an internet-based cognitive behaviour therapy program. Aust N Z J Psychiatry 2006 Jan;40(1):59-62. [doi: 10.1080/j.1440-1614.2006.01743.x] [Medline: 16403040]

Abbreviations

eHealth: electronic health
IQR: interquartile range
mHealth: mobile health

Edited by G Eysenbach; submitted 05.07.18; peer-reviewed by S Schueller; comments to author 05.08.18; revised version received 16.09.18; accepted 22.10.18; published 14.12.18.

Please cite as:
Baumel A, Kane JM
Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset
J Med Internet Res 2018;20(12):e11491
URL: http://www.jmir.org/2018/12/e11491/
doi:10.2196/11491
PMID:30552077

Original Paper

Efficacy of a Parent-Based, Indicated Prevention for Anorexia Nervosa: Randomized Controlled Trial

Corinna Jacobi[1], PhD; Kristian Hütter[1], Dipl Psych; Ulrike Völker[2], PhD; Katharina Möbius[3], Dipl Psych; Robert Richter[4], Dipl Psych; Mickey Trockel[5], MD; Megan Jones Bell[6], PhD; James Lock[5], MD; C Barr Taylor[5], MD

[1]Klinische Psychologie & E-Mental Health, Institut für Klinische Psychologie und Psychotherapie, Universität Dresden, Dresden, Germany

[2]Klinik für Psychiatrie und Psychotherapie, Städtisches Klinikum Görlitz, Görlitz, Germany

[3]Klinik für Psychosomatische Medizin und Psychotherapie, Städtisches Klinikum Görlitz, Görlitz, Germany

[4]Sächsische Bildungsagentur, Regionalstelle Leipzig, Leipzig, Germany

[5]Department of Psychiatry, Stanford University School of Medicine, Stanford, CA, United States

[6]Headspace, Santa Monica, CA, United States

Corresponding Author:
Corinna Jacobi, PhD
Klinische Psychologie & E-Mental Health
Institut für Klinische Psychologie und Psychotherapie
Universität Dresden
Chemnitzer Strasse 46
Dresden, 01187
Germany
Phone: 49 35146338576
Fax: 49 351463372
Email: corinna.jacobi@tu-dresden.de

Abstract

Background: Web-based preventive interventions can reduce risk and incidence of bulimia and binge eating disorders among young high-risk women. However, their specific effects on core symptoms of anorexia nervosa (AN) are rather weak.

Objective: The primary objective of this study was to evaluate the efficacy of an indicated, parent-based, Web-based preventive program Eltern als Therapeuten (*E@T*) in reducing risk factors and symptoms of AN.

Methods: Girls aged between 11 and 17 years were screened by selected risk factors and early symptoms of AN. At-risk families were then randomized to E@T or an assessment-only control condition. Assessments took place at pre- and postintervention (6 weeks later) and at 6- and 12-month follow-up (FU).

Results: A total of 12,377 screening questionnaires were handed out in 86 German schools, and 3941 including consent returned. Overall, 477 (447/3941, 12.10%) girls were identified as at risk for AN and 256 of those could be contacted. In all, 66 families (66/256, 25.8% of those contacted) were randomized to the E@T or a wait-list control condition, 43 (43/66, 65%) participated in postassessments, and 27 (27/66, 41%) in 12-month FUs. Due to low participation and high dropout rates of parents, recruitment was terminated prematurely. At 12-month FU, girls' expected body weight (EBW) percentage was significantly greater for intervention participants compared with control participants (group by time interaction beta=21.0 [CI 5.81 to 36.13], *P*=.007; group by time squared interaction beta=−15.5 [CI −26.6 to −4.49], *P*=.007; estimated Cohen *d*=0.42]. No other significant effects were found on risk factors and attitudes of disturbed eating.

Conclusions: Despite a significant increase in girls' EBW percentage, parental participation and adherence to the intervention were low. Overall, parent-based, indicated prevention for children at risk for AN does not seem very promising, although it might be useful for parents who engage in the intervention.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 18614564; http://www.isrctn.com/ISRCTN18614564 (Archived by WebCite at http://www.webcitation.org/74FTV1EpF).

(J Med Internet Res 2018;20(12):e296) doi:10.2196/jmir.9464

KEYWORDS

anorexia nervosa; indicated prevention; internet; parental intervention; randomized controlled trial

Introduction

Background

Anorexia nervosa (AN) is a serious condition with a prevalence estimated between 0.3% and 0.7% among adolescent females based on Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition (DSM-IV) criteria [1-4]. AN can be accompanied by severe medical complications, including significant growth retardation, pubertal delay or interruption, and peak bone mass reduction [5,6]. Furthermore, the mortality rates associated with AN are significantly elevated when compared with standard population norms [7]. Approximately 60% of all eating disorder patients have a lifetime affective disorder [8], 35% of AN patients also suffer from obsessive-compulsive disorder, and there is a moderate overlap of AN and avoidant personality disorder. Furthermore, physical and psychological functioning and distress can be as severe in adolescents with AN regardless of presenting weight [9]. In addition, AN is associated with increased health care utilization and health care costs [10-12]. Krauth et al [12] reported yearly overall costs for AN of €195.4 million in Germany (€64.9 million through hospitalization, convalescence benefits, and rehabilitation as well as €130.5 million indirect costs through inability to work and premature death).

Given the seriousness of the disorder, the poor prognosis, and the associated burden and costs, early preventive interventions are of crucial importance. If these interventions target modifiable and potent risk factors, this could reduce both the onset of the disorder or mitigate core symptoms of the disorder before the onset. A few longitudinally assessed risk factors for eating disorders have been identified [13,14], but not all are suitable for preventive approaches, for example, pre- and perinatal risk factors have been confirmed in several studies but are not modifiable, and early childhood health and eating problems have also been confirmed in several studies but would not be suitable targets of preventive interventions for older children or young adults.

The factors weight and shape concerns and dieting, on the other hand, represent the most potent, modifiable, and confirmed risk factors for eating disorders in general. However, these factors are not specific for distinct eating disorder diagnoses, such as AN. In addition to these longitudinally assessed risk factors, a number of probable (retrospectively assessed) risk factors were found. At the start of the study, perfectionism and obsessive-compulsive symptoms seemed to be the best candidates as they both were modifiable and showed some specificity for AN [13-15].

Prior Work

Several previous reviews and meta-analyses suggest that preventive interventions for eating disorders (ED) in general reduce risk factors for, symptoms of, and—in few cases—even onset of mostly bulimic or binge eating-type ED [16-21]. Prevention programs reviewed in these analyses include the whole range of interventions from universal to indicated programs, school-based versus individually based, and are directed at age ranges from younger children to young adults. Effect sizes of the core risk factor outcomes range from low to high depending on the selection criteria applied and included samples, data analysis, and consideration of sensitivity analyses.

Although most of these meta-analyses included technology or Web-based interventions to some degree, 2 explicitly addressed effects of technology- or Web-based interventions only [16,20] with slightly different results. Beintner et al [16] conducted a cross-cultural comparison of 10 randomized controlled trials using the Web-based prevention program *Student Bodies* and found small to medium mean post and follow-up (FU) effects on drive for thinness, negative body image, and weight concerns. Loucas et al [20] found overall small post and FU effects on drive for thinness, weight and shape concern, and dietary restraint in 8 of the 13 Web-based prevention trials, including the Student Bodies intervention, and small or inconclusive effects for interventions in the remaining studies.

However, in the absence of confirmed risk factors for specific ED diagnoses, preventive interventions in general are usually not specifically directed at individuals at risk for specific ED diagnoses, such as AN but rather ED in general. This lack of diagnostic specificity of interventions is even more evident when only targeted interventions for individuals at risk are considered or moderators for intervention types are analyzed [18]. Participants in these studies are usually selected based on nonspecific risk factors for ED such as weight concerns, dieting, or body dissatisfaction. Early symptoms of specific ED categories (eg, subthreshold binge eating, compensatory behaviors, and body mass index [BMI]) are rarely used for selection. On the basis of the studies included in 1 meta-analysis [18], the mean BMI of young adult participants in these interventions was 23.3, and selection criteria did not include low BMI to determine risk status. It therefore seems likely that individuals at higher risk for AN were not reached by these interventions, and tailored preventive interventions for these individuals need to be developed.

Current treatment approaches for adults with AN have shown only limited effects [22], but there is considerable evidence supporting the effectiveness of family-oriented treatments [23-25] for adolescents. Family-based treatment (FBT [26]) has also been recommended by the American Psychiatric Association (APA [27]) and the National Institute for Health and Care Excellence (NICE) guidelines [28] as first-line treatment for adolescents with AN. A preventive approach, targeting risk factors and early symptoms for AN combined with elements of FBT, could therefore be beneficial in preventing the onset of the disorder in high-risk adolescents.

Thus, as part of a pilot study for a subsequent randomized controlled trial, we developed a family-based intervention called *Parents Act Now* targeting individuals at risk for AN. The 6-week intervention was directed at parents, originally developed in the United States, and subsequently translated into

German (*Eltern als Therapeuten or E@T*). The pilot study, conducted in parallel in the United States and Germany, examined the feasibility, acceptability, as well as short-term effects of the intervention in 46 adolescent females aged 11 to 17 years [29]. Overall, 11% of girls screened at the Germany site and 24% of girls screened at the US site met the risk criteria for AN. Parents accessed the majority of the Web-based sessions and rated the program favorably. At postassessment, we found a reduction in risk status for 16 out of the 19 participants. Participants remained stable or reported increased EBW percentage and decreased eating disorder attitudes and behaviors. However, the pilot study was also characterized by parents' rather low willingness to participate in and low compliance with the intervention. To address these problems, we made a number of changes to the intervention itself (eg, addressing potential denial and downplaying of eating problems in the first session) and the assessment procedure (eg, adding a motivational enhancement module to the first assessment where parents received feedback on the risk status of their daughters) before conducting the main study.

Goal of This Study

Following this pilot study, the major objective of this study was to determine the efficacy of the parent-based, Web-based, indicated preventive intervention *E@T* in comparison with a wait-list control group. We hypothesized that children of parents participating in the intervention would show an improvement in core AN symptoms, that is, weight loss, overvaluation of weight and shape, and restraint eating.

Methods

Design

We conducted a randomized controlled trial including parents and their daughters recruited from schools in Saxony, Germany. Eligible participants were randomized either to E@T or a wait-list control group. Assessments took place before randomization (at baseline, T1), at postintervention (6 weeks after baseline, T2), and at 6- and 12-month FU (8 [T3] and 14 [T4] months after baseline). Baseline, postintervention, and FU assessments were—with few exceptions—conducted in face-to-face settings. Both parental and child consents were required.

Participants

To be included in the study, girls had to be aged between 11 and 17 years and fulfill criteria of being at risk for AN based on the screening results. We defined at risk as a combination of factors selected from the following 3 categories [15]: (1) A: established risk factors for AN as high weight and shape concerns and drive for thinness (defined by either scoring ≥42 on the Weight Concerns Scale [30,31] or ≥24.1 on the Eating Disorder Inventory (EDI-2) subscale Drive for Thinness [32]), (2) B: early symptoms of AN indicated by low weight (defined as <90% EBW; Centers for Disease Control and Prevention, 2001) or significant weight loss (5% in the past 6 months), and (3) C: the presence of 1 out of the 4 probable risk factors, for example, high levels of perfectionism defined by scoring ≥78.0 on the Frost Multidimensional Perfectionism Scale [33], amenorrhea, excessive exercise, and a family history of an eating disorder. To be included, criterion B was mandatory and either criterion A or C (or both) was additionally required. In a previous study [15], the overall prevalence of the combination of these factors in a sample of 1562 adolescent girls was 10.8% and it increased from 9.5% to 16.5% between ages 11 and 16 years.

Exclusion criteria were the presence of a full-syndrome eating disorder in the past 6 months, current major depression, current substance abuse or dependence, and suicidal ideation.

Procedures

We asked the authorities of the school district of Saxony, Germany, for permission to conduct screenings in all high schools and secondary high schools throughout Saxony. Following their consent, 170 schools were individually invited to participate. Recruitment was completed in 86 schools (34 high schools and 52 secondary high schools) and followed a 2-step procedure. First, high-risk girls were identified through screens in participating schools after a short introduction of the study in class provided by trained research assistants. Questionnaires, including screening questions to be filled out by the children and few questions to be filled out by parents (daughters' current weight, height, and weight loss in the past 6 months, family history of eating disorders, internet access, and willingness to participate in an internet prevention program), were completed at home, and consent forms and questionnaires were collected approximately 1 week later in schools. If children screened positive, parents and children were invited for face-to-face baseline assessments. During these assessments, we conducted separate interviews with parents and children to assess children's eating and general pathology, to exclude ED diagnoses, and to obtain parental demographic information (education level, occupation, marital status, daughter's number and age of siblings, and daughter's type of school and current grade). Because the results of the pilot study suggested some problems with parental motivation, we included a manual-based motivational assessment and enhancement module (adapted from motivational interviewing) to guide interviewers' feedback to this first assessment. Children also filled out a number of self-report questionnaires. Children's height was measured to the nearest millimeter using a calibrated stadiometer and weight was measured to the nearest 0.1 kg using a digital scale.

Interview results were directly entered into a database (*MACRO*), which also contained algorithms to determine probable ED and other diagnoses and conduct randomization. Following separate interviews, 1 interviewer provided detailed verbal feedback on the daughter's risk factor status, eating disorder, and general pathology to parents and the daughter together and discussed inconsistent results from both interviews with them. Interviews were conducted by experienced graduate students and by research assistants who had received intensive training before conducting the interviews. To improve and maintain interview quality, all interviews were recorded and interviewers received verbal feedback on the recordings by an experienced graduate student. In a second step, parents of eligible children were randomly assigned to the E@T intervention or the assessment-only control group.

The study was approved by the local human subjects' committee (#EK172052010). Due to an oversight, the trial registration was delayed while participant recruitment had already started, following adaptations made to the intervention itself and the planned procedures. Participant recruitment and post and FU assessments took place between October 2010 and May 2014.

Intervention Eltern als Therapeuten (E@T)

The parental intervention is based on the first phase of the family-based treatment for AN by Lock [26], the parent guide by Lock and Le Grange [34], and an internet-based intervention to prevent eating disorders for adolescents [35]. The intervention *E@T* consists of a 6-session Web-based program for parents accessible over the course of 6 weeks and moderated by eating disorder experts (graduate-level clinical psychologists in training under supervision). The intervention also includes a moderated Web-based discussion group for parents, weekly monitoring journals related to their daughter's weight, eating and exercise with feedback provided by moderators, videos, and 2 phone calls to enable individualized feedback on the daughter's problems with eating, weight and shape concerns, and referral to other resources if necessary. Adolescents received a brief handout describing the purpose of the study written for a general audience in clear, lay terms, at a 6-grade reading level. A more detailed description of the intervention is summarized in a previous report [29].

Measures

Screening

The screening questionnaire consisted of 61 questions covering established risk factors, possible risk factors, and early symptoms of AN: weight and shape concerns based on the Weight Concerns Scale (WCS), a 5-item self-report screening questionnaire to identify students at risk for developing an ED [31]. Previous studies have shown that 10% of girls in the highest quartile of the WCS subsequently develop a subthreshold or full-syndrome ED. The German validation of the WCS [30] has a high test-retest reliability (r=.95).

In addition, drive for thinness was based on the respective 7-item subscale of the EDI-2 [32,36], self-reported height and weight, weight loss, and the presence of an ED in the past 6 months; perfectionism was based on the Frost Multidimensional Perfectionism Scale (MPS-F) [33,37]. The EDI-2 drive for thinness subscale has been shown to have high internal consistency (Cronbach alpha=.88). The MPS-F consists of 35 items covering 6 subscales (with 4 to 9 items each). The internal consistency for the subscales (Cronbach alpha) varies between .70 and .90 [33]. Questions related to the presence of secondary amenorrhea, excessive exercise, and family history of an ED in at least one family member were also included in the screen. Secondary amenorrhea was assessed by asking whether girls' menses had already started and, if so, whether they had missed menses in the past 3 months and if they took contraceptive medication. Amenorrhea was coded *yes* if menses had started but had been missed in the past 3 months or if menses had started but the use of contraceptives was endorsed. To endorse excessive exercise, girls had to indicate that they exercised in the past 4 weeks to lose weight, to influence body shape or body fat, to burn more calories, and to receive an average score of 3 (*sometimes*) or lower on a 1 (*always*) to 5 (*never*) scale asking if they were afraid of becoming upset or if they were feeling guilty when they had to skip exercise and if they exercised in spite of being sick or injured. These 4 questions were based on the Eating Disorder Examination (EDE) interview [38-40]. Parental history of ED was obtained from both girls and parents based on questions from the risk factor interview by Fairburn et al [41].

The primary outcomes were weight normalization (defined by change in EBW percentage, objectively measured) and other core AN symptoms, such as daughters' self-reported weight and shape concerns (assessed by the WCS), restraint, and frequency of driven exercise based on the EDE twelfth Edition [38-40]. The EDE is a semistructured interview that measures ED psychopathology on the 4 subscales: restraint (5 items), eating concern (5 items), weight concern (5 items), and shape concern (8 items), which can be aggregated to a total score and also generates ED diagnoses based on DSM-IV (text revision) criteria. The internal consistency (Cronbach alpha) varies between .73 and .86 for the subscales and is .93 for the total score. Secondary outcomes were EDE weight concern, EDE shape concern, EDE eating concern, EDI-2 drive for thinness, and EDI-2 body dissatisfaction. The latter EDI-2 subscale also has an internal consistency of .88.

Furthermore, the Schedule for Affective Disorders and Schizophrenia for School-Age Children [42] was used to assess present and past episodes of psychopathology based on DSM-IV Axis I criteria [43] in girls. At baseline, parents completed the Parent Motivation Inventory (PMI; [44]) to assess their motivation and confidence to address their daughter's eating problem. The PMI is a 25-item scale with a high internal consistency of .96 (Cronbach alpha) [44].

At postintervention and 6- and 12-month FU, only the EDE interview was conducted with parents on the daughter, and with daughters themselves; in addition, daughters' height and weight were measured and daughters filled out the WCS and EDI-2 subscales drive for thinness, bulimia, and body dissatisfaction and answered questions regarding treatment utilization. Whenever possible, subjective reasons for parents declining participation in the study after initial contact and feedback on daughters' risk status were assessed qualitatively.

We assessed adherence to the intervention by mean number of sessions opened and overall percentage of program pages opened. These data were retrieved from the program log files.

Randomization and Masking

The randomization algorithm, which was integrated into the database MACRO, was provided by the independent Centre for Clinical Trials (Koordinierungszentrum für Klinische Studien) at TU Dresden. Children were stratified by age and EBW percentage and randomized in a ratio of 1:1 to E@T or the waiting list control condition after parents had given informed consent. Parents and psychologists involved in the moderation of the E@T program could not be masked to intervention allocation. Assessors who conducted T1 to T4 diagnostic assessments could not be blinded to the intervention condition

but were neither involved in the moderation of the intervention nor in final data analyses.

Statistical Analyses

The initial study sample size estimate was based on a power calculation for a single end-point group comparison at 12 months. Therefore, the sample size calculation was based on a 2-tailed *t* test for independent samples. For this calculation, the estimated between-group effect size at 12-month FU was *d*=0.5 between the intervention and the control group for the primary outcomes: EBW percentage and other core AN symptoms. Using these parameters, the estimated sample size needed to achieve 80% power was 64 participants per group (or a total of 128 subjects). Considering a combined noncompliance and loss to FU rate of 30%, an estimated 91 participants per group would have been required to attain an adequate sample size for this study. We subsequently employed a mixed-effects model approach to assess all group differences over time. Mixed-effects models use all available data points when participants are lost to FU. This approach is likely to render higher statistical power than a *t* test used to assess group differences at a single time point. Therefore, the a priori power analysis conducted for this study (based on an independent samples *t* test) may have overestimated the sample size needed for this study.

All analyses were conducted as intention-to-treat (ITT) analyses including all randomized participants. Differences between the intervention and the control group on primary outcomes (EBW percentage, weight and shape concerns, EDE restraint, and driven exercise) as well as secondary outcomes were tested using mixed-effects models to account for the nested data structure of 4 observations across time within individual participants [45]. Total observation time was set at 1.0. Each measurement time point was set at its corresponding fraction of 1.0. This 0 to 1 time variable was then multiplied by itself to create a variable for time squared, which enables specification of quadratic regression models assessing intervention effects on change and on rate of change in outcome variables. A group by time interaction term was specified as to estimate the effect of the intervention. A group by time squared interaction term was also specified to estimate the effect of intervention on rate of change using a quadratic regression model. Cohen *d* was calculated by dividing the mixed-effects model derived intervention effect estimate by the pooled SD of the particular measure at baseline.

Differences on screening variables between participating and nonparticipating parents and children were analyzed using 2-tailed *t* test for continuous screening variables and chi-square test for dichotomous variables. These exploratory analyses were not corrected for multiple comparisons.

All analyses were performed using the software programs SPSS 22 (Statistical Package for the Social Sciences) and HLM7 (Hierarchical Linear and Nonlinear Modeling).

Results

Recruitment

Between October 2010 and December 2012, a total of 12,377 screening questionnaires were handed out in schools, 4416 (4416/12,337, 35.79%) returned and 3941 (3941/12,377, 31.84%) included consent. At baseline, 99 interviews were conducted with parents and screen-positive children. In total, 33 families could not be enrolled because they did not fulfill inclusion criteria any more during these assessments (mostly because children's EBW percentage was in the normal range when measured objectively) or refused randomization. Finally, 66 families were randomized, 32 to the E@T intervention and 34 to the control condition. At 12-month FU, the dropout rate in the E@T condition was 65.6% and in the control condition was 52.9%, and they were not significantly different. Figure 1 presents a Consolidated Standards of Reporting Trials flow diagram.

Due to a much lower response rate of parents of girls screened for risk status and the high rate of parents of at-risk girls not willing to participate, it would have taken at least twice as long to recruit the originally planned sample size. The funding of the study was therefore stopped before the originally planned sample size had been achieved, but FUs of all randomized parents were still included.

Sample

Of participants with informed consent, 12.10% (477/3941) met predefined criteria for at risk. Overall, 47.8% of the sample fulfilled the combination of criteria A (high WCS or EDI drive for thinness), B (low weight or significant weight loss), and C (high levels of perfectionism, amenorrhea, excessive exercise, or a family history of an ED); 28.4% fulfilled the combination of criteria B and C; and 23.9% fulfilled the combination of criteria A and B. Regarding criterion B, 47% of the sample endorsed low body weight and 53% of the sample endorsed significant weight loss in the past 6 months.

Table 1 presents baseline characteristics of the sample. Included girls were on average about 14 years old, and the average EBW percentage was in the normal range. Overall, girls showed only few ED symptoms in the 4 weeks before baseline. Current or past comorbid major depressive disorder, separation anxiety disorder, social phobia, and specific phobia were also low.

Figure 1. Consolidated Standards of Reporting Trials diagram of participant flow. E@T: Eltern als Therapeuten.

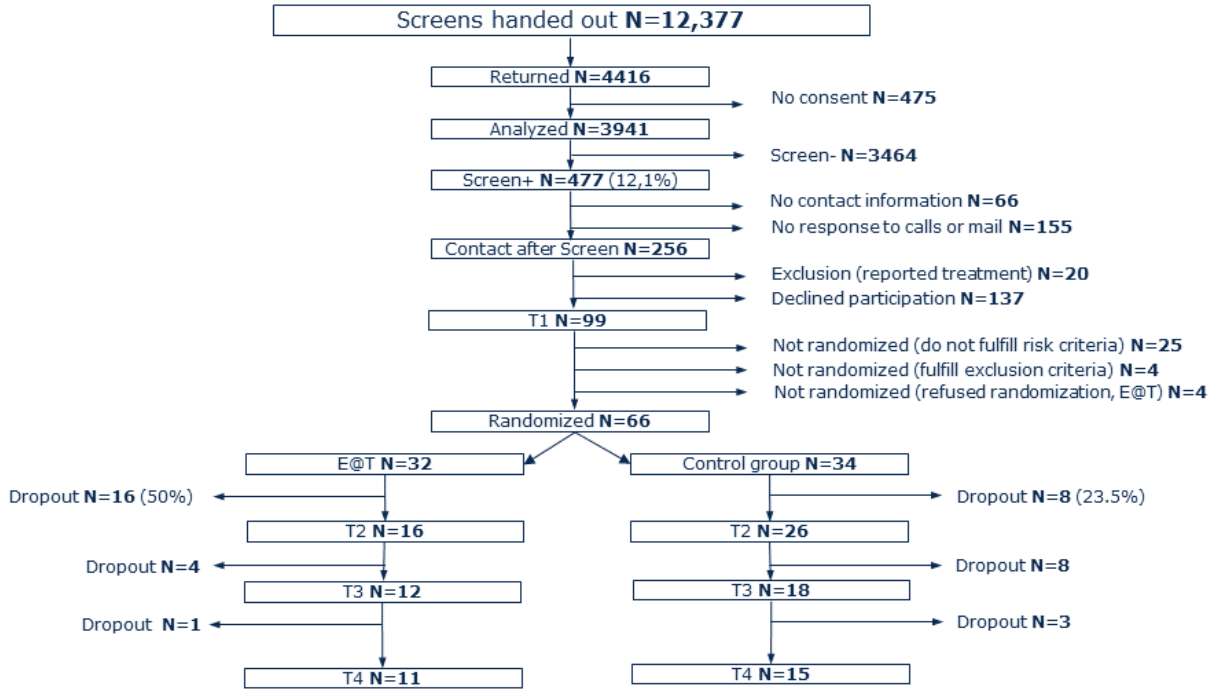


Table 1. Sample characteristics of all randomized children at baseline.

Characteristics	E@T[a] (N=32)	Control group (N=34)
Age in years, mean (SD)[b]	13.8 (1.5)	13.7 (1.6)
Percentage of expected body weight, mean (SD)	98.8 (12.3)	99.1 (13.4)
Objective binge episodes (past month), mean (SD)[b]	0 (0)	0 (0)
Subjective binge episodes (past month), mean (SD)[b]	0.8 (4.4)	0.1 (0.3)
Fasting (days past month), mean (SD)[b]	0.1 (0.3)	0.3 (1.2)
Vomiting (episodes past month), mean (SD)[b]	0 (0)	0.3 (1.7)
Laxative use (episodes past month), mean (SD)[b]	0 (0)	0.1 (0.5)
Excessive exercise (days past month), mean (SD)[b]	4.7 (8.6)	5.1 (9.1)
Comorbidity, n (%)[c]		
Separation anxiety disorder	0 (0)	3 (9)
History of major depression	2 (6)	2 (6)
Social phobia	0 (0)	1 (3)
Specific phobia	1 (3)	0 (0)

[a]E@T: Eltern als Therapeuten.

[b]According to the Eating Disorder Examination [38-40].

[c]According to Schedule for Affective Disorders and Schizophrenia for School-Age Children [42].

Screening Differences Between Participating and Nonparticipating Parents and Children

We found no significant differences in the frequencies of endorsing screening criteria A and C between parents who participated in the study and those who refused to participate. However, children of participating parents endorsed significantly more of all 3 screening criteria (52.1% vs 37.2%; *P*=.02) and showed significantly higher levels of weight concerns (WCS; mean= 47.9 vs mean=40.5; *P*=.03) compared with children of nonparticipating parents. This might be indicative of higher levels of impairment of children of participating parents. In addition, when parent-reported and daughters' self-reported weight loss was compared, the discrepancy between the 2 estimates was significantly smaller for nonparticipating parents compared with participating parents (1.2 vs 2.15 kg; *P*=.03).

Intervention Adherence and Acceptance

On average, intervention group parents opened 28% of program pages (median 16%), 2.7 of 6 sessions (median 2.0), and logged on to the program 3.4 times (median 3.0; range 0-11). In total, 29% of randomized parents never logged on to the program at all and only 16% opened more than 75% of program pages. However, participating intervention group parents overall rated the program quite favorably as *good* (mean=2.2; SD=0.94; range=1-5; scale from 1 [*very good*] to 6 [*very poor*]), rated the program content and group moderation on average between *good* and *very good* (means 1.8 and 1.7, respectively), and reported they would *very much* recommend the program to other parents of at-risk children (mean=3.60; SD=0.74; range=2-4; scale scores from 1 [*not at all*] to 4 [*very much*]).

Primary and Secondary Outcomes

Results of ITT analyses of primary and secondary outcomes based on daughters' self-report are summarized in Table 2. Of the primary outcomes, only 1 significant difference between the intervention and the control group was found: between preintervention and 12-month FU, girls of the intervention group gained significantly more and faster weight as indicated by change in percentage of EBW compared with girls in the control group. There was a significant time-squared by group interaction indicating the effect of intervention on EBW percentage was curvilinear. The greatest effect of intervention on EBW percentage occurred early during the observation period. The total effect of intervention on EBW percentage can be estimated by adding the estimated group by time interaction effect with the group by time squared interaction effect (21.0–15.5=5.5%), which estimates a 5.5% greater increase in EBW percentage in intervention group participants compared with control group participants. The effect size (*d=0*.42) is in the small to medium range. No other significant differences were found between groups on child- and parent-reported secondary outcomes. In both groups, no new onset full-syndrome DSM-IV diagnoses of AN were observed over time.

Multimedia Appendices 1 and 2 provide means and SDs for primary and secondary outcomes for both groups (based on daughters' self-report and parental report) at all assessment points.

Parent-Reported Reasons for Unwillingness to Participate

Whenever possible, we asked the parents who declined participation in the study after being told that their daughters had screened positive to give reasons for their unwillingness to participate (Table 3). The majority of these parents responded that they did not perceive the identified risk factors and early symptoms in their daughters as problematic and, accordingly, participation in a preventive intervention as necessary or useful. Other frequently reported reasons were lack of time and daughter's own unwillingness to participate. A relatively large proportion of parents also reported *all-clear* given by the pediatrician (ie, the pediatrician did not consider the daughter's weight loss problematic or explicitly advised parents not to participate in the study). In a considerable proportion of cases, parents also reported a change in measures included to define risk status in the screening (eg, a weight gain after screening) or revised the previously reported screening criteria (eg, family history of ED). Some parents, however, also seemed to be afraid to worsen the current condition of their daughter ("let sleeping dogs lie") by getting engaged in the intervention or reported too many other current problems to get further engaged. A small proportion of children were reported to be already in treatment because of eating or other mental health problems.

Table 2. Intervention effects on outcome variables estimated with mixed-effects models.

Effect	Group*time (95% CI)	*t* ratio	*P* value	Cohen *d*
Percentage of expected body weight	21.0 (5.81 to 36.13)	2.76	.007	0.42[a]
Group*time squared effect	−15.5 (−26.6 to −4.49)	−2.81	.007	0.42[a]
Excessive exercise	0.82 (−3.97 to 5.62)	0.34	.73	0.09
Weight Concerns Scale	2.02 (−8.03 to 12.08)	0.40	.69	0.08
EDI-2[b] bulimia	2.32 (−1.95 to 6.61)	1.09	.28	0.29
EDI-2 drive for thinness	1.86 (−1.05 to 4.77)	1.28	.21	0.27
EDI-2 body dissatisfaction	2.77 (−0.36 to 5.89)	1.76	.08	0.34
EDE[c] total score	0.04 (−0.41 to 0.5)	0.19	.85	0.04
EDE dietary destraint	−0.04 (−0.63 to 0.55)	−0.13	.90	−0.03
EDE eating concern	0.12 (−0.53 to 0.77)	0.38	.71	0.13
EDE weight concern	0.03 (−0.72 to 0.77)	0.07	.94	0.02
EDE shape concern	−0.03 (−0.73 to 0.67)	−0.09	.93	−0.02

[a]Estimated Cohen *d* for percentage of expected body weight is the sum of the standardized effects for group by time plus group by time squared.

[b]EDI-2: Eating Disorder Inventory.

[c]EDE: Eating Disorder Examination.

Table 3. Parent-reported reasons for declining participation (N=137 parents; multiple answers possible).

Parent-reported reasons for declining participation	Endorsements, n
Do not see risk factors and symptoms as problematic	89
Pediatrician does not see a problem or does not recommend study participation	18
Lack of time	16
Interview canceled, not attended, no response, or no reason given	30
Daughter declines participation	13
Change in risk status since screening (weight gain, exercise, and family history of eating disorder)	15
Too many other problems	5
Afraid to raise awareness for eating disorder problems	5
Currently in treatment for eating disorder or other mental health problem	3

Discussion

Principal Findings

The objective of this study was to evaluate the efficacy of a parent-based, targeted preventive intervention for children and adolescents at risk for AN compared with an assessment-only control group. The intervention was specifically developed to target early symptoms and potential risk factors for AN that distinguish E@T from other preventive interventions for ED. It also incorporated elements from the current most promising treatment approach for adolescents with AN, that is, family-based treatment. This trial was preceded by a pilot study with overall encouraging results conducted in both the United States and Germany. We found that—over the course of the study and at the 12-month FU and based on ITT analyses—at-risk girls, whose parents had participated in E@T, gained significantly more and faster weight based on change in percentage of EBW compared with girls in the control group. Although the effect size of this change is in the small to medium range, previous Web-based prevention trials for ED in general usually do not find differences in BMI [17,46,47]. In addition, low weight was one of the risk factors or early symptoms we hoped to change through the intervention. However, these results must be considered in the context that few parents were willing to enroll and engage in the study and no other significant effects on primary or secondary outcomes were found in the ITT analyses. In interpreting the results, it is also important to note that means of outcome measures included in the screening (ie, WCS and EDI drive for thinness) dropped between screening and preintervention assessment. Furthermore, ED and weight-related measures improved in participants of both groups who completed postintervention and FU measures, which limit the potential to see differences. The reasons for the improvement in the control groups are not known but might be indicative of regression to the mean effects. Participants underwent detailed ED interviews after initial screening over the course of more than a year, which in itself may have raised awareness for risk factors and symptoms in parents and may have contributed to improvements in both groups.

Limitations

The results of this study need to be discussed in the context of the following limitations: (1) small sample size because of low screening completion rates possibly resulting in too little power to establish efficacy; (2) low rates of eligible participants agreeing to participate; (3) low parental engagement in the intervention; and (4) high dropout rates, which again may have affected power, randomization and, thus, also conclusions drawn from the analysis. Parents' (low) willingness to partake in a Web-based intervention aimed at reducing their daughter's risk of AN was the kernel of this study and warrants further exploration. Of screens distributed in schools, 35.7% were returned. At the beginning of recruitment, this rate was 20% but was increased by a number of strategies (eg, letter of recommendation of school authorities directed at individual schools, increasing awareness for the study by increased press releases, and offering incentives to girls). Furthermore, even for girls identified as being at risk, parental willingness to participate in the study was low. Only about half of identified families provided contact information, and of those contacted, only about 16% could be randomized. Although parents receiving the intervention, on average, rated the program favorably, they accessed less than a third of all program pages and less than half of the sessions. Using a standardized measure of engagement might have provided information to explain this discrepancy. Adherence is a well-known problem for Web-based interventions in general [48]. However, compared with targeted preventive interventions for ED, in which adherence usually ranges between 50% and 80% [49], adherence to E@T was clearly lower. Along with low engagement rates in the intervention, the study was also characterized by high dropout rates, that is, over 50% in the control group and 65.6% in the intervention group. These rates exceed dropout rates of both targeted intervention trials for ED in general [46,47] and of those reported for family-based treatment trials for AN, which average between 15% and 25% [50-52].

Comparison With Prior Work

In the absence of specific, parent-based prevention trials for girls at risk for AN, we can only compare our results with more general, parent-based preventive studies. For example, compared with parents referred to outpatient treatment for child conduct problems in the validation sample of the PMI [44], parental motivation in this preventive trial was much lower. As included children had not already developed a mental health problem requiring treatment, parents may have been more reluctant to engage in the intervention. In our pilot study [29], we also found

a positive correlation between daughter's risk status and parental engagement: at the US site, parents of children who already met criteria for AN showed higher levels of engagement with the intervention than those of children at risk for AN. A recent review of interventions involving parents that aim to prevent body dissatisfaction or eating disorders [53] identified 20 studies, 12 of which presented data on the effects of involving parents in prevention programs. A quarter of these studies revealed significant problems with parental recruitment and motivation, despite daughters being screened at-risk [29,54,55]. Although Hart et al [53] concluded that preventive interventions involving parents may have some benefit, they also expressed concern over the finding *that measuring and communicating a child's at-risk status does not appear to improve parent engagement with prevention programs*.

On the other hand, even with pediatric long-term medical conditions, such as asthma, cystic fibrosis, HIV, diabetes [56], or life-threatening conditions requiring pediatric organ transplantation [57], parents' and caregivers' nonadherence to prescribed treatments is reported to be a common problem. Low adherence or denial may therefore represent a more general problem of parents when confronted with chronic or potentially threatening health conditions of their child.

Given the low parental engagement, participation, and completion rates, this sample is likely to be biased toward parents who are more willing to respond to a perceived risk for AN in their daughter (eg, higher motivated parents or parents more willing to acknowledge these risks). Thus, the observed intervention effect on percentage of EBW likely applies to this group of parents. This interpretation is supported in part by the reasons parents gave for unwillingness to participate that we gathered from parents that could be contacted. The majority of these parents did not consider the identified risk factors and early symptoms in their daughters as severe enough to get engaged or do so despite their daughter's refusal to participate. As included children were, on average, at the time of the preintervention assessment not markedly underweight, parents of these normal-weight children fulfilling the weight loss criterion may have not perceived other risk factors, such as increased weight concerns, as problematic.

The comparison of parents willing to participate and those refusing to participate, on the other hand, shows that daughters of participating parents had even higher levels of ED-related impairment and parents unwilling to participate may underestimate their daughters' weight loss. Thus, although parents of daughters with higher levels of ED risk factors and symptoms were more willing to participate in the study, symptoms of AN may have needed to be even more pronounced for most parents to engage at all or to engage more consistently in a preventive intervention. Alternatively, given the insidious course AN onset can take, parents may have needed more time and further evidence to realize and accept these risk factors and symptoms to motivate their engagement.

A recent systematic review [58] suggested 6 categories of reasons for parents' and caregivers' nonadherence to prescribed treatments in pediatric long-term medical conditions, including concerns or fears of the condition or the recommended treatment, difficulty following the treatment regimen, children's resistance to treatment, perceived threats and strains to family relationships, parental priorities to preserve *normal life*, and (negative) input from and relationship with health professionals. Some of these reasons may explain parents' low engagement and adherence in this study. Future studies, therefore, should address these potential barriers to engagement and parental level of readiness to engage more explicitly.

Conclusions

In conclusion, the intervention showed small effects on only 1 outcome, and given the few parents who were willing to enroll and engage in the study, the intervention does not have the potential for wide-scale acceptance as we hoped for. It may be more beneficial for parents willing to face their daughter's initial problems, when these have become more pronounced or as the first step for parents of children with full syndrome AN before getting engaged in outpatient or inpatient treatment. However, this will need to be demonstrated in subsequent studies. The parent-based intervention did show some promise for the subsample of children of parents willing to engage in the assessments and in the intervention, even when only administered in a relatively small dose. Together with detailed interviews and feedback on ED risk factors and symptoms, children at risk may benefit from the intervention. Next steps in developing a population-based intervention targeting parents with children at risk for AN would be to (1) consider reasons why parents did not find the identified risk factors compelling, (2) develop better and more effective ways to convey information about risk factors, and (3) identify strategies to resolve parents' concerns with engagement. Preventive interventions for ED may generally need to educate parents more explicitly about the potential dangers of early signs of disordered eating, such as dieting or weight loss in a child. Finally, a strategy for making the intervention more readily accessible is needed; although we had hoped to offer the original intervention as part of the curriculum in the participating schools, we were ultimately not allowed to do so. By offering the program in a systematic normative manner to parents in the school setting, potential avoidance and stigma, which likely interfered with engagement in the intervention, might be reduced.

Acknowledgments

The study was funded by a grant by the Schweizerische Anorexia Nervosa Stiftung (Project No 02-09 B). The authors would like to thank all parents and children who participated in this study for their contributions. Authors would also like to thank Dr Ina Beintner for support with the data analyses. The authors acknowledge support by the Open Access Publishing Funds of the SLUB/TU Dresden. The authors would like to thank the following student assessors and research assistants: Dagmar Uhlmann, Francie Kriegel, Juliane Penner, Sören Paul, Katharina Abel, Nicole Müller, Marco Holzmann, Melanie Hausdorf, Paula von

Bloh, Denise Kain, Carolin Müller, Jana Bittn er, Julia Fleischhack, Katharina Wick, Claudia Lange, Constanze Nicklisch, Katharina Liebscher, Anne-Kathrin Reinhardt, Anna Wagner, Johannes Sperling, Claudia Recksiedler, Lotta Maaß, Linda Eberlein, Monique Lippmann, Maria Heinig, Janka Lubinova, Anne Ueck, Christin Partzsch, and Dominique Schwartze.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Primary and Secondary Outcomes (Daughters` Self-Report).

[PDF File (Adobe PDF File), 30KB - jmir_v20i12e296_app1.pdf]

Multimedia Appendix 2

Secondary Outcomes (Parent-Reported Outcomes).

[PDF File (Adobe PDF File), 22KB - jmir_v20i12e296_app2.pdf]

References

1. Hoek HW, van Hoeken D. Review of the prevalence and incidence of eating disorders. Int J Eat Disord 2003 Dec;34(4):383-396. [doi: 10.1002/eat.10222] [Medline: 14566926]
2. Hoek HW, van Harten PN, Hermans KME, Katzman MA, Matroos GE, Susser ES. The incidence of anorexia nervosa on Curaçao. Am J Psychiatry 2005 Apr;162(4):748-752. [doi: 10.1176/appi.ajp.162.4.748] [Medline: 15800148]
3. van Son GE, van Hoeken D, Bartelds AIM, van Furth EF, Hoek HW. Time trends in the incidence of eating disorders: a primary care study in the Netherlands. Int J Eat Disord 2006 Nov;39(7):565-569. [doi: 10.1002/eat.20316] [Medline: 16791852]
4. Zipfel S, Löwe B, Reas DL, Deter HC, Herzog W. Long-term prognosis in anorexia nervosa: lessons from a 21-year follow-up study. Lancet 2000 Feb 26;355(9205):721-722. [doi: 10.1016/S0140-6736(99)05363-5] [Medline: 10703806]
5. Golden NH, Katzman DK, Kreipe RE, Stevens SL, Sawyer SM, Rees J, et al. Eating disorders in adolescents. J Adolesc Health 2003 Dec;33(6):496-503. [doi: 10.1016/j.jadohealth.2003.08.004]
6. Rome ES, Ammerman S, Rosen DS, Keller RJ, Lock J, Mammel KA, et al. Children and adolescents with eating disorders: the state of the art. Pediatrics 2003 Jan;111(1):e98-108. [Medline: 12509603]
7. Rosling AM, Sparén P, Norring C, von Knorring AL. Mortality of eating disorders: a follow-up study of treatment in a specialist unit 1974-2000. Int J Eat Disord 2011 May;44(4):304-310. [doi: 10.1002/eat.20827] [Medline: 21472749]
8. Herzog DB, Nussbaum KM, Marmor AK. Comorbidity and outcome in eating disorders. Psychiatr Clin North Am 1996 Dec;19(4):843-859. [Medline: 9045226]
9. Sawyer SM, Whitelaw M, Le Grange D, Yeo M, Hughes EK. Physical and psychological morbidity in adolescents with atypical anorexia nervosa. Pediatrics 2016 Apr;137(4):- [FREE Full text] [doi: 10.1542/peds.2015-4080] [Medline: 27025958]
10. Agras WS. The consequences and costs of the eating disorders. Psychiatr Clin North Am 2001 Jun;24(2):371-379. [Medline: 11416936]
11. Ágh T, Kovács G, Supina D, Pawaskar M, Herman BK, Vokó Z, et al. A systematic review of the health-related quality of life and economic burdens of anorexia nervosa, bulimia nervosa, and binge eating disorder. Eat Weight Disord 2016 Sep;21(3):353-364 [FREE Full text] [doi: 10.1007/s40519-016-0264-x] [Medline: 26942768]
12. Krauth C, Buser K, Vogel H. How high are the costs of eating disorders - anorexia nervosa and bulimia nervosa - for German society? Eur J Health Econ 2002;3(4):244-250. [doi: 10.1007/s10198-002-0137-2] [Medline: 15609150]
13. Jacobi C, Hayward C, de Zwaan M, Kraemer HC, Agras WS. Coming to terms with risk factors for eating disorders: application of risk terminology and suggestions for a general taxonomy. Psychol Bull 2004 Jan;130(1):19-65. [doi: 10.1037/0033-2909.130.1.19] [Medline: 14717649]
14. Jacobi C, Hütter K, Fittig E. Psychosocial risk factors for eating disorders. In: Agras WS, Robinson A, editors. The Oxford Handbook of Eating Disorders. New York: Oxford University Press; 2018.
15. Völker U, Jacobi C, Jones M, Lock J, Taylor CB. Potential risk factors and early symptoms of anorexia nervosa: prevalence in 11–16-year-old girls. Advances in Eating Disorders 2013 Jul 02;2(1):19-30. [doi: 10.1080/21662630.2013.794521]
16. Beintner I, Jacobi C, Taylor CB. Effects of an Internet-based prevention programme for eating disorders in the USA and Germany--a meta-analytic review. Eur Eat Disord Rev 2012 Jan;20(1):1-8. [doi: 10.1002/erv.1130] [Medline: 21796737]
17. Le LK, Barendregt JJ, Hay P, Mihalopoulos C. Prevention of eating disorders: A systematic review and meta-analysis. Clin Psychol Rev 2017 Apr;53:46-58. [doi: 10.1016/j.cpr.2017.02.001] [Medline: 28214633]
18. Stice E, Shaw H, Marti CN. A meta-analytic review of eating disorder prevention programs: encouraging findings. Annu Rev Clin Psychol 2007;3:207-231. [doi: 10.1146/annurev.clinpsy.3.022806.091447] [Medline: 17716054]

19. Watson HJ, Joyce T, French E, Willan V, Kane RT, Tanner-Smith EE, et al. Prevention of eating disorders: A systematic review of randomized, controlled trials. Int J Eat Disord 2016 Sep;49(9):833-862. [doi: 10.1002/eat.22577] [Medline: 27425572]
20. Loucas CE, Fairburn CG, Whittington C, Pennant ME, Stockton S, Kendall T. E-therapy in the treatment and prevention of eating disorders: A systematic review and meta-analysis. Behav Res Ther 2014 Dec;63:122-131 [FREE Full text] [doi: 10.1016/j.brat.2014.09.011] [Medline: 25461787]
21. Pratt BM, Woolfenden SR. Interventions for preventing eating disorders in children and adolescents. Cochrane Database Syst Rev 2002(2):CD002891. [doi: 10.1002/14651858.CD002891] [Medline: 12076457]
22. Hay PJ, Claudino AM, Touyz S, Abd Elbaky G. Individual psychological therapy in the outpatient treatment of adults with anorexia nervosa. Cochrane Database Syst Rev 2015 Jul 27(7):CD003909. [doi: 10.1002/14651858.CD003909.pub2] [Medline: 26212713]
23. Couturier J, Kimber M, Szatmari P. Efficacy of family-based treatment for adolescents with eating disorders: a systematic review and meta-analysis. Int J Eat Disord 2013 Jan;46(1):3-11. [doi: 10.1002/eat.22042] [Medline: 22821753]
24. Fisher CA, Hetrick SE, Rushford N. Family therapy for anorexia nervosa. Cochrane Database Syst Rev 2010 Apr 14(4):CD004780. [doi: 10.1002/14651858.CD004780.pub2] [Medline: 20393940]
25. Lock J. An Update on Evidence-Based Psychosocial Treatments for Eating Disorders in Children and Adolescents. J Clin Child Adolesc Psychol 2015;44(5):707-721. [doi: 10.1080/15374416.2014.971458] [Medline: 25580937]
26. Lock J, Le Grange D. Treatment manual for anorexia nervosa: a family-based approach. New York: Guilford Press; 2001.
27. American Psychiatric Association. Practice Guideline for the Treatment of Patients with Eating Disorders. Arlington, VA: American Psychiatric Association; 2006.
28. National Institute for Health and Excellence. Clinical Guidelines. Eating Disorders: Recognition and Treatment. London: National Institute for Health and Care Excellence (UK); 2017.
29. Jones M, Völker U, Lock J, Taylor CB, Jacobi C. Family-based early intervention for anorexia nervosa. Eur Eat Disord Rev 2012 May;20(3):e137-e143. [doi: 10.1002/erv.2167] [Medline: 22438094]
30. Grund K. Validierung der Weight Concerns Scale zur Erfassung von Essstörungen Dissertation/Thesis, Unpublished. Trier: Universität Trier; 2003.
31. Jacobi C, Abascal L, Taylor CB. Screening for eating disorders and high-risk behavior: caution. Int J Eat Disord 2004 Nov;36(3):280-295. [doi: 10.1002/eat.20048] [Medline: 15478137]
32. Paul T, Thiel A. Eating Disorder Inventory-2. Deutsche Version. Göttingen: Hogrefe; 2004.
33. Altstötter-Gleich C, Bergemann N. Testgüte einer deutschsprachigen Version der Mehrdimensionalen Perfektionismus Skala von Frost, Marten, Lahart und Rosenblate (MPS-F). Diagnostica 2006 Jul;52(3):105-118. [doi: 10.1026/0012-1924.52.3.105]
34. Lock J, Le Grange D. Help Your Teenager Beat an Eating Disorder. New York: Guilford Press; 2006.
35. Jacobi C, Morris L, Beckers C, Bronisch-Holtze J, Winter J, Winzelberg AJ, et al. Maintenance of internet-based prevention: a randomized controlled trial. Int J Eat Disord 2007 Mar;40(2):114-119. [doi: 10.1002/eat.20344] [Medline: 17080447]
36. Garner DM. EDI-2. Eating Disorder Inventory-2. Professional Manual. 2nd ed. Odessa, Florida: Psychological Assessment Resources; 1991.
37. Frost RO, Marten P, Lahart C, Rosenblate R. The dimensions of perfectionism. Cogn Ther Res 1990 Oct;14(5):449-468. [doi: 10.1007/Bf01172967]
38. Fairburn C, Cooper Z. The Eating Disorder Examination. In: Fairburn CG, T. WG. editors. Binge Eating: Nature, Assessment and Treatment. 12 ed. New York: The Guildford Press; 1993:317-360.
39. Hilbert A, Tuschen-Caffier B. Eating Disorder Examination: Deutschsprachige Übersetzung. Münster: Verlag für Psychotherapie; 2006.
40. Hilbert A, Tuschen-Caffier B, Ohms M. Eating Disorder Examination: Deutschsprachige Version des strukturierten Essstörungsinterviews. Diagnostica 2004:98-106 [FREE Full text]
41. Fairburn CG, Welch SL, Doll HA, Davies BA, O'Connor ME. Risk factors for bulimia nervosa. A community-based case-control study. Arch Gen Psychiatry 1997 Jun;54(6):509-517. [Medline: 9193191]
42. Delmo C, Weiffenbach O, Gabriel M, Stadler C, Poustka F. Adhs-essen. Frankfurt am Main: Klinik für Psychiatrie und Psychotherapie des Kindes- und Jugendalters; 2001. Diagnostisches Interview Kiddie-Sads-Present and Lifetime Version (K-SADS-PL). Auflage der deutschen Forschungsversion URL: http://www.adhs-essen.com/PDF/K-SADS_Fragebogen.pdf [accessed 2017-12-04] [WebCite Cache ID 6vSMbr37m]
43. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (4th ed.). Washington, DC: Author; 1994.
44. Nock MK, Photos V. Parent Motivation to Participate in Treatment: Assessment and Prediction of Subsequent Participation. J Child Fam Stud 2006 May 3;15(3):333-346. [doi: 10.1007/s10826-006-9022-4]
45. Raudenbush S, Bryk A. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks: Sage Publications Inc; 2002.
46. Jacobi C, Völker U, Trockel MT, Taylor CB. Effects of an Internet-based intervention for subthreshold eating disorders: a randomized controlled trial. Behav Res Ther 2012 Feb;50(2):93-99. [doi: 10.1016/j.brat.2011.09.013] [Medline: 22137366]

47. Taylor CB, Kass AE, Trockel M, Cunning D, Weisman H, Bailey J, et al. Reducing eating disorder onset in a very high risk sample with significant comorbid depression: A randomized controlled trial. J Consult Clin Psychol 2016 May;84(5):402-414 [FREE Full text] [doi: 10.1037/ccp0000077] [Medline: 26795936]
48. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
49. Beintner I, Jacobi C, Taylor CB. Participant adherence to the Internet-based prevention program StudentBodies™ for eating disorders — A review. Internet Interventions 2014 Mar;1(1):26-32. [doi: 10.1016/j.invent.2014.03.001]
50. Agras WS, Lock J, Brandt H, Bryson SW, Dodge E, Halmi KA, et al. Comparison of 2 family therapies for adolescent anorexia nervosa: a randomized parallel trial. JAMA Psychiatry 2014 Nov;71(11):1279-1286. [doi: 10.1001/jamapsychiatry.2014.1025] [Medline: 25250660]
51. Lock J, Agras WS, Bryson S, Kraemer HC. A comparison of short- and long-term family therapy for adolescent anorexia nervosa. J Am Acad Child Adolesc Psychiatry 2005 Jul;44(7):632-639. [doi: 10.1097/01.chi.0000161647.82775.0a] [Medline: 15968231]
52. Lock J, Le Grange D, Agras WS, Moye A, Bryson SW, Jo B. Randomized clinical trial comparing family-based treatment with adolescent-focused individual therapy for adolescents with anorexia nervosa. Arch Gen Psychiatry 2010 Oct;67(10):1025-1032 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2010.128] [Medline: 20921118]
53. Hart LM, Cornell C, Damiano SR, Paxton SJ. Parents and prevention: a systematic review of interventions involving parents that aim to prevent body dissatisfaction or eating disorders. Int J Eat Disord 2015 Mar;48(2):157-169. [doi: 10.1002/eat.22284] [Medline: 24796581]
54. Bruning BJ, Winzelberg AJ, Abascal LB, Taylor CB. An evaluation of an Internet-delivered eating disorder prevention program for adolescents and their parents. J Adolesc Health 2004 Oct;35(4):290-296. [doi: 10.1016/j.jadohealth.2003.10.010] [Medline: 15450542]
55. Varnado-Sullivan PJ, Zucker N, Williamson DA, Reas D, Thaw J, Netemeyer SB. Development and implementation of the body logic program for adolescents: A two-stage prevention program for eating disorders. Cognitive and Behavioral Practice 2001 Jun;8(3):248-259. [doi: 10.1016/S1077-7229(01)80061-4]
56. World Health Organisation. Adherence to Long-term Therapies: Evidence for Action. Geneva: Author; 2003.
57. Dew MA, Dabbs AD, Myaskovsky L, Shyu S, Shellmer DA, DiMartini AF, et al. Meta-analysis of medical regimen adherence outcomes in pediatric solid organ transplantation. Transplantation 2009 Sep 15;88(5):736-746 [FREE Full text] [doi: 10.1097/TP.0b013e3181b2a0e0] [Medline: 19741474]
58. Santer M, Ring N, Yardley L, Geraghty AWA, Wyke S. Treatment non-adherence in pediatric long-term medical conditions: systematic review and synthesis of qualitative studies of caregivers' views. BMC Pediatr 2014 Mar 04;14:63 [FREE Full text] [doi: 10.1186/1471-2431-14-63] [Medline: 24593304]

Abbreviations

AN: anorexia nervosa
BMI: body mass index
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition
E@T: Eltern als Therapeuten
EBW: expected body weight
ED: eating disorder
EDE: Eating Disorder Examination
EDI: Eating Disorder Inventory
FU: follow-up
ITT: intention-to-treat
MPS-F: Frost Multidimensional Perfectionism Scale
PMI: Parent Motivation Inventory
WCS: Weight Concerns Scale

Edited by G Eysenbach; submitted 14.12.17; peer-reviewed by S Agras, R Schuster; comments to author 29.01.18; revised version received 20.03.18; accepted 29.06.18; published 14.12.18.

Please cite as:
Jacobi C, Hütter K, Völker U, Möbius K, Richter R, Trockel M, Jones Bell M, Lock J, Taylor CB
Efficacy of a Parent-Based, Indicated Prevention for Anorexia Nervosa: Randomized Controlled Trial
J Med Internet Res 2018;20(12):e296
URL: http://www.jmir.org/2018/12/e296/
doi:10.2196/jmir.9464
PMID:30552078

Original Paper

Internet-Based Cognitive Behavioral Therapy With Real-Time Therapist Support via Videoconference for Patients With Obsessive-Compulsive Disorder, Panic Disorder, and Social Anxiety Disorder: Pilot Single-Arm Trial

Kazuki Matsumoto[1,2], MPH; Chihiro Sutoh[3], MD, PhD; Kenichi Asano[2], PhD; Yoichi Seki[1,2], PhD; Yuko Urao[1,2], PhD; Mizue Yokoo[2], PhD; Rieko Takanashi[2], PhD; Tokiko Yoshida[2], MPH; Mari Tanaka[2], PhD; Remi Noguchi[2], PhD; Shinobu Nagata[2], PhD; Keiko Oshiro[2], PhD; Noriko Numata[1,2], PhD; Motohisa Hirose[2], PhD; Kensuke Yoshimura[4], MD, MPH; Kazue Nagai[5], PhD; Yasunori Sato[6], PhD; Taishiro Kishimoto[7], MD, PhD; Akiko Nakagawa[1,2], MD, PhD; Eiji Shimizu[2,3], MD, PhD

[1]United Graduate School of Child Development, Osaka University, Kanazawa University, Hamamatsu University School of Medicine, Chiba University and University of Fukui, Osaka, Japan

[2]Research Center for Child Mental Development, Graduate School of Medicine, Chiba University, Chiba, Japan

[3]Department of Cognitive Behavioral Physiology, Graduate School of Medicine, Chiba University, Chiba, Japan

[4]Research Center for Medical Economics Administration, Chiba University Hospital, Chiba, Japan

[5]Reseach and Education Center of Health Sciences, Gunma University Graduate School of Health Sciences, Gunma, Japan

[6]Department of Preventive Medicine and Public Health, Keio University, Tokyo, Japan

[7]Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan

Corresponding Author:
Chihiro Sutoh, MD, PhD
Department of Cognitive Behavioral Physiology
Graduate School of Medicine
Chiba University
1-8-1, Chuo-ku, Chiba-shi
Chiba,
Japan
Phone: 81 43 226 2027
Email: csutoh@graduate.chiba-u.jp

Related Article:

Comment in: https://www.jmir.org/2020/8/e13234

Abstract

Background: Cognitive behavioral therapy (CBT) is the first-line treatment for adults with obsessive-compulsive disorder (OCD), panic disorder (PD), and social anxiety disorder (SAD). Patients in rural areas can access CBT via the internet. The effectiveness of internet-delivered cognitive behavioral therapy (ICBT) has been consistently shown, but no clinical studies have demonstrated the feasibility of ICBT with real-time therapist support via videoconference for OCD, PD, and SAD at the same time.

Objectives: This study aimed to evaluate the feasibility of videoconference-delivered CBT for patients with OCD, PD, or SAD.

Methods: A total of 30 Japanese participants (mean age 35.4 years, SD 9.2) with OCD, SAD, or PD received 16 sessions of individualized videoconference-delivered CBT with real-time support of a therapist, using tablet personal computer (Apple iPad Mini 2). Treatment involved individualized CBT formulations specific to the presenting diagnosis; all sessions were provided by the same therapist. The primary outcomes were reduction in symptomatology, using the Yale-Brown obsessive-compulsive scale (Y-BOCS) for OCD, Panic Disorder Severity Scale (PDSS) for PD, and Liebowitz Social Anxiety Scale (LSAS) for SAD. The secondary outcomes included the EuroQol-5 Dimension (EQ-5D) for Quality of Life, the Patient Health Questionnaire

(PHQ-9) for depression, the Generalized Anxiety Disorder (GAD-7) questionnaire for anxiety, and Working Alliance Inventory-Short Form (WAI-SF). All primary outcomes were assessed at baseline and at weeks 1 (baseline), 8 (midintervention), and 16 (postintervention) face-to-face during therapy. The occurrence of adverse events was observed after each session. For the primary analysis comparing between pre- and posttreatments, the participants' points and 95% CIs were estimated by the paired *t* tests with the change between pre- and posttreatment.

Results: A significant reduction in symptom of obsession-compulsion (Y-BOCS=−6.2; Cohen *d*=0.74; 95% CI −9.4 to −3.0, *P*=.002), panic (PDSS=−5.6; Cohen *d*=0.89; 95% CI −9.83 to −1.37; *P*=.02), social anxiety (LSAS=−33.6; Cohen *d*=1.10; 95% CI −59.62 to −7.49, *P*=.02) were observed. In addition, depression (PHQ-9=−1.72; Cohen *d*=0.27; 95% CI −3.26 to −0.19; *P*=.03) and general anxiety (GAD-7=−3.03; Cohen *d*=0.61; 95% CI −4.57 to −1.49, *P*<.001) were significantly improved. Although there were no significant changes at 16 weeks from baseline in EQ-5D (0.0336; Cohen *d*=-0.202; 95% CI −0.0198 to 0.00869; *P*=.21), there were high therapeutic alliance (ie, WAI-SF) scores (from 68.0 to 73.7) throughout treatment, which significantly increased (4.14; 95% CI 1.24 to 7.04; *P*=.007). Of the participants, 86% (25/29) were satisfied with videoconference-delivered CBT, and 83% (24/29) preferred videoconference-delivered CBT to face-to-face CBT. An adverse event occurred to a patient with SAD; the incidence was 3% (1/30).

Conclusions: Videoconference-delivered CBT for patients with OCD, SAD, and SAD may be feasible and acceptable.

(J Med Internet Res 2018;20(12):e12091) doi:10.2196/12091

KEYWORDS

clinical trial; cognitive behavioral therapy; feasibility study; obsessive-compulsive disorder; panic disorder; social anxiety disorder; videoconference

Introduction

Background

Obsessive-compulsive disorder (OCD), panic disorder (PD), and social anxiety disorder (SAD) are the most common mental disorders and incur a huge burden throughout the lifespan [1-4]. Cognitive behavioral therapy (CBT) has been found to be effective in treating all of the 3 disorders [5-16]. Although CBT is an effective treatment, it is difficult for all patients to receive CBT because of problems of access to treatment such as expensive specialized medical treatment, lack of therapists, and uneven urban distribution. Known as telemental health, a new therapeutic approach, born out of technological innovation after the internet revolution, has solved these problems. A patient can now receive treatment from remote distance via the internet, regardless of physical or psychological barriers such as time, distance, and stigma from remote distance by the internet [17,18]. For example, in internet-delivered CBT (ICBT), users can receive their programmed treatment at any time, 24 hours a day. In addition, the therapist's involvement with the patient can be reduced, optimizing treatment costs. Furthermore, a therapist guide can be employed via remote treatment, by using videoconference in real time, without compromising linguistic and nonverbal communication with the patient as much as possible.

ICBT has been shown to be effective for OCD, PD, and SAD from several randomized controlled trials [19-35]. A recent systematic review and meta-analyses using 20 studies comparing ICBTs and face-to-face CBTs showed that ICBT and face-to-face treatment produced equivalent overall effects [36]. Target diseases in this meta-analysis included SAD, PD, and depression. In the other meta-analyses, comparing the effectiveness of ICBT and face-to-face CBT with anxiety disorders including OCD (defined according to the Diagnostic and Statistical Manual of Mental Disorder III [37], III-R [38], IV [39], and IV-TR or the International Classification of Diseases 9 or 10 [40-42]), there were no clear differences between them [43]. ICBT can be roughly divided into the 3 categories depending on how the therapist participates in the program [44]: programs with no therapist assistance [45]; programs with assistance, where the therapists assistance is minimal [46]; and live conversations on the internet, where the therapist is fully involved using videoconference [47,48]. Although the effectiveness of ICBT has been suggested, there is little knowledge about ICBT including videoconference. In telemental health, which is an important theme when considering the optimization of social resources, the effectiveness of ICBT should be examined based on the therapist's degree of involvement. In this study, we examined the feasibility of ICBT with real-time videoconference, where the therapists were fully involved in treatment.

Videoconference-Delivered Cognitive Behavioral Therapy

In videoconference-delivered CBT, the therapist and patient use video and audio links to have a therapeutic conversation, as in face-to-face CBT, which includes nonverbal information such as expressions, body language, voice volume, and tone. Therefore, among the ICBT options, this allows the patient and the therapist to have the strongest therapeutic relevance. In a situation where real-time communication with the patient is important, it is considered to be a powerful approach; for example, when the therapist can enhance motivation through role playing by performing internal sensory exposure in front of a patient with serious panic PD.

Mental health services for remote populations via videoconferences have yielded high satisfaction [17,49-52]. A systematic review of videoconference-delivered CBT trials, using 20 controlled studies, uncontrolled studies, case series, and case studies for anxiety disorders including posttraumatic stress disorder (PTSD) using the criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), suggested

a medium to large effect size [53]. At the same time, there were no clinical studies about videoconference-delivered CBT for all of OCD, PD, and SAD. Comorbidity has often been seen between OCD, PD, and SAD [54]; clinical trials of CBT with additional remote systems to determine videoconference-delivered CBT adaptation to normal clinical scenes are thus meaningful. A randomized controlled trial compared OCD using professional videoconferencing equipment (n=10) with self-help (n=10) and wait-list (n=10) controls [51]. The results indicated large effect sizes (Cohen *d*=2.1-2.5); in addition, 60% (6/10) of participants in the videoconference condition achieved clinically significant changes at posttreatment and 50% (5/10) did so at a 3-month follow-up. Another controlled trial compared manualized CBT for PD and agoraphobia by videoconference (n=11) with in-person CBT (n=10) [52]. The results indicated no significant difference between the conditions, but 81% (9/11) of participants in the videoconferencing condition were panic-free at posttreatment and 91% (10/11) at a 9-month follow-up; this indicated a very large effect size for all panic and agoraphobia symptoms. An uncontrolled trial, using 24 participants with SAD, applied acceptance-based behavioral therapy via videoconference [50] and found that 54% (13/24) of participants did not meet DSM-IV-TR criteria for SAD at posttreatment. Large effect sizes for all social anxiety symptoms (Cohen *d*=1.23-1.91) were obtained at post-therapy and at a 3-month follow-up. In addition, it is known that videoconference-delivered therapy can develop a similarly strong therapeutic alliance with psychotherapy clients as in-person therapy [49,55].

Nevertheless, during this trial, the only national university hospital with a CBT specialized outpatient in Japan was at our facility alone, the Chiba University; most patients do not have access to this treatment. For patients in Japan to receive CBT via the internet, it was necessary to examine the feasibility of videoconference-delivered CBT.

Objective of This Study

Previous studies have consistently demonstrated the effectiveness of videoconference-delivered CBT, but they have small samples sizes and were conducted in Western countries. Therefore, this pilot study utilized an open trial to examine the feasibility and preliminary effectiveness of in-home CBT via videoconference for Japanese adults with OCD [49], PD, and SAD. It was hypothesized that videoconference-delivered CBT would be effective to reduce the symptomatology for each disorder from pre- to posttreatment and acceptable to Japanese participants.

Methods

Study Design

This study was conducted as a single-arm, open trial at the academic outpatient clinic of the Cognitive Behavioral Therapy Center of Chiba University Hospital between March 2017 and March 2018. Since this trial was the first to employ an individual videoconference-delivered CBT intervention design against OCD, PD and SAD in Japan, a single-arm trial examining baseline predictors rather than effectiveness was considered to be an appropriate design [56].

Ethics and Dissemination

This trial was approved by the Institutional Review Board of Chiba University Hospital (reference number: G28038). The clinical trial registration number was UMIN000026609.

If an individual wished to participate in the trial, they had to contact the study trial office, where they were informed about the study objectives and asked to confirm whether they were willing to participate. Furthermore, they were assured absolute anonymity. They had to fill out an informed consent form for participation in this study. All participants were informed that they could continue receiving conventional drug treatments from their primary doctors. We practiced videoconferencing twice with participants.

Participants and Eligibility Criteria

The study participants were recruited through posters and leaflets placed at medical institutions in Chiba Prefecture, through the official Web-based advertisements at the Cognitive Behavioral Therapy Center of Chiba University Hospital and by referrals from their primary care doctors or psychiatrists. After email or telephone screening through Web-based app, the participants visited our center and were diagnosed with OCD, SAD, or PD using the Mini-International Neuropsychiatric Interview [57,58].

Inclusion criteria for this study included informed consent to participate in the study; having a primary diagnosis of OCD, PD, or SAD; aged between 19 and 65 years, and having access to the internet at home. Comorbid mental disorders—including major depressive disorder, other anxiety disorders, and eating disorders—were permitted if they were clearly secondary, considering that this trial should reflect routine clinical practice. The exclusion criteria were organic brain damage, dementia, psychotic disorders, serious drug dependence, recurrent suicidal and antisocial behaviors, and severe somatic conditions. Participants, who used psychotropic drugs, including selective serotonin reuptake inhibitors and benzodiazepines, were asked to report all of changes regarding pharmacotherapy during the study period.

Intervention

The participants entered a Web conference room by clicking a URL in an email sent from their therapists. The intervention was conducted at a 50-min session once a week for 16 weeks. The modules were derived from previous studies on in-person CBT for OCD [59], PD [60], and SAD [61] in Japan and included psychoeducation, exposure exercises, behavioral experiments, and homework assignments.

Therapists and Therapy Quality Control

Videoconference-delivered CBT was delivered by 12 therapists, who were experienced in face-to-face CBT for patients with OCD, PD, and/or SAD (including 7 clinical psychologists, 2 psychiatric social workers, 1 nurse, 1 psychiatric pharmacist, and 1 psychiatrist). Therapists were trained in CBT programs for patients with OCD, PD, and SAD and attended weekly group-supervision sessions with other therapists as well as undergoing individual supervision by a senior supervisor. All therapists had completed a CBT training course (Chiba

Improving Access to Psychological Therapies project: Chiba-IAPT) [62]. Of the therapists, 6 were female, with an average age of 43.5 years (SD 7.5) and an average of 2.2 years of clinical experience (SD 6.4) at the beginning of the study. Senior supervisors assessed the quality of the videoconference-delivered CBT sessions using the Cognitive Therapy Scale-Revised [63,64], a revision of the Cognitive Therapy Scale designed by Young and Beck (unpublished data [65,66]).

Visual Aids

The use of visual aids facilitates the learning process by enhancing motivation and understanding of complex matters [67]. Therefore, visual aids were used in each program to enhance the participants' understanding. The visual aids consisted of several slides including key concepts of CBT. The therapists conducted CBT sessions with the visual aids by using the screen-sharing function of the videoconference software, sending them to the participants as password-protected homework slides by email after each session.

Hardware

The therapists used 2 Surface Pro 2 computers—2-in-1 detachable produced by Microsoft , running Windows 10 Pro (Microsoft Corporation, US). The display size was 10.6 inches and a resolution of 1920×1080 pixels. Each participant was lent an iPad Mini 2 (Apple Inc, US) with a 7.9-inch display and a 2048×1536-pixel resolution.

Software for Videoconference

A total of 3 licenses for videoconference software (Cisco WebEx, Milpitas, CA, USA) were used in this trial. This system has been awarded ISO27001 certification (regarding handling of information security) and SSAE16 (Statement of Standards for Attestation Engagements No. 16: former SAS 70) compliance certification (issued by a third party). It also complies with the United States *Health Insurance Portability and Accountability Act*. WebEx's use of a switching network along with a 128-bit Secure Sockets Layer encryption and public key infrastructure is regarded by the Japan Ministry of Health, Labor and Welfare to have solved the problem of safety, as reported in their "Guidelines on Safety Management of Medical Information System Version 4.3" in March 2016 [68]. Since the stability and safety of the software are excellent and it sufficiently protects personal information, we judged WebEx to be sufficiently reliable for this study.

Measures of Primary Outcomes

The most commonly used scales to measure symptoms of each disorder were used as follows: the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) for OCD symptoms [69,70], the Panic Disorder Severity Scale (PDSS) for PD symptoms [71,72], and the Liebowitz Social Anxiety Scale (LSAS) for SAD symptoms [73,74].

To calculate responsiveness to treatment and the remission rate after the intervention, the criteria of the previous studies regarding the severity rating scale of each disorder was used [62,64,65]. For OCD, treatment response was defined as a 35% or greater reduction in the total Y-BOCS score, and remission was defined as a final Y-BOCS score of ≤14 [75]. For PD, treatment response was defined as a 40% or greater reduction in total PDSS score, and remission was defined as a final PDSS score of ≤5 [76]. For SAD, treatment response was defined as a 31% or greater reduction in total LSAS score, and remission was defined as a final LSAS score of ≤36 [73].

Measures of Secondary Outcomes

We used the EuroQol-5 Dimension (EQ-5D) to measure health-related quality of life [77,78]. This trial measured the psychological bond between therapist and participant using the Working Alliance Inventory-Short Form (WAI-SF) [79], depressive symptoms using the Patient Health Questionnaire-9 (PHQ-9) [80,81], and generalized anxiety symptoms using the Generalized Anxiety Disorder-7 (GAD-7) [82,83]. The definition of the response in PHQ-9 and GAD-7 was defined as a 50% reduction in total score. We used 7-point Likert scale format to measure participants' satisfaction about videoconference-delivered CBT as follows: "Very dissatisfied," "Dissatisfied," "Slightly dissatisfied," "Neutral," "Slightly satisfied," "Satisfied," and "Very satisfied." In addition, the participants were asked about preference of videoconference-delivered CBT or face-to-face CBT as follows: "If you could choose in the future, would you wish to receive treatment with either face-to-face or videoconference CBT?" Participants answered using a 7-point Likert scale as follows: "Clearly prefer face-to-face," "Prefer face-to-face," "Slightly prefer face-to-face," "Neutral," "Slightly prefer videoconference-delivered CBT," "Prefer videoconference-delivered CBT," and "Clearly prefer videoconference-delivered CBT."

Data Setting and Locations

Participant and therapists used the Numbers app for iOS to run the digital questionnaires, and the therapist asked each participant to answer them by themselves on the tablet PC at weeks 1, 8, and 16. Each participant sent an email with the completed questionnaires of all primary outcomes (Y-BOCS, PDSS, and LSAS) and part of secondary outcomes (EQ-5D, PHQ-9, and GAD-7) to their therapist before each session and sent an email with the completed questionnaires of the secondary outcomes (WAI-SF and satisfaction/preference) attached after session. The therapist checked outcomes and evaluated the symptoms during the session, collaborating with the participant. The collected data were registered to the server of DATATRAK ONE (DATATRAK International Inc, US) as Web case registration system by the lead author and managed by the data management office of Chiba University. This study adhered to the CONSORT-EHEALTH guidelines for improving and standardizing the report of Web-based and mobile health interventions [84].

Adverse Events

To confirm the occurrence of adverse events after intervention, the therapist asked the patient about their physical and mental condition at the end of each session and instructed all participants to report all adverse events by email.

Statistical Analysis

Statistical analysis and reporting of this trial were conducted in accordance with the CONSORT-EHEALTH guidelines [84]. All statistical analyses were described in the statistical analysis plan, which was fixed before the database lock. All efficacy analyses were primarily based on the full analysis set, which included all patients who had received at least one session of the videoconference-delivered CBT treatment. For baseline variables, summary statistics were constructed, employing frequencies and proportions for categorical data and means and SDs for continuous variables. Baseline variables were compared using the Fisher exact test for categorical outcomes and the unpaired *t* test for continuous variables. For the primary analysis comparing between pre- and posttreatments, the points and their 95% CIs were estimated by the paired *t* tests with the change at week 16 from baseline in EQ-5D index scores for all of the patients in Y-BOCS for OCD, PDSS for PD, and in LSAS for SAD. For comparison among the 3 disorders, a one-way analysis of variance was used. Analyses of secondary outcomes were performed in the same manner as the primary analysis.

In addition, we calculated Cohen *d* pre-post effect sizes by calculating the mean differences between pre- and posttreatments, dividing by the pooled SDs. We also adopted the criteria that a Cohen *d* of >0.20 was a small effect, that of >0.50 was a medium effect, and that of >0.80 was a large effect [85]. All *P* values were two-sided; a value of $P<.05$ was considered statistically significant. Statistical analyses were performed using SAS software version 9.4 (SAS Institute, Cary, North Carolina, USA) and the R statistical program version 2.13 (The R Foundation, Vienna, Austria).

Results

Recruitment

Figure 1 shows the participant flow. A total of 37 patients applied to participate through our website. After email or telephone screening, 6 patients were excluded; 1 did not meet one of the inclusion criteria due to epilepsy and 5 declined to participate because of long distance to our hospital. After the screening, 31 attended the face-to-face baseline assessment, and one did not meet one of the inclusion criteria due to high risk of suicide. Finally, 30 patients were enrolled to the study.

Attrition

Of the participants eligible to take part in the study, 1 SAD participant with major depressive disorder dropped out after 9 sessions because of worsening of his depressive state. The remaining 97% (29/30) completed the full course of videoconference-delivered CBT. All data at point each assessment (screening, session 1, session 8, and session 16) were statistically analyzed.

Clinical and Demographic Characteristics

The sample included 30 participants (6 males and 24 females), aged 20 to 54 years (mean 35.4 years, SD 9.2), education 10 to 19 years (mean 14.8 years, SD 2.1). Apart from primary diagnoses, a summary of the participants' demographic and diagnostic information is presented in Table 1. Moreover, 15 participants continued to receive pharmacotherapy during the trial (4 fluvoxamine, 2 escitalopram, 1 sertraline, 1 paroxetine hydrochloride, 1 duloxetine, 1 mirtazapine, 1 trazodone, 1 ethyl loflazepate, 1 alprazolam, and 1 clotiazepam).

Primary Outcomes

There were significant reductions for each symptoms of obsession-compulsion (Y-BOCS=−6.2; Cohen *d*=0.74; 95% CI −9.4 to −3.0; *P*=.002), panic (PDSS=−5.6; Cohen *d*=0.89; 95% CI −9.83 to −1.37; *P*=.02), and social anxiety (LSAS=−33.3; Cohen *d*=1.10; 95% CI −59.62 to −7.49; *P*=.02). Of the participants with OCD, 20% (2/10) showed a treatment response, whereas 40% (4/10) went into remission [75]. Of the participants with PD, 60% (6/10) showed a treatment response and 50% (5/10) went into remission [76]. Of the participants with SAD, 44% (4/9) showed a treatment response and 22% (2/9) went into remission [73].

Secondary Outcomes

Figure 2 shows the change in the primary outcomes. Table 2 shows the mean change in the EQ-5D scores, at 16 sessions from baseline. The adjusted mean changes of the EQ-5D for all of the 3 disorders was 0.0336 (95% Cl −0.0198 to 0.0869; *P*=.21), which showed that it was not significant and showed a small effect size (Cohen *d*=- 0.202).

Figure 3 also shows the change in the secondary outcomes. The mean changes in the PHQ-9 and GAD-7 scores reflected significant decreases in total participants for the 3 disorders (Table 3). The mean change in the WAI-SF reflected a significant increase for the total sample. Table 4 shows participants' satisfaction with and preferences for videoconference-delivered CBT. As the ratings "very satisfied" and "satisfied" were combined, the majority of participants (86%, 25/29) reported being satisfied with videoconference-delivered CBT. In that case, the ratings "slightly prefer videoconference-delivered CBT," "prefer videoconference-delivered CBT," and "clearly prefer videoconference-delivered CBT" were combined; 83% (24/29) of the participants preferred videoconference-delivered CBT to face-to-face CBT. Conversely, 7% (2/29) of them preferred face-to-face CBT to videoconference-delivered CBT.

Figure 1. Participant flow.

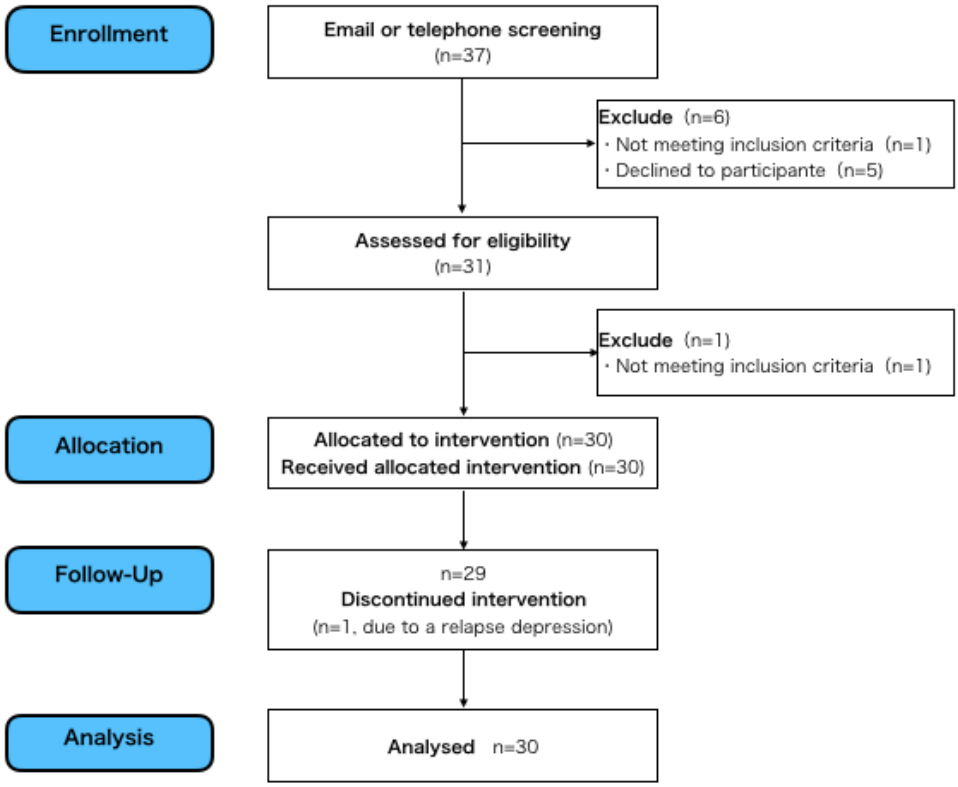


Table 1. Clinical and demographic characteristics of participants (N=30).

Characteristics	All	OCD[a]	PD[b]	SAD[c]
Age (years), mean (SD)	35.4 (9.2)	37.7 (6.9)	38.8 (9.8)	29.7 (8.6)
Sex, n (%)				
Male	6 (20)	2 (6)	0 (0)	4 (13)
Female	24 (80)	8 (27)	10 (33)	6 (20)
Employment, n (%)	11 (36)	5 (17)	2 (6)	4 (13)
Combined pharmacotherapy, n (%)	15 (50)	6 (20)	5 (17)	4 (13)
Videophone use experience, n (%)	15 (50)	4 (13)	6 (20)	5 (17)
Comorbid disorders, n (%)				
Depression	5 (17)	1 (3)	1 (3)	3 (10)
Panic/agoraphobia	2 (6)	2 (6)	N/A[d]	N/A
PTSD[e]	1 (3)	N/A	1 (3)	N/A
Alcohol dependence	1 (3)	N/A	N/A	1 (3)
Bulimia nervosa	1 (3)	N/A	N/A	1 (3)
GAD[f]	3 (10)	2 (6)	N/A	1 (3)

[a]OCD: obsessive-compulsive disorder.
[b]PD: panic disorder.
[c]SAD: social anxiety disorder.
[d]N/A: not applicable.
[e]PTSD: posttraumatic stress disorder.
[f]GAD: generalized anxiety disorder.

Figure 2. Change of primary outcomes. LSAS: Liebowitz Social Anxiety Scale; PDSS: Panic Disorder Severity Scale; Y-BCOS: Yale-Brown Obsessive-Compulsive Scale.

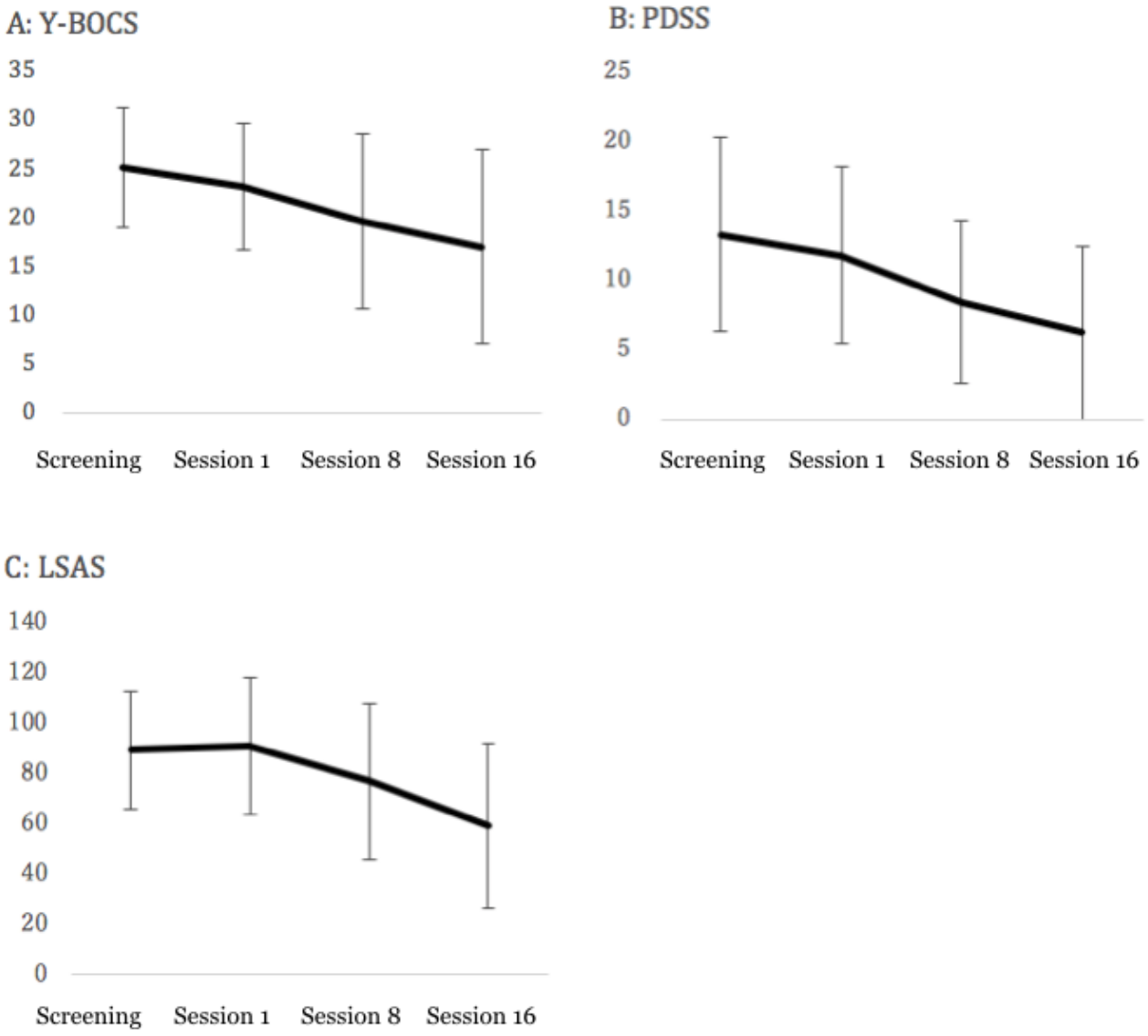


Table 2. Mean change in EuroQol-5 Dimension score.

Disorder	Mean change (95% CI)	*P* value (paired *t* test)	*P* value (*F* test)
All (n=29)	0.0336 (−0.0198 to 0.0869)	.21	N/A[a]
OCD[b] (n=10)	0.0488 (−0.0577 to 0.1553)	.33	.91
PD[c] (n=10)	0.0305 (−0.0393 to 0.1003)	.35	.91
SAD[d] (n=9)	0.0201 (−0.1188 to 0.1591)	.75	.91

[a]N/A: not applicable.

[b]OCD: obsessive-compulsive disorder.

[c]PD: panic disorder.

[d]SAD: social anxiety disorder.

Figure 3. Change of secondary outcomes. EQ-5D-5L: EuroQol-5 Dimension; GAD-7: Generalized Anxiety Disorder-7; PHQ-9: Patient Health Questionnaire-9; WAI-SF: Working Alliance Inventory-Short Form.

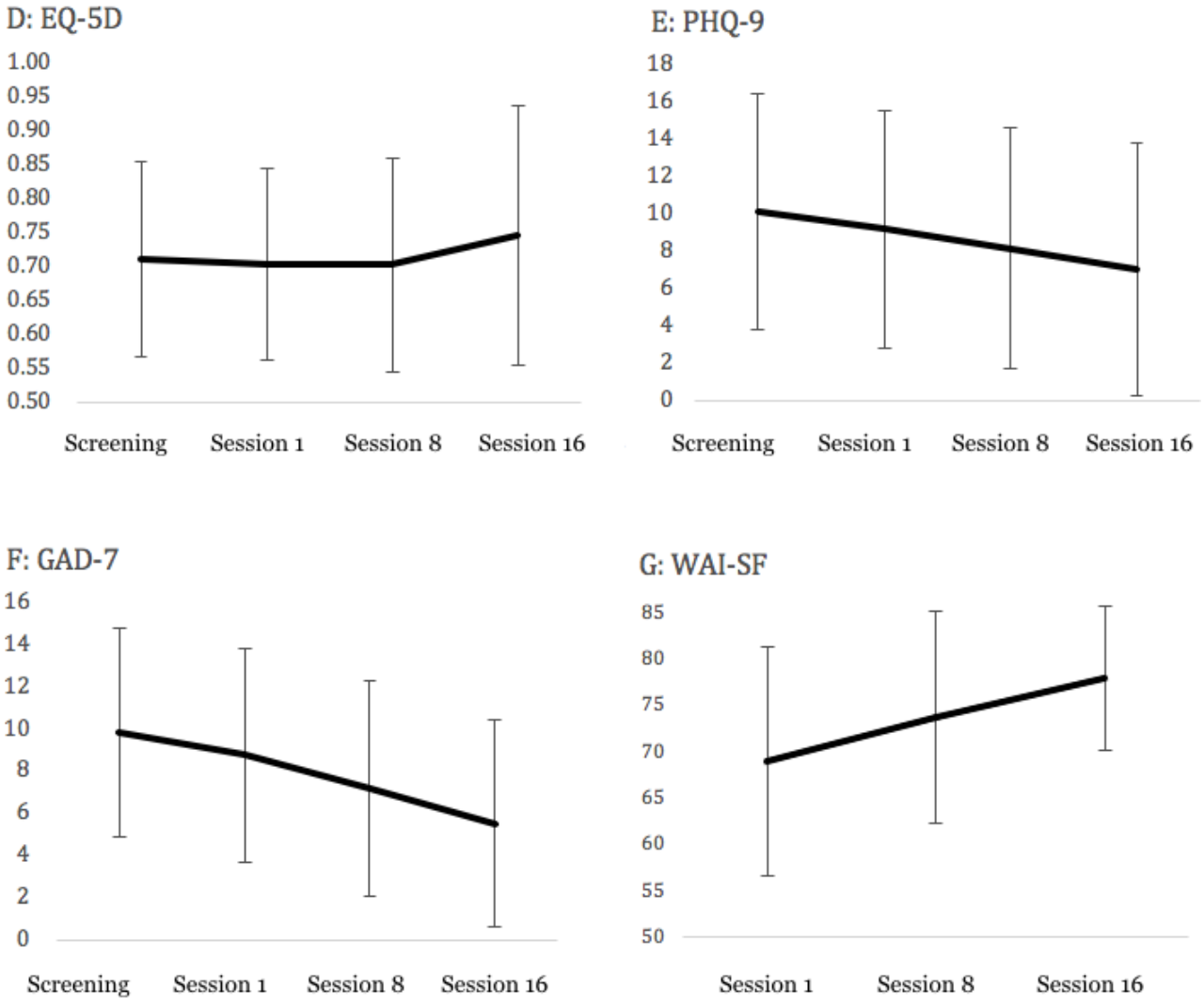


Table 3. Mean changes in secondary outcomes.

Measures and Disorder	Mean change, *t* value (95% CI)	*P* value (paired *t* test)	*P* value (*F* test)
PHQ-9[a]			
All (n=29)	−1.72 (−3.26 to −0.19)	.03	N/A[b]
OCD[c] (n=10)	−1.70 (−4.52 to 1.12)	.21	.27
PD[d] (n=10)	−0.30 (−1.65 to 1.05)	.63	.27
SAD[e] (n=9)	−3.33 (−7.56 to 0.89)	.11	.27
GAD-7[f]			
All (n=29)	−3.03 (−4.57 to −1.49)	<.001	N/A
OCD (n=10)	−.50 (−6.37 to −0.63)	.002	.68
PD (n=10)	−2.10 (−3.43 to 0.77)	.006	.68
SAD (n=9)	−3.56 (−8.02 to 0.91)	.10	.68
WAI-SF[g]			
All (n=29)	4.14 (1.24 to 7.04)	.007	N/A
OCD (n=10)	3.00 (−1.77 to 7.77)	.19	.85
PD (n=10)	4.80 (−0.42 to 10.02)	.07	.85
SAD (n=9)	4.67 (−2.66 to 11.99)	.18	.85

[a]PHQ-9: Patient Health Questionnaire-9.
[b]N/A: not applicable.
[c]OCD: obsessive-compulsive disorder.
[d]PD: panic disorder.
[e]SAD: social anxiety disorder.
[f]GAD-7: Generalized Anxiety Disorder-7.
[g]WAI-SF: Working Alliance Inventory-Short Form.

Table 4. Satisfaction and preference.

Answer options	n (%)
Satisfaction	
Very dissatisfied	0 (0)
Dissatisfied	0 (0)
Slightly dissatisfied	1 (3)
Neutral	0 (0)
Slightly satisfied	3 (10)
Satisfied	9 (31)
Very satisfied	16 (55)
Preference	
Clearly prefer face-to-face	0 (0)
Prefer face-to-face	1 (3)
Slightly prefer face-to-face	1 (3)
Neutral	3 (10)
Slightly prefer videoconference-delivered CBT[a]	10 (34)
Prefer videoconference-delivered CBT	6 (21)
Clearly prefer videoconference-delivered CBT	8 (28)

[a]CBT: cognitive behavioral therapy.

Adverse Events

A total of 3 patients reported adverse events, including depression relapse, headache, and feeling of exhaustion. The depressive symptoms of 1 SAD patient with depressive disorder worsened between the ninth and tenth sessions, when he was travelling with his friend. We identified this as a serious adverse event at the tenth session. He wanted to decline continuing with videoconference-delivered CBT and dropped out of the trial at that time. At 6 months after he received pharmacotherapy from his psychiatrist, he recovered the depressive episode. In addition, 1 PD participant reported a headache at the fourth session but recovered in the same day. Furthermore, 1 OCD patient reported a feeling of exhaustion after the 4th session but recovered in the same day.

Discussion

Principal Findings

This study examined the feasibility of videoconference-delivered CBT in adult patients with mild to severe OCD, PD, and SAD. Interventions based on CBT were conducted for each group divided by primary diagnosis, and examination of symptom improvement and acceptance of patients was conducted before and after the intervention. We use different criteria for each disease looking at the rate of responders to treatment (defined as a 35% reduction in Y-BOCS obsessive-compulsive symptoms, a 40% reduction in PDSS panic symptoms, and a 31% reduction in LSAS social anxiety symptoms); patients' satisfaction were also confirmed by using therapeutic alliance and patients' treatment acceptance. Improvement of the symptoms was confirmed in 3 disorders; it was found that the therapeutic alliance was achieved at a high level, and patients' satisfaction was extremely high. Therefore, this study showed the feasibility of ICBT with real-time support of therapists to Japanese patients except for depression [86].

Feasibility of Videoconference-Delivered Cognitive Behavioral Therapy

Regarding the other primary outcomes, the calculated Cohen *d* for pre- to posttreatment were 0.74 for the Y-BOCS, 0.89 for the PDSS, and 1.10 for the LSAS. The Cohen *d* scores were classified as medium and large. Though it is difficult to compare these studies because the characteristics of the patients and/or the methods were different, these results seem similar to those of our face-to-face CBT studies [59-61]. These medium and large effect sizes were also found in the previous studies of videoconference-delivered CBT (Cohen *d*=1.4-2.5) [50-52]. A previous videoconference-delivered CBT study for OCD reported that the treatment response rate was 60% (6/10). According to a systematic review about face-to-face CBT studies conducted between 2000 and 2014, the response rates were 43.3% for OCD, 53.2% for PD, and 45.3% for SAD [87]. The response rates in this study were 20% (2/10) for OCD, 60% (6/10) for PD, and 44% (4/9) for SAD. Previous videoconference-delivered CBT studies reported that the remission rate for PD was 81% (9/11) [52] and that for SAD was 54% (13/24) [50]. The remission rates in our studies were 40% (4/10) for OCD, 50% (5/10) for PD, and 22% (2/9) for SAD. Although comparisons of these results must be done with caution, the response and remission rates of this study were comparable with those in the previous studies of in-person CBT and videoconference-delivered CBT. In the future, it will be necessary to verify the effectiveness of our videoconference-delivered CBT through a randomized controlled trial or noninferiority trial in comparison with face-to-face CBT or videoconference-delivered CBT with different methodology.

There was no significant change (*P*=.21) in the EQ-5D scores for all of the 3 disorders, the calculated Cohen *d* from pre- to posttreatment was −0.202 and classified as small. In addition, there were no significant differences in changes of the EQ-5D score among the 3 disorder groups (*P*=.91). As described by the previous reports [88], our findings suggested that the EQ-5D was responsive in videoconference-delivered CBT for OCD, PD, and SAD.

For the PHQ-9, a significant reduction between pre- and posttreatment was observed for the entire sample (*P*=.03) There were no significant differences in the PHQ-9 changes among the 3 disorders (*P*=.27). The effect size for the 3 disorders was small (Cohen *d*=0.27). After dividing each disorder, the effect sizes ranged from small to medium (OCD: Cohen *d*=0.23; PD: Cohen *d*=0.07; SAD: Cohen *d*=0.64). As for the GAD-7, a significant reduction pre- and posttreatment was observed for the entire sample (*P*<.001). Although the change reflected a medium effect size for the entire sample (Cohen *d*= 0.61), all were medium for each disorder (OCD: Cohen *d*=0.75; PD: Cohen *d*=0.79; SAD: Cohen *d*=0.67). There were significant differences in changes between the OCD and PD groups (OCD: *P*=.002; PD: *P*=.006; SAD: *P*=.10) at week 16. A previous study reported a response rate in GAD-7 of 50.9% following a computerized CBT program, for 1062 adults who had GAD-7 scores of 10 or more at baseline [89] including 75 patients with GAD, 47 with PD, 40 with SAD, and 18 with PTSD. In this study, the treatment response rate for the GAD-7 was 45% (13/29). The response rates of PHQ-9 and GAD-7 in this study were similar to those in a prior study [89]. Taken together, our results suggest that videoconference-delivered CBT for OCD, PD, and SAD might secondarily ameliorate symptoms for generalized anxiety and depression.

More than half of the participants (55%, 16/29) who completed the videoconference-delivered CBT reported the highest level of satisfaction ("very satisfied") with treatment via videoconferencing, whereas 31% (9/29) reported that they were "satisfied" (the second highest level). In other words, 86% (25/29) of participants reported being satisfied with videoconference-delivered CBT. These results are consistent with those of previous studies [50,90]. Furthermore, 83% (24/29) of the participants preferred videoconference-delivered CBT to face-to-face CBT. Conversely, 7% (2/29) preferred face to face CBT to videoconference-delivered CBT. Taken together, these results indicated that videoconference-delivered CBT was generally accepted by Japanese participants with OCD and anxiety disorders.

The therapeutic alliance indicated by the total scores of WAI-SF significantly improved throughout the treatment, from 68.9 (SD 12.3) at week 1 to 77.9 (SD 7.7) at week 16 (*P*=.007). The mean

scores of the WAI-SF items were 5.7 (SD 0.97) at pretreatment, 6.1 (SD 0.95) at midtreatment, and 6.5 (SD 0.63) at posttreatment. These results were comparable with those of a previous study on videoconference-delivered CBT, where the WAI-SF scores increased from 5.22 (SD 0.42) to 5.60 (SD 0.90) in patients with SAD [50] and were 5.80 (SD 0.90) in patients with OCD [51]. Furthermore, these therapeutic alliance scores were comparable with those in a previous study on in-person CBT, where the WAI-SF scores increased from 5.78 (SD 0.94) to 5.93 (SD 0.90) in patients with PD and from 5.32 (SD 0.87) to 5.57 (SD 0.85) in patients with SAD [90]. Considering that lower alliance is known to be associated with dropout [90], the high alliance scores in this study can explain low dropout rate (3%, 1/30).

The dropout rate of this study was 3% (1/30), as 97% (29/30) completed the videoconference-delivered CBT treatment. Dividing the 3 disorders, the dropout rate of 10% (1/10) of participants with SAD was comparable with that of a previous study of videoconference-delivered CBT (17%, 4/24) [50] as well as a previous meta-analysis of 587 studies of in-person CBT between 1990 and 2010 (18%) [91]. The dropout rate of 0% for OCD was comparable with that of a previous study of videoconference-delivered CBT (0%, 0/10) [51] as well as the results of a meta-analysis of studies on in-person CBT published between 1993 and 2014 (15%) [92]. The dropout rate of 0% for PD was similarly comparable with a previous study on videoconference-delivered CBT (0%, 0/11) [52] as well as a meta-analysis of in-person CBT studies published between 1993 and 2002 (12.7%) [93].

Limitations

This study has some limitations, including its small sample size, lack of a control group, unstandardized outcomes (satisfaction/preference of videoconference-delivered CBT), and long-term follow-up. Without a placebo control group and pharmacotherapy group, it remains unknown whether the observed improvements in symptom severity were merely the natural course, a result of the drug, or the effect of the intervention. Future studies should employ psychological placebo conditions and pharmacotherapy conditions. Thus, a 3-armed randomized controlled trial comparing pill placebo as the control group, videoconference-delivered CBT patients on antidepressants, and videoconference-delivered CBT patients who are drug-free should be designed and performed. We have been conducting a randomized controlled trial that includes the pharmacotherapy condition to provide greater insight into this CBT for PD since December 2016. In addition, we intend to conduct similar trials for OCD and SAD in the near future.

Conclusions

This study demonstrated the feasibility of CBT with real-time support by the therapist to remotely treat adult patients with symptoms of obsessive-compulsion or anxiety, examining the reduction in symptoms before and after the intervention, and patient acceptance. As it was found that the mutual relationship between therapists and patients can be built on a high level by patients and that patients felt satisfaction about remote treatment with real-time therapist support via videoconference, we believe that videoconference-delivered CBT can be easily implemented on a larger scale in present Japan where the internet is easily accessible. Future research should aim at increasing the reach of intervention and determining whether the intervention is indeed more approachable to people who are young patients or those with a low socioeconomic status. Related to this matter, because a patient's understanding level and information communication skills probably influence the effectiveness of remote treatment, future studies should be made on designs that consider the contents of support of the therapist beyond the absence or presence of guides.

Acknowledgments

The authors would like to thank Ms Ota and Ms Suwabe, nurses at the Cognitive Behavior Therapy Center, Chiba University Medical School Hospital for the smooth coordination between the therapists and the patients. They also thank Professor Hanaoka, Mr Oki, Ms Yamasaki, and other staff of the Clinical Research Center, Chiba University Medical School Hospital for ensuring the rigor of the study plan. Accurate data management was carried out thanks to Ms Shin and Professor Hanawa in the Data Management Center, Chiba University Medical School Hospital. The authors express their heartfelt appreciation to everyone for their cooperation.

This study was funded by the Japan Agency for Medical Research and Development (AMED; "Medical Digital Revolution Realization Project"; Second Public Inquiry, "Study on Database Infrastructure Construction for Medical Images and Others," aimed at utilization of Information Communication Technology and Artificial Intelligence, Clinical Research Evidence of Distant Psychiatric Medicine to establish guideline through the accumulation of databases and database construction for data utilization; grant 16929962). The AMED played no role in the collection, management, analysis, or interpretation of data and will not have any impact on the trial and the publication of the results.

Conflicts of Interest

None declared.

References

1. Kawakami N, Takeshima T, Ono Y, Uda H, Hata Y, Nakane Y, et al. Twelve-month prevalence, severity, and treatment of common mental disorders in communities in Japan: preliminary finding from the World Mental Health Japan Survey

2002-2003. Psychiatry Clin Neurosci 2005 Aug;59(4):441-452. [doi: 10.1111/j.1440-1819.2005.01397.x] [Medline: 16048450]
2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
3. Kessler RC, Angermeyer M, Anthony JC, Demyttenaere K, Gasquet I, Gluzman S, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176 [FREE Full text] [Medline: 18188442]
4. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. Int J Methods Psychiatr Res 2012 Sep;21(3):169-184 [FREE Full text] [doi: 10.1002/mpr.1359] [Medline: 22865617]
5. Marks IM, Swinson RP, Başoğlu M, Kuch K, Noshirvani H, O'Sullivan G, et al. Alprazolam and exposure alone and combined in panic disorder with agoraphobia. A controlled study in London and Toronto. Br J Psychiatry 1993 Jun;162:776-787. [Medline: 8101126]
6. Clark DM, Salkovskis PM, Hackmann A, Middleton H, Anastasiades P, Gelder M. A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. Br J Psychiatry 1994 Jun;164(6):759-769. [Medline: 7952982]
7. Clark DM, Ehlers A, McManus F, Hackmann A, Fennell M, Campbell H, et al. Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo-controlled trial. J Consult Clin Psychol 2003 Dec;71(6):1058-1067. [doi: 10.1037/0022-006X.71.6.1058] [Medline: 14622081]
8. Foa EB, Liebowitz MR, Kozak MJ, Davies S, Campeas R, Franklin ME, et al. Randomized, placebo-controlled trial of exposure and ritual prevention, clomipramine, and their combination in the treatment of obsessive-compulsive disorder. Am J Psychiatry 2005 Jan;162(1):151-161. [doi: 10.1176/appi.ajp.162.1.151] [Medline: 15625214]
9. Mitte K. A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. J Affect Disord 2005 Sep;88(1):27-45. [doi: 10.1016/j.jad.2005.05.003] [Medline: 16005982]
10. Sousa MB, Isolan LR, Oliveira RR, Manfro GG, Cordioli AV. A randomized clinical trial of cognitive-behavioral group therapy and sertraline in the treatment of obsessive-compulsive disorder. J Clin Psychiatry 2006 Jul;67(7):1133-1139. [Medline: 16889458]
11. Hofmann SG, Smits JA. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. J Clin Psychiatry 2008 Apr;69(4):621-632 [FREE Full text] [Medline: 18363421]
12. Rosa-Alcázar AI, Sánchez-Meca J, Gómez-Conesa A, Marín-Martínez F. Psychological treatment of obsessive-compulsive disorder: a meta-analysis. Clin Psychol Rev 2008 Dec;28(8):1310-1325. [doi: 10.1016/j.cpr.2008.07.001] [Medline: 18701199]
13. Acarturk C, Cuijpers P, van Straten A, de Graaf R. Psychological treatment of social anxiety disorder: a meta-analysis. Psychol Med 2009 Feb;39(2):241-254. [doi: 10.1017/S0033291708003590] [Medline: 18507874]
14. Ougrin D. Efficacy of exposure versus cognitive therapy in anxiety disorders: systematic review and meta-analysis. BMC Psychiatry 2011 Dec 20;11:200 [FREE Full text] [doi: 10.1186/1471-244X-11-200] [Medline: 22185596]
15. Mayo-Wilson E, Dias S, Mavranezouli I, Kew K, Clark DM, Ades AE, et al. Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. Lancet Psychiatry 2014 Oct;1(5):368-376 [FREE Full text] [doi: 10.1016/S2215-0366(14)70329-3] [Medline: 26361000]
16. Öst LG, Havnen A, Hansen B, Kvale G. Cognitive behavioral treatments of obsessive-compulsive disorder: a systematic review and meta-analysis of studies published 1993-2014. Clin Psychol Rev 2015 Aug;40:156-169. [doi: 10.1016/j.cpr.2015.06.003] [Medline: 26117062]
17. Ruskin PE, Silver-Aylaian M, Kling MA, Reed SA, Bradham DD, Hebel JR, et al. Treatment outcomes in depression: comparison of remote treatment through telepsychiatry to in-person treatment. Am J Psychiatry 2004 Aug;161(8):1471-1476. [doi: 10.1176/appi.ajp.161.8.1471] [Medline: 15285975]
18. Sorvaniemi M, Ojanen E, Santamäki O. Telepsychiatry in emergency consultations: a follow-up study of sixty patients. Telemed J E Health 2005 Aug;11(4):439-441. [doi: 10.1089/tmj.2005.11.439] [Medline: 16149889]
19. Nieminen K, Malmquist A, Wijma B, Ryding E, Andersson G, Wijma K. Nulliparous pregnant women's narratives of imminent childbirth before and after internet-based cognitive behavioural therapy for severe fear of childbirth: a qualitative study. BJOG 2015 Aug;122(9):1259-1265 [FREE Full text] [doi: 10.1111/1471-0528.13358] [Medline: 25817045]
20. Mathiasen K, Andersen TE, Riper H, Kleiboer AA, Roessler KK. Blended CBT versus face-to-face CBT: a randomised non-inferiority trial. BMC Psychiatry 2016 Dec 05;16(1):432 [FREE Full text] [doi: 10.1186/s12888-016-1140-y] [Medline: 27919234]
21. Calear AL, Christensen H. Review of internet-based prevention and treatment programs for anxiety and depression in children and adolescents. Med J Aust 2010 Jun 7;192(11 Suppl):S12-S14. [Medline: 20528700]
22. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. J Med Internet Res 2005;7(1):e7 [FREE Full text] [doi: 10.2196/jmir.7.1.e7] [Medline: 15829479]

23. Klein B, Meyer D, Austin DW, Kyrios M. Anxiety online: a virtual clinic: preliminary outcomes following completion of five fully automated treatment programs for anxiety disorders and symptoms. J Med Internet Res 2011;13(4):e89 [FREE Full text] [doi: 10.2196/jmir.1918] [Medline: 22057287]
24. Dear BF, Staples LG, Terides MD, Fogliati VJ, Sheehan J, Johnston L, et al. Transdiagnostic versus disorder-specific and clinician-guided versus self-guided internet-delivered treatment for social anxiety disorder and comorbid disorders: a randomized controlled trial. J Anxiety Disord 2016 Aug;42:30-44 [FREE Full text] [doi: 10.1016/j.janxdis.2016.05.004] [Medline: 27261562]
25. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
26. Andersson E, Enander J, Andrén P, Hedman E, Ljótsson B, Hursti T, et al. Internet-based cognitive behaviour therapy for obsessive-compulsive disorder: a randomized controlled trial. Psychol Med 2012 Oct;42(10):2193-2203 [FREE Full text] [doi: 10.1017/S0033291712000244] [Medline: 22348650]
27. Wootton BM, Dear BF, Johnston L, Terides MD, Titov N. Remote treatment of obsessive-compulsive disorder: a randomized controlled trial. J Obsess-Compuls Relat Disord 2013 Jan 20;2:375-384 [FREE Full text] [doi: 10.1016/j.jocrd.2013.07.002]
28. Bergström J, Andersson G, Ljótsson B, Rück C, Andréewitch S, Karlsson A, et al. Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. BMC Psychiatry 2010;10:54 [FREE Full text] [doi: 10.1186/1471-244X-10-54] [Medline: 20598127]
29. Carlbring P, Nilsson-Ihrfelt E, Waara J, Kollenstam C, Buhrman M, Kaldo V, et al. Treatment of panic disorder: live therapy vs. self-help via the Internet. Behav Res Ther 2005 Oct;43(10):1321-1333. [doi: 10.1016/j.brat.2004.10.002] [Medline: 16086983]
30. Kiropoulos LA, Klein B, Austin DW, Gilson K, Pier C, Mitchell J, et al. Is internet-based CBT for panic disorder and agoraphobia as effective as face-to-face CBT? J Anxiety Disord 2008 Dec;22(8):1273-1284. [doi: 10.1016/j.janxdis.2008.01.008] [Medline: 18289829]
31. Wims E, Titov N, Andrews G, Choi I. Clinician-assisted Internet-based treatment is effective for panic: a randomized controlled trial. Aust N Z J Psychiatry 2010 Jul;44(7):599-607. [doi: 10.3109/00048671003614171] [Medline: 20560847]
32. Andersson G, Carlbring P, Holmström A, Sparthan 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] [Medline: 16881775]
33. Andrews G, Davies M, Titov N. Effectiveness randomized controlled trial of face to face versus Internet cognitive behaviour therapy for social phobia. Aust N Z J Psychiatry 2011 Apr;45(4):337-340. [doi: 10.3109/00048674.2010.538840] [Medline: 21323490]
34. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther 2011 Mar;49(3):158-169. [doi: 10.1016/j.brat.2010.12.007] [Medline: 21255767]
35. Hedman E, Andersson G, Ljótsson B, Andersson E, Rück C, Mörtberg E, et al. Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: a randomized controlled non-inferiority trial. PLoS One 2011;6(3):e18001 [FREE Full text] [doi: 10.1371/journal.pone.0018001] [Medline: 21483704]
36. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
37. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Third Edition. Arlington, VA, United States: American Psychiatric Publishing; 1980.
38. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised. Arlington, VA, United States: American Psychiatric Publishing; 1987.
39. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Arlington, VA, United States: American Psychiatric Publishing; 1994.
40. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Arlington, VA, United States: American Psychiatric Publishing; 2000.
41. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, Ninth Revision. Geneva, Switzerland: World Health Organization; 1979.
42. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision. Geneva, Switzerland: World Health Organization; 1999.
43. 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]
44. Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CEA, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. Clin Psychol Psychother 2013;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]

45. 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]
46. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered internet psychotherapy for depression in primary care: a randomized controlled trial. Lancet 2009 Aug 22;374:628-634 [FREE Full text] [Medline: 19700005]
47. Storch EA, Caporino NE, Morgan JR, Lewin AB, Rojas A, Brauer L, et al. Preliminary investigation of web-camera delivered cognitive-behavioral therapy for youth with obsessive-compulsive disorder. Psychiatry Res 2011 Oct 30;189(3):407-412. [doi: 10.1016/j.psychres.2011.05.047] [Medline: 21684018]
48. Rees CS, Maclaine E. A systematic review of videoconference-delivered psychological treatment for anxiety disorders. Aust Psychol 2015 Jul 20;50(4):259-264 [FREE Full text]
49. Kishimoto T, Tarumi S, Nakamae T, Mimura M. [Exposure response prevention utilizing telemedicine-case series of three cases]. Jpn J Clin Psychiatry 2016 Aug 12;45(12):1603-1609 [FREE Full text]
50. Yuen EK, Herbert JD, Forman EM, Goetter EM, Juarascio AS, Rabin S, et al. Acceptance based behavior therapy for social anxiety disorder through videoconferencing. J Anxiety Disord 2013 May;27(4):389-397. [doi: 10.1016/j.janxdis.2013.03.002] [Medline: 23764124]
51. Vogel PA, Solem S, Hagen K, Moen EM, Launes G, Håland AT, et al. A pilot randomized controlled trial of videoconference-assisted treatment for obsessive-compulsive disorder. Behav Res Ther 2014 Oct 30;63:162-168. [doi: 10.1016/j.brat.2014.10.007] [Medline: 25461792]
52. Bouchard S, Paquin B, Payeur R, Allard M, Rivard V, Fournier T, et al. Delivering cognitive-behavior therapy for panic disorder with agoraphobia in videoconference. Telemed J E Health 2004;10(1):13-25. [doi: 10.1089/153056204773644535] [Medline: 15104911]
53. Sorvaniemi M, Ojanen E, Santamäki O. Telepsychiatry in emergency consultations: a follow-up study of sixty patients. Telemed J E Health 2005 Aug;11(4):439-441. [doi: 10.1089/tmj.2005.11.439] [Medline: 16149889]
54. McGrath JJ, Saha S, Al-Hamzawi A, Andrade L, Benjet C, Bromet EJ, et al. The bidirectional associations between psychotic experiences and DSM-IV mental disorders. Am J Psychiatry 2016 Oct 01;173(10):997-1006 [FREE Full text] [doi: 10.1176/appi.ajp.2016.15101293] [Medline: 26988628]
55. Simpson SG, Reid CL. Therapeutic alliance in videoconferencing psychotherapy: a review. Aust J Rural Health 2014 Dec;22(6):280-299. [doi: 10.1111/ajr.12149] [Medline: 25495622]
56. Mohr DC, Spring B, Freedland KE, Beckner V, Arean P, Hollon SD, et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. Psychother Psychosom 2009;78(5):275-284. [doi: 10.1159/000228248] [Medline: 19602916]
57. 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]
58. Muramatsu K, Miyaoka H, Kamijima K, Muramatsu Y, Yoshida M, Otsubo T, et al. The patient health questionnaire, Japanese version: validity according to the mini-international neuropsychiatric interview-plus. Psychol Rep 2007 Dec;101(3 Pt 1):952-960. [doi: 10.2466/pr0.101.3.952-960] [Medline: 18232454]
59. Nakatani E, Nakagawa A, Nakao T, Yoshizato C, Nabeyama M, Kudo A, et al. A randomized controlled trial of Japanese patients with obsessive-compulsive disorder--effectiveness of behavior therapy and fluvoxamine. Psychother Psychosom 2005;74(5):269-276. [doi: 10.1159/000086317] [Medline: 16088264]
60. Seki Y, Nagata S, Shibuya T, Yoshinaga N, Yokoo M, Ibuki H, et al. A feasibility study of the clinical effectiveness and cost-effectiveness of individual cognitive behavioral therapy for panic disorder in a Japanese clinical setting: an uncontrolled pilot study. BMC Res Notes 2016 Oct 07;9(1):458 [FREE Full text] [doi: 10.1186/s13104-016-2262-5] [Medline: 27717407]
61. Yoshinaga N, Matsuki S, Niitsu T, Sato Y, Tanaka M, Ibuki H, et al. Cognitive behavioral therapy for patients with social anxiety disorder who remain symptomatic following antidepressant treatment: a randomized, assessor-blinded, controlled trial. Psychother Psychosom 2016;85(4):208-217 [FREE Full text] [doi: 10.1159/000444221] [Medline: 27230862]
62. Kobori O, Nakazato M, Yoshinaga N, Shiraishi T, Takaoka K, Nakagawa A, et al. Transporting cognitive behavioral therapy (CBT) and the Improving Access to Psychological Therapies (IAPT) project to Japan: preliminary observations and service evaluation in Chiba. J Ment Health Train Educ Pract 2014;9:155-166 [FREE Full text]
63. Blackburn IM, James IA, Milne DL, Baker C, Standart S, Garland A, et al. The Revised Cognitive Therapy Scale (CTS-R): psychometric properties. Behav Cogn Psychother 2001;29:431-446 [FREE Full text]
64. Shimizu E, Kobori O. Technique and therapeutic effect of Cognitive Therapy Scale-Revised. Jpn J Clin Psychiatry 2012;41(8):969-979.
65. Young JE, Beck AT. Cognitive Therapy Scale: rating manual. Philadelphia, PA: University of Pennsylvania; 1980 Aug 22. URL: https://cdn.ymaws.com/www.academyofct.org/resource/collection/24743CF7-351E-4335-9E93-83F26EF675A3/CTRS_Manual.pdf [accessed 2018-12-02] [WebCite Cache ID 74Mm51FFC]
66. Young JE, Beck AT. Cognitive Therapy Scale. Philadelphia, PA: University of Pennsylvania; 1988.
67. Shabiralyani G, Hasan KS, Hamad N, Iqbal N. Impact of visual aids in enhancing the learning process case research: district Dera Ghazi Khan. J Educ Pract 2015;6:226-233 [FREE Full text]

68. Ministry of Health, Labour and Welfare, Japan. 2016. Guideline on Safety Management of Medical Information System Version 4 URL: https://www.mhlw.go.jp/stf/shingi2/0000119588.html [accessed 2018-12-02] [WebCite Cache ID 74MmGbpmW]
69. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, et al. The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Arch Gen Psychiatry 1989 Nov;46(11):1006-1011. [Medline: 2684084]
70. Hamagaki S, Takagi S, Urushihara Y, Ishisaka Y, Matsumoto M. [Development and use of the Japanese version of the self-report Yale-Brown Obsessive Compulsive Scale]. Seishin Shinkeigaku Zasshi 1999;101(2):152-168. [Medline: 10375975]
71. Houck PR, Spiegel DA, Shear MK, Rucci P. Reliability of the self-report version of the Panic Disorder Severity Scale. Depress Anxiety 2002;15(4):183-185. [doi: 10.1002/da.10049] [Medline: 12112724]
72. Katagami M. The self-report version of the Panic Disorder Severity Scale: reliability and validity of the Japanese version. Jpn J Psychosom Med 2007;47:331-338 [FREE Full text]
73. Liebowitz MR. Social phobia. Mod Probl Pharmacopsychiatry 1987;22:141-173. [Medline: 2885745]
74. Asakura S, Inoue S, Sasaki F, Sasaki Y, Kitagawa N, Inoue T, et al. Consideration of validity and reliability for Liebowitz Social Anxiety Scale (LSAS) Japanese version. Clin Psychiatry 2002;177:1084.
75. Farris SG, McLean CP, Van Meter PE, Simpson HB, Foa EB. Treatment response, symptom remission, and wellness in obsessive-compulsive disorder. J Clin Psychiatry 2013 Jul;74(7):685-690 [FREE Full text] [doi: 10.4088/JCP.12m07789] [Medline: 23945445]
76. Furukawa TA, Katherine SM, Barlow DH, Gorman JM, Woods SW, Money R, et al. Evidence-based guidelines for interpretation of the Panic Disorder Severity Scale. Depress Anxiety 2009;26(10):922-929 [FREE Full text] [doi: 10.1002/da.20532] [Medline: 19006198]
77. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
78. Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ 2002 Jun;11(4):341-353. [doi: 10.1002/hec.673] [Medline: 12007165]
79. Tracey TJ, Kokotovic AM. Factor structure of the Working Alliance Inventory. Psychol Assess 1989;1(3):207-210. [doi: 10.1037/1040-3590.1.3.207]
80. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. J Am Med Assoc 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
81. Muramatsu K. Patient Health Questionnaire (PHQ-9, PHQ-15) Japanese version and Generalized Anxiety Disorder-7 Japanese version?up to date. Stud Clin Psychol 2014;7:35-39 [FREE Full text]
82. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing Generalized Anxiety Disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
83. Muramatsu K, Miyaoka H, Uesima K, Muramatsu Y, Fuse K, Yoshimine F, et al. Examination of validity and usefulness of GAD-7 Japanese version. Jpn Soc Psychosom Med 2010;50(6):592.
84. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
85. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
86. Nakao S, Nakagawa A, Oguchi Y, Mitsuda D, Kato N, Nakagawa Y, et al. Web-based cognitive behavioral therapy blended with face-to-face sessions for major depression: randomized controlled trial. J Med Internet Res 2018 Sep 21;20(9):e10743 [FREE Full text] [doi: 10.2196/10743] [Medline: 30249583]
87. Loerinc AG, Meuret AE, Twohig MP, Rosenfield D, Bluett EJ, Craske MG. Response rates for CBT for anxiety disorders: need for standardized criteria. Clin Psychol Rev 2015 Dec;42:72-82. [doi: 10.1016/j.cpr.2015.08.004] [Medline: 26319194]
88. Payakachat N, Ali MM, Tilford JM. Can The EQ-5D detect meaningful change? A systematic review. Pharmacoeconomics 2015 Nov;33(11):1137-1154 [FREE Full text] [doi: 10.1007/s40273-015-0295-6] [Medline: 26040242]
89. Williams MD, Sawchuk CN, Shippee ND, Somers KJ, Berg SL, Mitchell JD, et al. A quality improvement project aimed at adapting primary care to ensure the delivery of evidence-based psychotherapy for adult anxiety. Br Med J Open Qual 2018;7(1):e000066 [FREE Full text] [doi: 10.1136/bmjoq-2017-000066] [Medline: 29333493]
90. Haug T, Nordgreen T, Öst LG, Tangen T, Kvale G, Hovland OJ, et al. Working alliance and competence as predictors of outcome in cognitive behavioral therapy for social anxiety and panic disorder in adults. Behav Res Ther 2016 Feb;77:40-51. [doi: 10.1016/j.brat.2015.12.004] [Medline: 26708332]
91. Swift JK, Greenber RP. A treatment by disorder meta-analysis of dropout from psychotherapy. J Psychother Integr 2014;24:193-207 [FREE Full text]
92. Öst LG, Havnen A, Hansen B, Kvale G. Cognitive behavioral treatments of obsessive-compulsive disorder. A systematic review and meta-analysis of studies published 1993-2014. Clin Psychol Rev 2015 Aug;40:156-169. [doi: 10.1016/j.cpr.2015.06.003] [Medline: 26117062]

93. Mitte K. A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. J Affect Disord 2005 Sep;88(1):27-45. [doi: 10.1016/j.jad.2005.05.003] [Medline: 16005982]

Abbreviations

AMED: Agency for Medical Research and Development
CBT: cognitive behavioral therapy
DSM: Diagnostic and Statistical Manual of Mental Disorders
EQ-5D: EuroQol-5 Dimension
GAD-7: Generalized Anxiety Disorder-7
ICBT: internet-delivered CBT
LSAS: Liebowitz Social Anxiety Scale
OCD: obsessive-compulsive disorder
PD: panic disorder
PDSS: Panic Disorder Severity Scale
PHQ-9: Patient Health Questionnaire-9
PTSD: posttraumatic stress disorder
SAD: social anxiety disorder
WAI-SF: Working Alliance Inventory-Short Form
Y-BOCS: Yale-Brown Obsessive-Compulsive Scale

Edited by C Eichenberg; submitted 05.09.18; peer-reviewed by K Mathiasen, E Aboujaoude, M Ak; comments to author 26.09.18; revised version received 29.10.18; accepted 11.11.18; published 17.12.18.

Please cite as:
Matsumoto K, Sutoh C, Asano K, Seki Y, Urao Y, Yokoo M, Takanashi R, Yoshida T, Tanaka M, Noguchi R, Nagata S, Oshiro K, Numata N, Hirose M, Yoshimura K, Nagai K, Sato Y, Kishimoto T, Nakagawa A, Shimizu E
Internet-Based Cognitive Behavioral Therapy With Real-Time Therapist Support via Videoconference for Patients With Obsessive-Compulsive Disorder, Panic Disorder, and Social Anxiety Disorder: Pilot Single-Arm Trial
J Med Internet Res 2018;20(12):e12091
URL: http://www.jmir.org/2018/12/e12091/
doi:10.2196/12091
PMID:30559094

Original Paper

The Effectiveness of a Web-Based Computer-Tailored Physical Activity Intervention Using Fitbit Activity Trackers: Randomized Trial

Corneel Vandelanotte[1], PhD; Mitch J Duncan[2], PhD; Carol A Maher[3], PhD; Stephanie Schoeppe[1], PhD; Amanda L Rebar[1], PhD; Deborah A Power[4], MedSc; Camille E Short[5], PhD; Christopher M Doran[4], PhD; Melanie J Hayman[1], PhD; Stephanie J Alley[1], PhD

[1]Physical Activity Research Group, Appleton Institute, School of Health, Medical and Applied Sciences, Central Queensland University, Rockhampton, Australia

[2]School of Medicine & Public Health, Faculty of Health and Medicine, The University of Newcastle, Newcastle, Australia

[3]School of Health Sciences, Alliance for Research in Exercise, Nutrition and Activity, University of South Australia, Adelaide, Australia

[4]School of Health, Medical and Applied Sciences, Central Queensland University, Rockhampton, Australia

[5]School of Medicine, Freemasons Foundation Centre for Men's Health, University of Adelaide, Adelaide, Australia

Corresponding Author:
Corneel Vandelanotte, PhD
Physical Activity Research Group, Appleton Institute
School of Health, Medical and Applied Sciences
Central Queensland University
Building 7, Bruce Highway
Rockhampton, 4700
Australia
Phone: 61 749232183
Email: c.vandelanotte@cqu.edu.au

Abstract

Background: Web-based interventions that provide personalized physical activity advice have demonstrated good effectiveness but rely on self-reported measures of physical activity, which are prone to overreporting, potentially reducing the accuracy and effectiveness of the advice provided.

Objective: This study aimed to examine whether the effectiveness of a Web-based computer-tailored intervention could be improved by integrating Fitbit activity trackers.

Methods: Participants received the 3-month *TaylorActive* intervention, which included 8 modules of theory-based, personally tailored physical activity advice and action planning. Participants were randomized to receive the same intervention either with or without Fitbit tracker integration. All intervention materials were delivered on the Web, and there was no face-to-face contact at any time point. Changes in physical activity (Active Australia Survey), sitting time (Workforce Sitting Questionnaire), and body mass index (BMI) were assessed 1 and 3 months post baseline. Advice acceptability, website usability, and module completion were also assessed.

Results: A total of 243 Australian adults participated. Linear mixed model analyses showed a significant increase in total weekly physical activity (adjusted mean increase=163.2; 95% CI 52.0-274.5; *P*=.004) and moderate-to-vigorous physical activity (adjusted mean increase=78.6; 95% CI 24.4-131.9; *P*=.004) in the Fitbit group compared with the non-Fitbit group at the 3-month follow-up. The sitting time and BMI decreased more in the Fitbit group, but no significant group × time interaction effects were found. The physical activity advice acceptability and the website usability were consistently rated higher by participants in the Fitbit group. Non-Fitbit group participants completed 2.9 (SD 2.5) modules, and Fitbit group participants completed 4.4 (SD 3.1) modules.

Conclusions: Integrating physical activity trackers into a Web-based computer-tailored intervention significantly increased intervention effectiveness.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12616001555448; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=371793 (Archived by WebCite at http://www.webcitation.org/73ioTxQX2)

(J Med Internet Res 2018;20(12):e11321) doi:10.2196/11321

KEYWORDS

online, internet, tracking, health behavior change, advanced activity trackers, wearables

Introduction

Background

Regular physical activity is recommended to reduce the risk of developing chronic disease (eg, diabetes, cardiovascular disease, and cancer), mental health problems, mortality, and morbidity [1,2]. Unfortunately, in Australia, and in most other developed and developing nations, the majority of the population is not meeting the physical activity recommendations [1,3]. This causes a large burden of disease, reduced quality of life, and high health care costs [2,4]. As such, the search for cost-effective interventions that can effectively increase physical activity levels in large populations is ongoing [5].

In this regard, Web-based computer-tailored interventions have demonstrated promising outcomes. Computer-tailored interventions aim to mimic face-to-face interactions with health professionals and provide highly detailed and personally relevant behavior change information [6,7]. However, unlike face-to-face experience, they have a wide reach with access to unlimited numbers of Web users at low cost [6,7]. Personalized physical activity advice is provided after participants complete 1 or more Web-based surveys. On the basis of participant responses and using IF-THEN algorithms (eg, IF not meeting activity guideline, THEN provide advice to increase activity levels), relevant feedback is selected from a large database with all possible response options [8]. Although a systematic review found that 80% of studies that provided Web-based personalized physical activity advice reported positive results at 3 months, the effect sizes were relatively small and less than half of the studies (47%) found significant effects 6 months after starting the study, meaning that intervention effects are not maintained [6].

As such, there is scope to improve the effectiveness of computer-tailored interventions. An important limitation is that they depend on Web-based self-report physical activity measures to generate personalized advice. It is well known that many people overestimate their self-reported activity levels by a large margin [9]. For example, an Australian study showed that 24% of the general population (and up to 58% in certain subgroups) overreported their activity levels [9]. Inaccurate self-reported physical activity can lead to participants being provided with incorrect advice [10]. For example, because of overreporting, someone might receive the message that they are meeting the activity guidelines and do not need to become more active, when this is actually not the case. When this happens, the intervention is not providing accurate and credible advice to participants and will, therefore, not be as effective as it could be [10,11]. Therefore, new techniques to increase the effectiveness of computer-tailored interventions are needed.

The proliferation of sophisticated activity trackers (eg, Fitbit) provides a unique opportunity to improve the effectiveness of computer-tailored interventions. These advanced activity trackers can measure steps, heart rate, energy expenditure, sleep, sedentary behavior, and physical activity intensity (ie, light, moderate, or vigorous intensity) [12]. Furthermore, they allow for automated data uploads to websites or apps via a wireless connection. As such, these activity trackers can objectively and accurately assess physical activity through continuous monitoring [13]. The data generated by these activity trackers can then conveniently and seamlessly be integrated into computer-tailored advice without the burden of repeated Web-based surveys, thus increasing the potential for providing computer-tailored advice that is more credible and effective when compared with using less reliable self-reports [11]. Moreover, replacing the Web-based surveys by activity trackers may lead to greater intervention adherence, as participants in previous computer-tailored studies have systematically reported that there are too many questions that need to be answered before they receive their personalized advice [14,15].

Objectives

Therefore, the objective of this 2-group randomized trial was to examine whether a Web-based computer-tailored intervention using Fitbit activity trackers to generate personalized feedback is more effective in increasing physical activity and engaging participants compared with a computer-tailored intervention using traditional self-reports.

Methods

Procedures and Participants

Participants were recruited across Australia using random digit dialing (conducted by the Population Research Lab at Central Queensland University [CQUniversity]), Facebook advertisements, flyers, posters, word-of-mouth, and email lists (ie, people who signed up to the Web-based 10,000 Steps program [only those who had not used the program for at least 12 months were invited], CQUniversity alumni). Those interested were directed to a landing page on the intervention website to complete a screening tool that determined eligibility. Eligible participants were aged 18 years or above, living in Australia, had a smartphone and computer with internet access, scored 2 or more out of 5 on the Internet Self-Confidence Scale [16], able to speak and read English, had a body mass index (BMI) between 25 and 40, engaged in less than 150 min per week of moderate-to-vigorous physical activity [17,18], had no prior experience in using an activity tracker, had not participated in a physical activity intervention within the last 12 months, and were able to safely increase physical activity assessed through the physical activity readiness questionnaire (PAR-Q) [19]. Those not meeting PAR-Q standards were instructed to obtain medical clearance before participation was allowed.

After completing the Web-based screening tool, eligible participants completed Web-based baseline surveys (see Measures section below). After completing baseline assessments, participants were randomized into 1 of the 2 groups

in a ratio of 1:1 using a random list generator and provided with access to the *TaylorActive* intervention (see Intervention section below). All participants received access to the *TaylorActive* intervention; however, only 1 group (the *Fitbit* group) received a Fitbit activity tracker to monitor physical activity objectively, and the other group (the *Non-Fitbit* group) did not. In the *Fitbit* group, participants were posted a Fitbit Flex, along with instructions on how to use it and sync data from the Fitbit to the *TaylorActive* website. They only received access to the *TaylorActive* intervention 7 days following receipt of the Fitbit, so that it could collect physical activity data that could then be immediately synced with the *TaylorActive* website upon first use. Access was not delayed in the non-Fitbit group, as participants were able to self-report the last week of activity immediately. Follow-up measures were assessed 1 and 3 months post baseline. Participants in both groups received up to 3 reminder emails and 2 phone calls/text messages when they did not complete the surveys within the desired time frame. There was no face-to-face contact with participants throughout the entire duration of this study; all procedures were Web-based, via phone or postal mail. Participants who complied with all study procedures received an Aus $50 incentive for their participation; those in the Fitbit group were able to decline the incentive in exchange for keeping the Fitbit they received (they were only informed about this option at the end of the study).

All participants provided informed consent, ethical approval was obtained from the CQUniversity Human Ethics Committee (H1608-227), and the trial was registered at the Australian New Zealand Clinical Trails Registry (ACTRN12616001555448). All data were collected and analyzed in 2016 and 2017.

Intervention

Participants in both groups received access to a computer-tailored physical activity intervention named *TaylorActive* [20]. The behavior change content of this intervention was developed in line with the theory of planned behavior [21], self-determination theory [22], and social cognitive theory [23]. Specifically, content was focused on enhancing intrinsic motivation, self-efficacy, and intentions for increasing activity levels. In addition, training was provided on self-regulatory strategies to enhance the enactment of intentions into behavior through effective goal-setting, action planning, use of social support, overcoming barriers, problem solving, decision making, relapse prevention, and self-monitoring [21-23].

On the basis of short Web-based surveys, participants in both groups were provided with behavior change content across 8 modules of personal physical activity advice delivered over a 3-month period. The first 4 modules were delivered weekly; the next 4 modules were delivered every 14 days. The 8 modules were organized in a set order and the *next* module could only be accessed when the previous module was completed. All modules were released at a set time point based on participants' study start date. If participants did not access newly available modules, they received up to 3 reminder emails. To generate the personalized module content in the non-Fitbit group, participants were asked questions about how active they have been the previous week in conjunction with questions relating to individual, social, environmental, and theory-based correlates of physical activity behavior. On the basis of the answers of participants, and through applying IF-THEN algorithms, personally relevant physical activity content was automatically selected from a database. In the first session, participants were asked to select their preference of 1 of 5 motivations to be physically active: (1) to improve or maintain good health, (2) to increase fitness, (3) to increase strength, (4) to lose weight, or (5) to feel better (improve mood and/or reduce stress). The feedback and physical activity goals were tailored according to participants' preferred motivation.

The only difference between groups was the way in which physical activity was assessed to provide personalized advice for the 8 modules. In the *non-Fitbit* group, participants completed an adapted version of the *Godin-Shephard Leisure-time Exercise Questionnaire* at the start of each module [24]. In the *Fitbit* group, physical activity was assessed using a Fitbit Flex (this device does not have a display other than 5 tiny LEDs; 1 LED illuminates for every 2000 steps taken; this device does not nudge or buzz or beep when participants have not been active for a while). Participants only needed to click 1 button on the *TaylorActive* website at the start of each module to import physical activity data collected using the Fitbit. The physical activity advice was structured in the same way for both groups, as equivalent variables were extracted from both assessment methods (light, moderate, vigorous, and total physical activity).

Participants in both groups also had access to a *Library* with generic educational information about physical activity; a total of 19 brief articles were available about different aspects of physical activity and what to do to increase physical activity levels (eg, "Are you physically fit?," "Getting motivated," and "Making time to be active"). Finally, participants in both groups were encouraged to complete an action plan at the end of each module [20]. Action plans are self-regulation strategies in the form of setting up a detailed plan that can lead to better goal attainment and help in behavior modification [25]. Practically, it meant that participants were asked very specific questions on how they would meet their activity goals: what activity they would do, where they would do it, when they would do it, how often they would do it, how long will each activity session be, and with whom they would do it. At the start of creating an action plan, participants were asked to set long-, medium-, and short-term goals to reach their physical activity objectives.

More in-depth details about this intervention can be found in the protocol paper for a different trial, only the "Intervention" section (starting on page 3) from that paper is relevant for the study described here [20]. As outlined in this protocol paper, there are in fact 2 versions of the *TaylorActive* intervention, 1 version in which all personalized feedback is provided as text on a webpage and the other version where feedback is delivered through personalized videos. As the main *TaylorActive* trial is still ongoing, it was unknown at the time of this study which version was more effective. As such, participants in this study were equally randomized to text and video versions. Any effects caused from these different versions were controlled for in the statistical analysis. Discussing the impact of the different versions of the *TaylorActive* intervention is outside the scope of this paper.

Measures

Basic demographic factors were assessed: sex, age, years of education, income (≤Aus $51,999; Aus $52,000-Aus $99,999; ≥Aus $100,000; don't know or no response), employment status (full-time, part-time or casual, unemployed), height (centimeters), and weight (kilograms). Height and weight measures were used to calculate BMI of participants.

The 8-item *Active Australia Survey* was used to measure changes in physical activity (please note: the Godin-Shephard Leisure-time Exercise Questionnaire was only used to provide participants in the non-Fitbit group with personalized activity advice; it was not used to assess study outcomes). This survey assesses frequency and duration of walking for transport, walking for recreation, moderate intensity physical activity, and vigorous intensity physical activity [26]. Total physical activity was calculated by summing the time spent in walking, moderate activity, and vigorous activity (weighted by 2) according to specified scoring guidelines [26]. Moderate-to-vigorous physical activity was also calculated and did not include the time spent walking. The Active Australia Survey has acceptable test-retest reliability (intraclass coefficient=.64) and validity (r=.61) in the Australian adult population and has been documented as a useful evaluative tool for detecting intervention-related change in physical activity [27,28].

Sitting time was measured using the 10-item *Workforce Sitting Questionnaire* [29]. Participants reported time (hours or minutes) spent sitting on usual working and nonworking days in relation to work, transport, television use, computer use, and other leisure time sitting. One question also assessed the number of days participants usually work in a week. Total sitting time was defined as the sum of sitting time in all domains for all days. This questionnaire has demonstrated adequate test-retest reliability and validity [29].

The acceptability of the physical activity advice, website usability, and Fitbit use were also assessed [14]. These questions were based on previously published work where advice acceptability of similar interventions was assessed [14]. Finally, module completion was tracked objectively through the intervention website.

Statistics

Analyses were conducted using SPSS version 24. Descriptive statistics of participants' demographics, total physical activity, moderate-to-vigorous physical activity, total sitting time, and BMI at baseline are presented. Group (Fitbit and non-Fitbit) comparisons were conducted using t tests for continuous variables and chi-square analyses for categorical variables. To test for a group (Fitbit or non-Fitbit) by time (baseline, 1 month, and 3 months) interaction on total weekly physical activity, a linear mixed model analysis was conducted. In total, 3 more separate linear mixed model analyses were conducted to test a group by time interaction effects on moderate-to-vigorous physical activity, sitting time, and BMI. All linear mixed model analyses applied restricted maximum likelihood estimation to reduce risk of bias from missing data [30]. All linear mixed model analyses were adjusted for age, sex, education, employment, income, version of the *TaylorActive* intervention (video or text), and BMI (with exception of the model what was examining BMI itself). The non-Fitbit group was the reference variable for group, and baseline was the reference variable for time.

Results

A total of 243 participants were randomized (see Figure 1 for participant flow). The majority of participants were female (182/243, 74.9%), employed full-time (129/243, 53.1%), and earned a yearly income over Aus $51,000 (179/243, 61.0%). The average age, BMI, and years of education were 51.5, 31.2, and 14.8, respectively. At baseline, participants engaged in 106.8 min per week of total physical activity and 36.6 min per week of moderate-to-vigorous physical activity; average daily sitting time was 10 hours a day. There were no between-group differences at baseline. Significantly more participants in the non-Fitbit group did not complete assessments at 1 month (57% vs 35%; χ^2_1=12.5; P<.001) and at 3 months (63% vs 36%; χ^2_1=17.4; P<.001) compared with the Fitbit group. Participant characteristics are reported in Table 1.

There were significant time effects at 1 and 3 months for both groups for total physical activity and also a significant time by group interaction at 3 months (adjusted mean difference=163.2 min; 95% CI 52.0-274.5; P=.004) though not at 1 month (see Table 2 and Figure 2). Total physical activity increased 119.3 min per week in the non-Fitbit group and 284.9 min per week in the Fitbit group at 3 months. Similarly, significant time effects were observed at 1 and 3 months for moderate-to-vigorous physical activity as well as a significant time by group interaction at 3 months (adjusted mean difference=78.6 min; 95% CI 24.4-131.9; P=.004) but again not at 1 month. Total moderate-to-vigorous physical activity increased 38.3 min per week in the non-Fitbit group and 117.2 min per week in the Fitbit group at 3 months. Although there was a significant time effect for sitting time in the Fitbit group at 3 months, no other statistically significant time effects or interaction effects were found. Sitting was, on average, reduced by 56 min per day in the non-Fitbit group and 101 min per day in the Fitbit group at 3 months. For BMI, significant time effects were found at both time points for the non-Fitbit group but only at 3 months for the Fitbit group; no interaction effects were observed. BMI was reduced by 1.07 in the non-Fitbit group and 1.54 in the Fitbit group.

Figure 1. Participant flowchart. CQUniversity: Central Queensland University, PAR-Q: physical activity readiness questionnaire, BMI: body mass index.

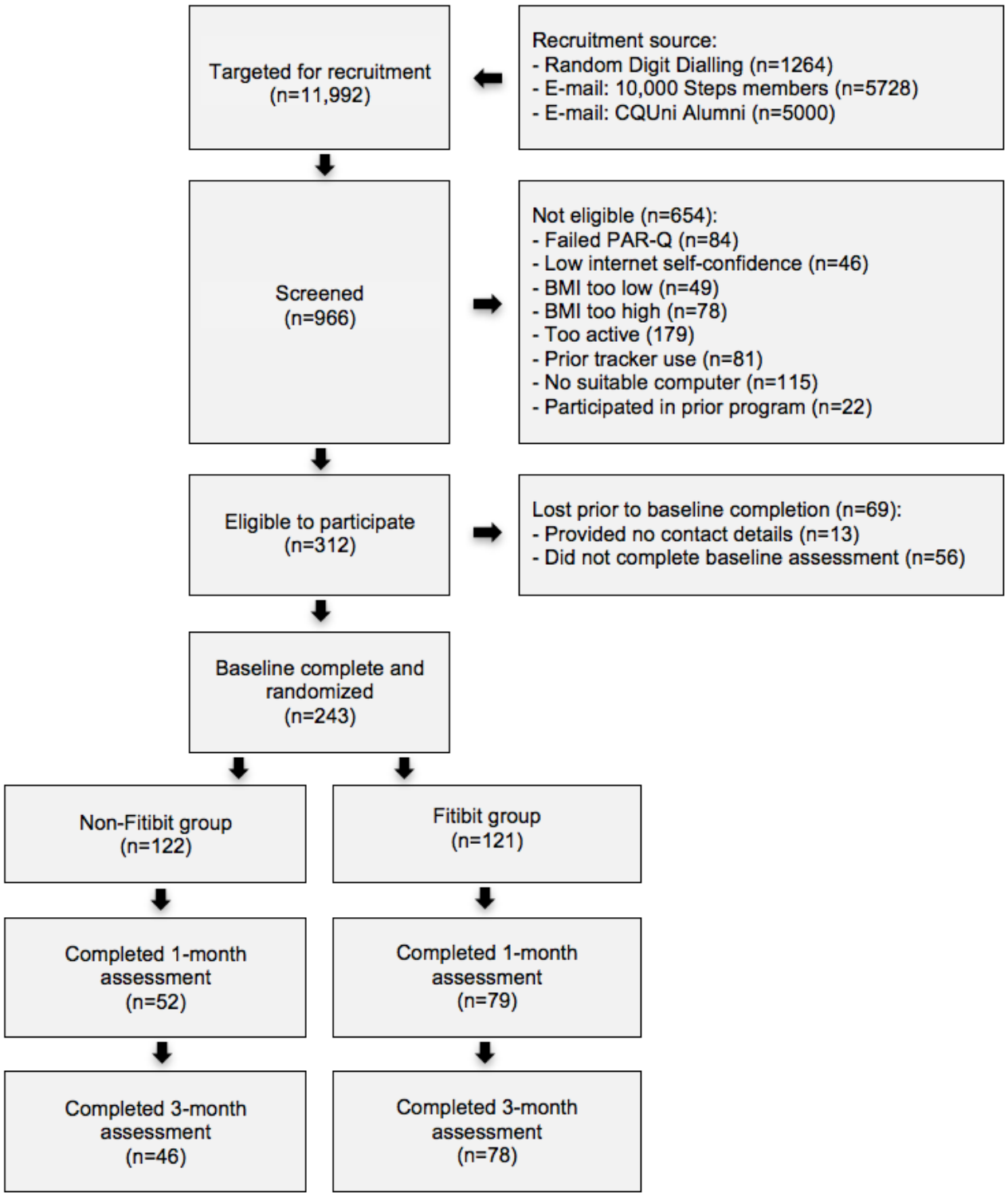


Table 1. Baseline participant characteristics as well as physical activity, moderate-to-vigorous physical activity, sitting time, and body mass index at all time points.

Baseline characteristics	All participants (N=243)	Non-Fitbit (n=122)	Fitbit (n=121)	*P* value[a]
Sex, n (%)				
Male	61 (25.1)	29 (23.8)	32 (26.4)	.63
Female	182 (74.9)	93 (76.2)	89 (73.6)	—[b]
Age in years, mean (SD)	51.5 (11.1)	51.5 (10.6)	51.6 (11.6)	.94
Education in years, mean (SD)	14.8 (3.4)	14.4 (3.0)	15.1 (3.7)	.09
Employment, n (%)				
Full time	129 (53.1)	67 (54.9)	62 (51.2)	.33
Part-time or casual	54 (22.2)	30 (24.6)	24 (19.8)	—
Other	60 (24.7)	25 (20.5)	35 (29.0)	—
Income, n (%)				
≤Aus $51,999	64 (26.3)	33 (27.0)	31 (25.6)	.14
Aus $52,000-Aus $99,999	81 (33.3)	46 (37.7)	35 (28.9)	—
≥Aus $100,000	67 (27.6)	33 (27.0)	34 (28.1)	—
Don’t know or no response	31 (12.8)	10 (8.2)	21 (17.4)	—
Recruitment source, n (%)				
10,000 steps database	79 (32.5)	41 (33.6)	38 (31.4)	.54
Population research lab	79 (32.5)	41 (33.6)	38 (31.4)	—
Facebook ads	28 (11.5)	16 (13.1)	12 (9.9)	—
Central Qqueensland University alumni database	17 (7.0)	6 (4.9)	11 (9.1)	—
Other	40 (16.4)	18 (14.8)	22 (18.1)	—
Body mass index, mean (SD)				
At baseline	31.2 (4.5)	31.1 (4.7)	31.4 (4.4)	.63
At 1 month	30.6 (4.3)	30.4 (4.5)	30.7 (4.2)	—
At 3 months	30.0 (4.5)	30.1 (4.6)	29.9 (4.4)	—
Total physical activity in minutes per week, mean (SD)				
At baseline	106.8 (147.4)	110.7 (150.7)	102.8 (144.4)	.67
At 1 month	300.1 (306.4)	250.2 (293.4)	333.0 (312.1)	—
At 3 months	329.2 (324.0)	230.0 (164.1)	387.7 (377.7)	—
Moderate-to-vigorous physical activity in minutes per week, mean (SD)				
At baseline	36.6 (76.5)	41.5 (80.4)	31.6 (72.4)	.31
At 1 month	109.3 (164.8)	87.3 (146.5)	123.8 (175.2)	—
At 3 months	123.2 (154.3)	79.8 (77.1)	148.8 (181.1)	—
Total sitting time in hours per day, mean (SD)				
At baseline	10.0 (3.6)	10.1 (3.3)	9.9 (3.8)	.59
At 1 month	9.3 (3.7)	9.2 (3.5)	9.3 (3.9)	—
At 3 months	8.6 (4.2)	9.2 (3.6)	8.2 (4.5)	—

[a]The *P* values reported are the outcomes of *t* tests (continuous variables) or chi-square tests (categorical variables) and only relate to comparing Fitbit and non-Fitbit groups at baseline (hence, no *P* values are reported for 1- and 3-month outcomes).

[b]Not applicable.

Table 2. Linear mixed models analysis comparing change in total physical activity, moderate-to-vigorous physical activity, sitting time, and body mass index between Fitbit and non-Fitbit groups at 1 and 3 months adjusted for baseline levels.

Characteristics[a]	Time-effects				Time by group interaction-effects (reference=non-Fitbit group)	
	Fitbit group		Non-Fitbit group		Adjusted mean difference from baseline (95% CI)	*P* value
	Adjusted mean difference from baseline (95% CI)	*P* value	Adjusted mean difference from baseline (95% CI)	*P* value		
Total physical activity (weekly minutes)						
1 month	222.93 (154.98 to 290.87)	<.001	152.00 (80.04 to 223.96)	<.001	77.89 (−23.30 to 179.07)	.13
3 months	270.12 (188.86 to 351.36)	<.001	110.24 (56.39 to 164.10)	<.001	163.26 (52.03 to 274.50)	.004
Moderate-to-vigorous physical activity (weekly minutes)						
1 month	89.59 (50.64 to 128.53)	<.001	50.90 (19.21 to 82.60)	.002	38.37 (−16.02 to 92.77)	.17
3 months	110.46 (72.38 to 148.54)	<.001	31.13 (5.02 to 57.24)	.002	78.65 (25.40 to 131.89)	.004
Sitting (daily minutes)						
1 month	−34.33 (−88.61 to 19.94)	.21	−40.20 (−99.38 to 18.98)	.18	8.58 (−71.8 to 88.98)	.83
3 months	−103.72 (−156.68 to −50.75)	<.001	−31.90 (−83.32 to 19.51)	.22	−70.10 (−147.74 to 7.53)	.08
Body mass index						
1 month	−0.20 (−0.41 to 0.01)	.06	−0.44 (−0.72 to −0.16)	.002	0.23 (−0.12 to 0.57)	.18
3 months	−0.72 (−1.04 to 0.40)	<.001	−0.62 (−1.03 to 0.21)	.004	−0.12 (−0.63 to 0.40)	.66

[a]Linear mixed models included all participants at all time points, as such N=243 for all analyses. Analyses were adjusted for age, gender, education, employment status, income, body mass index (BMI; the BMI model was not adjusted for BMI), and video or text advice. The reference variable for time was the baseline measure, and the reference variable for group was the non-Fitbit group.

Figure 2. Total physical activity and moderate-to-vigorous physical activity at baseline, 1 month and 3 months.

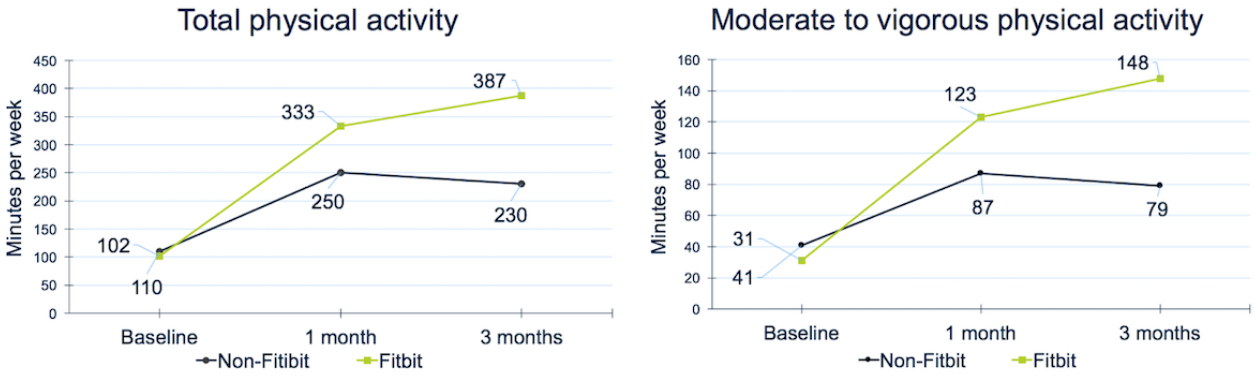


Table 3 presents outcomes on user acceptability of the advice, intervention website, and Fitbit. Both the physical activity advice acceptability and the website usability were consistently rated higher (though not always significantly higher) by participants in the Fitbit group. In terms of advice acceptability, significant differences were found for the questions *there were too many questions to access the advice,the advice taught me something new about my physical activity,* and *I shared the advice with others.* In terms of website usability, significant differences were found for the questions *I want to continue to use the website,the website is easy to use,* and *I used the website once per week or more.* The Fitbit group also indicated the use of the Fitbit itself was favorable and augmented the personal advice delivered through the website. For example, participants indicated (*agreed* or *strongly agreed*) that the value of the tailored advice was increased (74.4%), that the advice was more credible (67.9%), and more personally relevant (76.9%). The majority of participants (85.9%) thought it was easy to sync Fitbit data with the TaylorActive website.

Figure 3 demonstrates how much exposure participants had to the intervention content. A higher percentage of participants in the Fitbit group completed each module except the first one. Double the proportion of participants completed the final module in the Fitbit group compared with the non-Fitbit group (27.3% vs 13.9%). On average, non-Fitbit group participants completed 2.9 (SD 2.5) modules and Fitbit group participants completed 4.4 (SD 3.1) modules.

Table 3. Physical activity intervention acceptability, website use, and Fitbit use.

Acceptability and usability questions	Non-Fitbit (n=46)	Fitbit (n=78)	*P* value[a]
Advice acceptability (% agreed or agreed strongly)[b]			
Did you view all the advice	45.7	70.5	.16
There were too many questions to access the advice	30.5	16.7	.02
I changed my opinion about being active	28.3	46.1	.15
The tailored advice was credible	80.7	87.0	.36
The advice taught me something new about my physical activity	41.3	65.4	.03
Too much advice was provided per module	15.2	12.8	.44
The tailored advice helped me reach my goals	41.3	51.3	.74
I shared the advice with others	2.2	19.2	.006
Website usability (% agreed or agreed strongly)			
I want to continue to use the website	48.1	81.0	.003
The website is easy to use	67.3	82.3	.02
I like the presentation of the website (layout, colors)	57.7	68.3	.48
I used the website once per week or more	50.0	71.0	<.001
Fitbit usability[c] (% agreed or agreed strongly)			
The Fitbit improves the value of the tailored advice	—[d]	74.4	—
The Fitbit improves the credibility of the tailored advice	—	67.9	—
The Fitbit improves the personal relevance of the tailored advice	—	76.9	—
The Fitbit improves the user-friendliness of the tailored advice	—	69.3	—
It was easy to sync data between Fitbit and the intervention website	—	85.9	—
I wore the Fitbit every day during the study	—	73.1	—
The Fitbit helps me to increase my physical activity	—	83.5	—
I would like to continue using the Fitbit	—	91.2	—
The Fitbit is easy to use	—	96.2	—
The Fitbit is comfortable to wear	—	83.5	—

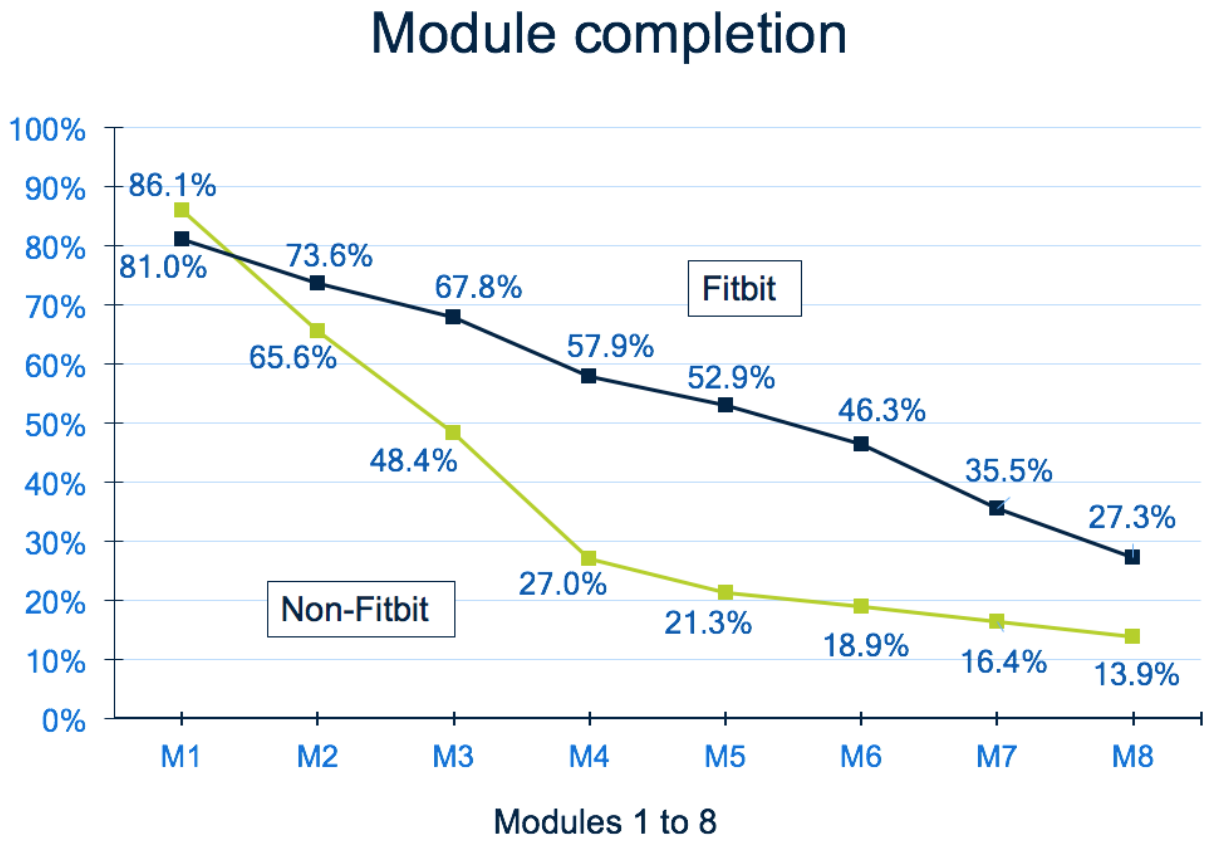
[a]The *P* values reported are the outcomes of *t* tests.

[b]All questions were scored on a 5-point Likert scale, only the sum of participants who *Agreed* or *Strongly Agreed* with each statement is presented in the table.

[c]Only participants in the Fitbit group were asked questions about Fitbit use.

[d]Not applicable.

Figure 3. Average module completion for the Fitbit and non-Fitbit group for each of the 8 available modules.



Discussion

Main Outcomes

The main aim of this study was to examine whether integrating a Fitbit activity tracker into a computer-tailored physical activity intervention increased the effectiveness of the intervention. The study findings clearly support the integration of activity trackers into a Web-based physical activity intervention that provides participants with personalized advice. Total physical activity increased more than twice as much in the Fitbit group, compared with the non-Fitbit group, and moderate-to-vigorous physical activity increased nearly 3 times as much at 3 months. The lack of significant interaction effects at 1 month may be explained by participants not having received all intervention content at this stage. These findings indicate that it takes some time to change behavior, and physical activity levels were still increasing at that point in time (see Figure 2).

To date, only a few other studies have examined the use of activity trackers (ie, mostly traditional pedometers) in combination with computer-tailored advice [10,11,31]. However, none of these trials directly compared the effectiveness of a computer-tailored intervention with and without activity trackers. For example, Compernolle et al [11] demonstrated the effectiveness of step-based computer-tailored advice that used pedometers but compared this with a no intervention control group. Another study by De Cocker et al [10] compared pedometer-based computer-tailored advice with a pedometer-only group; although the group that also received the tailored advice intervention increased their activity somewhat more than the pedometer-only group, the difference was not significant. Finally, Slootmaker et al [31,32] compared activity tracker–based physical activity advice with a usual care control group and did not see improvements in physical activity in either groups. Although innovative at the time (before the proliferation of smartphone), this study may have been ahead of its time, and the acceptability and user-friendliness of the technology may have been low. The use of smartphones and advanced activity trackers is now commonplace, and the technology is generally well designed and accepted, which may explain the better results in our study. This is confirmed by the strong acceptability outcomes observed in this study. All components of the intervention (advice acceptability, website usability, and Fitbit usability) were rated more highly in the Fitbit group compared with the non-Fitbit group. Remarkably, even the design of the intervention was rated higher in the Fitbit group, despite being identical across groups. The syncing of Fitbit data also received high ratings, despite first having to sync data with the Fitbit platform (this can happen automatically depending on app and phone settings) before being able to sync with the computer-tailored intervention. The impact of the Fitbit integration is also demonstrated in terms of module completion, with twice as many participants completing all computer-tailored modules in the Fitbit group.

Although the intervention did not focus on reducing sitting time (nor did the Fitbit buzz as a prompt for prolonged sitting), substantial reductions in sitting time were observed; a significant time effect at 3 months was found for the Fitbit group, which

reduced sitting by almost 12 hours per week. Many other physical activity interventions have also examined the impact on sitting time [11,33,34], and most of these studies show little to no effects on sitting time. Similarly, although the overall intervention did not focus on reducing weight nor included a diet component, substantial BMI reductions were found, with significant time effects in both groups. However, weight loss was the most popular motivation among participants for becoming more active, and a large proportion of participants did select this option (37.1%, data not reported in the Results section). For these participants only, the personalized physical activity advice they received incorporated a weight loss focus though recommending higher activity levels (no dietary advice was provided). Nevertheless, this finding is remarkable as weight loss interventions without a dietary component are often not very effective [35,36].

Strengths and Limitations

Despite the significant findings and the novelty of the study, several study limitations should be noted; as such, the study findings should be interpreted with some caution. First, the study did not have a control group or a tracker-only group; it is possible that outcomes in the Fitbit group are because of the Fitbit itself, and not because of the combined intervention. A more robust study design (including a *Fitbit-only* group) is needed to clarify this and disentangle these effects. On the other hand, higher website usability and acceptability in the Fitbit group suggests the computer-tailored website was genuinely contributing to the increase in physical activity, as participants could have chosen to only use the Fitbit and ignore the computer-tailored website, but rather they used it more than participants who did not receive a Fitbit. Second, the intervention groups were small and dropout was high. However, the posthoc power calculations demonstrated sufficient power to detect significant between-group differences for total physical activity (89.3%) and moderate-to-vigorous physical activity (83.7%) at the 3-month time point. The total lack of face-to-face interaction with participants (thus, low accountability), may have contributed to the high levels of dropout [37,38]. High dropout rates are common in Web-based interventions [39,40]. It was nevertheless interesting to observe that just providing participants with a Fitbit significantly increased retention. Many intervention studies have found higher dropout in intervention groups (or higher intensity intervention groups) compared with control groups because of the additional burden of actively participating and trying to improve health behavior [15]; this did not apply to our study. Third, although the Fitbit objectively assessed physical activity, we were not able to use Fitbit data to assess change over time as only 1 intervention group was provided with a Fitbit. Budgetary constraints meant we had to rely on a self-report measure to assess change over time, and although the Active Australia Survey has demonstrated it can detect change over time [28], the findings should be interpreted with caution. As the introduction points out, self-report physical activity measures are prone to overreporting [9]; however, in theory, the measurement error should be consistent across groups, so it is likely that the difference between groups is real, but the magnitude of the outcomes is less certain. Fourth, there was no longer-term follow-up to assess changes in behavioral outcomes. The 3-month assessment was immediately after the end of the intervention delivery, so behavior change maintenance effects and differences between groups could not be tested. Maintenance of physical activity improvements has been very difficult to achieve, with the majority of studies showing declines in activity levels after the intervention has finished [41,42]. Finally, although the accuracy and validity of commercial consumer-level activity trackers are high, there is room for improvement [12]. As such, in a small number of participants, the personalized advice generated using Fitbits may still have been somewhat inaccurate and indicated they were meeting guidelines when they were not in reality. Therefore, manufacturers are encouraged to continue to improve the quality of the devices, and researchers are encouraged to continue to assess their accuracy in validity studies.

Conclusions

In conclusion, integrating physical activity trackers into a Web-based computer-tailored intervention significantly increased intervention effectiveness in overweight or obese participants. Due to the technology-based nature of this intervention, it is possible to reach a large number of people at an acceptable cost and improve their physical activity behavior. As such, the potential of combining advanced activity trackers with sophisticated computer-tailored interventions is large. However, given the study limitations, follow-up studies with more robust designs (objective outcome measures and longer-term follow-up including control and tracker-only groups) are needed to confirm these outcomes.

Acknowledgments

The study was funded through a Central Queensland University infrastructure grant and through support funds as part of a National Heart Foundation of Australia Future Leader Fellowship (ID 100427). CV (ID 100427), MJD (ID 100029), and SS (ID 101240) were supported by National Heart Foundation of Australia Fellowships. CAM (ID 1125913), ALR (ID 1105926), SS (ID GNT1125586), and CES (ID 1090517) were supported by National Health and Medical Research Council Fellowships. Study and fellowship funders had no role in any part of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 617KB - jmir_v20i12e11321_app1.pdf]

References

1. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
2. Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van Mechelen W, Lancet Physical Activity Series 2 Executive Committee. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet 2016 Sep 24;388(10051):1311-1324. [doi: 10.1016/S0140-6736(16)30383-X] [Medline: 27475266]
3. Vandelanotte C, Duncan MJ, Caperchione C, Hanley C, Mummery WK. Physical activity trends in Queensland (2002 to 2008): are women becoming more active than men? Aust N Z J Public Health 2010 Jun;34(3):248-254. [doi: 10.1111/j.1753-6405.2010.00521.x] [Medline: 20618264]
4. Begg S, Vos T, Barker B. Australian Institute of Health and Welfare. 2007. The burden of disease in Australia 2003 URL: https://www.aihw.gov.au/reports/burden-of-disease/burden-of-disease-injury-australia-2003/contents/table-of-contents [accessed 2018-06-18] [WebCite Cache ID 70G5tG9LY]
5. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC, Lancet Physical Activity Series 2 Executive Committee. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. Lancet 2016 Sep 24;388(10051):1337-1348 [FREE Full text] [doi: 10.1016/S0140-6736(16)30728-0] [Medline: 27475273]
6. Broekhuizen K, Kroeze W, van Poppel MN, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012 Oct;44(2):259-286 [FREE Full text] [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
7. Vandelanotte C, De Bourdeaudhuij I, Brug J. Two-year follow-up of sequential and simultaneous interactive computer-tailored interventions for increasing physical activity and decreasing fat intake. Ann Behav Med 2007 Apr;33(2):213-219. [doi: 10.1080/08836610701310086] [Medline: 17447874]
8. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396. [doi: 10.1093/her/cyl096] [Medline: 16971674]
9. Vandelanotte C, Duncan MJ, Hanley C, Mummery WK. Identifying population subgroups at risk for underestimating weight health risks and overestimating physical activity health benefits. J Health Psychol 2011 Jul;16(5):760-769. [doi: 10.1177/1359105310390543] [Medline: 21421647]
10. De Cocker K, Spittaels H, Cardon G, De Bourdeaudhuij I, Vandelanotte C. Web-based, computer-tailored, pedometer-based physical activity advice: development, dissemination through general practice, acceptability, and preliminary efficacy in a randomized controlled trial. J Med Internet Res 2012 Apr 24;14(2):e53 [FREE Full text] [doi: 10.2196/jmir.1959] [Medline: 22532102]
11. Compernolle S, Vandelanotte C, Cardon G, De Bourdeaudhuij I, De Cocker K. Effectiveness of a web-based, computer-tailored, pedometer-based physical activity intervention for adults: a cluster randomized controlled trial. J Med Internet Res 2015 Feb 09;17(2):e38 [FREE Full text] [doi: 10.2196/jmir.3402] [Medline: 25665498]
12. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
13. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015 Mar 27;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
14. Vandelanotte C, De Bourdeaudhuij I. Acceptability and feasibility of a computer-tailored physical activity intervention using stages of change: project FAITH. Health Educ Res 2003 Jun;18(3):304-317. [Medline: 12828232]
15. Alley S, Jennings C, Plotnikoff RC, Vandelanotte C. Web-based video-coaching to assist an automated computer-tailored physical activity intervention for inactive adults: a randomized controlled trial. J Med Internet Res 2016 Dec 12;18(8):e223 [FREE Full text] [doi: 10.2196/jmir.5664] [Medline: 27520283]
16. Eastin M, LaRose R. Internet self-efficacy and the psychology of the digital divide. J Comp Mediated Commun 2000;6(1):25-56. [doi: 10.1111/j.1083-6101.2000.tb00110.x]
17. Elley C, Kerse N, Arroll B, Robinson E. Effectiveness of counselling patients on physical activity in general practice: cluster randomised controlled trial. Br Med J 2003 Apr 12;326(7393):793 [FREE Full text] [doi: 10.1136/bmj.326.7393.793] [Medline: 12689976]
18. Rose SB, Elley CR, Lawton BA, Dowell AC. A single question reliably identifies physically inactive women in primary care. N Z Med J 2008 Jan 25;121(1268):U2897. [Medline: 18256708]
19. Cardinal BJ, Esters J, Cardinal MK. Evaluation of the revised physical activity readiness questionnaire in older adults. Med Sci Sports Exerc 1996 Apr;28(4):468-472. [Medline: 8778552]

20. Vandelanotte C, Short C, Plotnikoff R, Hooker C, Canoy D, Rebar A, et al. TaylorActive--Examining the effectiveness of web-based personally-tailored videos to increase physical activity: a randomised controlled trial protocol. BMC Public Health 2015 Oct 05;15:1020 [FREE Full text] [doi: 10.1186/s12889-015-2363-4] [Medline: 26438225]
21. Ajzen I. A theory of planned behavior. In: Kuhl J, Beckman J, editors. Action-Control: From Cognition to Behaviour. Heidelberg: Springer; 1985:11.
22. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//110003-066X.55.1.68] [Medline: 11392867]
23. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs: Prentice Hall; 1985.
24. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]
25. Gollwitzer PM. Implementation intentions: strong effects of simple plans. Am Psychol 1999;54(7):493-503. [doi: 10.1037/0003-066X.54.7.493]
26. Australian Institute of Health and Welfare. The Active Australia Survey: a guide and manual for implementation, analysis and reporting. 2003. URL: http://www.aihw.gov.au/publication-detail/?id=6442467449 [accessed 2017-07-12] [WebCite Cache ID 6rtQe6Aq9]
27. Brown WJ, Trost SG, Bauman A, Mummery K, Owen N. Test-retest reliability of four physical activity measures used in population surveys. J Sci Med Sport 2004 Jun;7(2):205-215. [doi: 10.1097/00005768-200305001-01886] [Medline: 15362316]
28. Reeves MM, Marshall AL, Owen N, Winkler EA, Eakin EG. Measuring physical activity change in broad-reach intervention trials. J Phys Act Health 2010 Mar;7(2):194-202. [doi: 10.1016/j.jsams.2009.10.159] [Medline: 20484758]
29. Chau JY, van der Ploeg HP, Dunn S, Kurko J, Bauman AE. A tool for measuring workers' sitting time by domain: the Workforce Sitting Questionnaire. Br J Sports Med 2011 Dec;45(15):1216-1222. [doi: 10.1136/bjsports-2011-090214] [Medline: 21947817]
30. Allison P. Statistical Horizons. Haverford, PA, USA Handling Missing Data by Maximum Likelihood URL: http://www.statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf [accessed 2018-09-03] [WebCite Cache ID 7293YEzuY]
31. Slootmaker SM, Chin A Paw MJ, Schuit AJ, Seidell JC, van Mechelen W. Promoting physical activity using an activity monitor and a tailored web-based advice: design of a randomized controlled trial [ISRCTN93896459]. BMC Public Health 2005 Dec 15;5:134 [FREE Full text] [doi: 10.1186/1471-2458-5-134] [Medline: 16356182]
32. Slootmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, Van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. J Med Internet Res 2009 Jul 29;11(3):e27 [FREE Full text] [doi: 10.2196/jmir.1139] [Medline: 19674956]
33. Short CE, James EL, Girgis A, D'Souza MI, Plotnikoff RC. Main outcomes of the Move More for Life Trial: a randomised controlled trial examining the effects of tailored-print and targeted-print materials for promoting physical activity among post-treatment breast cancer survivors. Psychooncology 2015 Jul;24(7):771-778. [doi: 10.1002/pon.3639] [Medline: 25060288]
34. Gilson ND, Puig-Ribera A, McKenna J, Brown WJ, Burton NW, Cooke CB. Do walking strategies to increase physical activity reduce reported sitting in workplaces: a randomized control trial. Int J Behav Nutr Phys Act 2009 Jul 20;6:43 [FREE Full text] [doi: 10.1186/1479-5868-6-43] [Medline: 19619295]
35. Donnelly JE, Blair SN, Jakicic JM, Manore MM, Rankin JW, Smith BK, American College of Sports Medicine. American College of Sports Medicine Position Stand. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. Med Sci Sports Exerc 2009 Feb;41(2):459-471. [doi: 10.1249/MSS.0b013e3181949333] [Medline: 19127177]
36. Catenacci VA, Wyatt HR. The role of physical activity in producing and maintaining weight loss. Nat Clin Pract Endocrinol Metab 2007 Jul;3(7):518-529 [FREE Full text] [doi: 10.1038/ncpendmet0554] [Medline: 17581621]
37. 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] [Medline: 19296777]
38. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
39. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
40. 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] [Medline: 22786673]
41. Laitakari J, Vuori I, Oja P. Is long-term maintenance of health-related physical activity possible? An analysis of concepts and evidence. Health Educ Res 1996 Dec;11(4):463-477. [doi: 10.1093/her/11.4.463] [Medline: 10163955]
42. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. Soc Sci Med 2017 Dec;192:125-133. [doi: 10.1016/j.socscimed.2017.09.021] [Medline: 28965003]

Abbreviations

BMI: body mass index
CQUniversity: Central Queensland University
PAR-Q: physical activity readiness questionnaire

Edited by G Eysenbach; submitted 18.06.18; peer-reviewed by E Dove, J Boekhout-Oppeneer; comments to author 12.07.18; revised version received 03.09.18; accepted 04.09.18; published 18.12.18.

Please cite as:
Vandelanotte C, Duncan MJ, Maher CA, Schoeppe S, Rebar AL, Power DA, Short CE, Doran CM, Hayman MJ, Alley SJ
The Effectiveness of a Web-Based Computer-Tailored Physical Activity Intervention Using Fitbit Activity Trackers: Randomized Trial
J Med Internet Res 2018;20(12):e11321
URL: http://www.jmir.org/2018/12/e11321/
doi:10.2196/11321
PMID:30563808

Original Paper

Effects of a Participatory Ergonomics Intervention With Wearable Technical Measurements of Physical Workload in the Construction Industry: Cluster Randomized Controlled Trial

Mikkel Brandt[1,2], PhD; Pascal Madeleine[2], DSc, PhD; Afshin Samani[2], PhD; Jeppe ZN Ajslev[1], PhD; Markus D Jakobsen[1], PhD; Emil Sundstrup[1], PhD; Lars L Andersen[1,2], PhD

[1]National Research Centre for the Working Environment, Copenhagen, Denmark

[2]Sport Sciences, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Corresponding Author:
Mikkel Brandt, PhD
National Research Centre for the Working Environment
Lersø Parkalle 105
Copenhagen, DK-2100
Denmark
Phone: 45 39165342
Email: mbp@nrcwe.dk

Abstract

Background: Construction work frequently involves heavy physical work, and a reduction of the physical workload should have high priority. Technological development has made it possible to obtain field measurements with surface electromyography (sEMG), kinematics measured with inertial measurement units (IMUs), and video recordings. However, no studies have used these methods simultaneously to detect situations with excessive physical workload (events) during a working day. Thus, knowledge about these specific events may combat work-related risk factors. Participatory ergonomics (PE) has shown promising results, but whether it can be used as a tool to reduce the physical workload during construction work remains unknown.

Objective: This cluster randomized controlled trial investigated whether a PE intervention with technical measurements consisting of IMUs, sEMG, heart rate monitoring, and video recordings of physical workload could reduce the number of events with excessive physical workload during a working day. Furthermore, other outcomes were obtained from questionnaires.

Methods: A total of 80 male full-time construction workers (aged 19 to 67 years) were randomized at the cluster level (gang) to a PE intervention consisting of 3 workshops (7 gangs and 32 workers) or to a control group (8 gangs and 48 workers). The physical workload was recorded by technical measurements, that is, IMUs, sEMG, heart rate monitoring, and video recordings during a full working day at baseline and 3 and 6 months' follow-up. On the basis of the technical measurements, a custom-made computer program detected the situations (events) where the construction workers were exposed to excessive physical workload and used in the intervention. Differences in the number of events from baseline to follow-up between intervention and control were evaluated using linear mixed models (intention-to-treat), with individual nested in cluster as a random factor. Furthermore, questionnaires were filled out on test days.

Results: The results of the primary outcome showed no change in the number of events with excessive physical workload. However, compared with the control group, the other outcomes showed decreased general fatigue after a typical working day (P=.001) and increased influence on own work (P=.04).

Conclusions: This PE intervention with technical measurements did not reduce the number of events with excessive physical workload during construction work. However, the intervention led to decreased general fatigue and increased influence on own work.

Trial Registration: ClinicalTrials.gov NCT02498197; https://clinicaltrials.gov/ct2/show/NCT02498197 (Archived by WebCite at http://www.webcitation.org/74SZ3DIWS)

(J Med Internet Res 2018;20(12):e10272) doi:10.2196/10272

KEYWORDS

back pain; low back pain; shoulder pain; musculoskeletal pain; musculoskeletal diseases; occupational health; building industry; heavy industries; organizational ergonomics; action research

Introduction

Background

Work-related musculoskeletal disorders (WMSDs) such as low back pain and shoulder pain constitute a substantial problem for individuals, workplaces, and societies [1-3]. At the individual level, WMSDs increase risk of poor health, sick leave, and premature exit from the labor market [1,4,5]. For workplaces, workers with WMSDs have lower workability and are more likely to have long-term sickness absence [6,7]. For the societies, WMSDs lead to substantial expenses regarding treatment, lost production, and sickness absence [1,8]. Heavy physical work is a known risk factor for developing WMSDs [9,10]. In particular, heavy lifting, pushing or pulling, and working in awkward postures have been associated with low back pain [11] and sickness absence [4,12,13]. Construction work consists of a high degree of heavy physical work [14,15]. Consequently, a reduction of the physical workload to promote sustainable working careers [16] in construction work should have high priority. Moreover, a systematic review highlighted an urgent need for interventions focusing on reducing WMSDs in construction workers [17]. In addition, most field studies in the construction industry are based on self-reported measurements [17]. Hence, a more technical approach may enable objective evaluation of the loading and provide better grounds for targeted and effective interventions.

Technical Measurements

Technological development has made it possible to obtain field measurements with surface electromyography (sEMG) [18,19], kinematics measured with inertial measurement units (IMU) [20-23], or a combination [24]. However, no studies have used sEMG, IMU, and video recordings obtained simultaneously to detect events with excessive physical workload (events) during a working day. Thus, knowledge about these specific events may be an important tool for engaging workers to combat work-related risk factors.

Participatory Ergonomics

In participatory ergonomics (PE), the workers are involved in the decision processes. Systematic reviews have reported that PE has positive effects on musculoskeletal symptoms [25] and thereby may lead to increased productivity and reduced occupational risk factors [26]. Furthermore, a systematic review has shown that participatory responsibility concerning the identification of risk factors, development of solutions, and implementation is important to succeed in the participatory process [27]. Nevertheless, the evidence for preventing neck-shoulder and low back pain through ergonomics interventions is questionable because the number of randomized controlled trials are limited [28].

Objectives

This cluster randomized controlled trial investigated whether a PE intervention with technical measurements could reduce the number of events with excessive physical workload during a working day in the construction industry We hypothesized that the PE intervention involving both managers and workers would lead to a reduction in the number of events of excessive physical workload.

Methods

Study Design

This study was a 2-armed, parallel group, single-blinded, cluster randomized controlled trial with allocation concealment performed at construction sites across Denmark from May 2016 to June 2017. Clusters were defined as construction gangs. The organization of construction work, that is, working in construction gangs, was the reason for choosing a cluster design. The intervention consisted of 3 workshops based on individual technical measurements of excessive physical workload. The technical measurements to detect excessive physical workload have previously been validated in controlled laboratory settings [29] and were conducted at baseline and 3 and 6 months' follow-up.

Ethics

According to the Helsinki declaration, participants received written and oral information about the purpose and content of the study before signing the informed consent form. The study was approved by the local ethical committee of Frederiksberg and Copenhagen (H-3-2010-062) and registered with the Danish Data Protection Agency (215-57-0074) and ClinicalTrials.gov (NCT02498197). The reporting followed the CONSORT statements for cluster randomized trials [30] and CONSORT eHealth [31] (Multimedia Appendix 1). Design of the effect evaluation and process evaluation have previously been reported [32,33]. This study reports data solely from the effect evaluation.

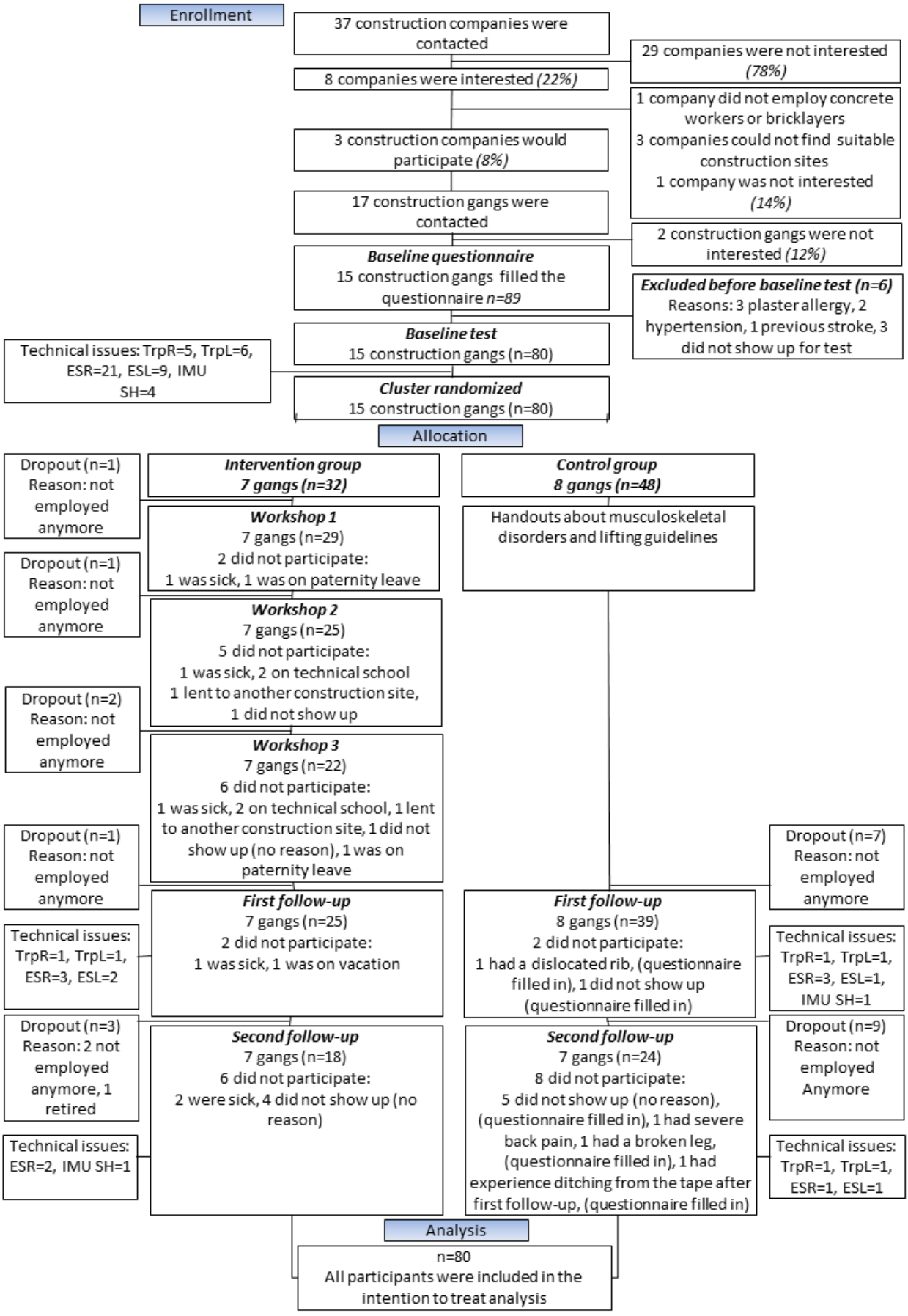
Participants

The inclusion criterion was full-time construction work. The exclusion criteria were life-threatening diseases and hypertension >160/100 mmHg. A total of 9 participants were excluded before the baseline test. Moreover, 80 participants (15 clusters (gangs)) met the inclusion criteria and completed the baseline test. The flow of participant enrollment is illustrated in Figure 1.

Randomization and Blinding

The randomization was performed by a researcher who was not involved in data collection (LLA). Block randomization of the construction gangs was chosen for practical reasons and was performed continuously as the baseline tests were completed. The researchers performing the data collection were not aware of the block size or group allocation. Blinding of participants is not possible in behavioral interventions. The data analyst and the statisticians were blinded to group allocation.

Figure 1. Participant's recruitment flowchart. TrpR: trapezius right; TrpL: trapezius left; ESR: erector spinae right; ESL: erector spinae left; IMU SH: inertial measurement unit, shank.



Intervention

The intervention was carried out at gang level and consisted of workshops or reading handouts for the intervention group and control group, respectively.

Intervention Group

The workshops were organized in a 3-phase structure inspired by an action research approach [34]. The programs aimed to create possibilities for change by enabling engagement between the technical measurements and the participants. The first workshop was designed with main inspiration from the *future workshop*; a type of workshop that usually consists of 3 phases: (1) critique, (2) utopia, and (3) realization [35]. In our design, the critique phase was replaced by an introduction of the video recordings of the participants' own work and a description of the physical workload measured in relation to each video recording. Subsequently, the participants decided which work situations should be modified during the intervention. In the *utopian phase*, the participants discussed in groups each selected work situation. The participants were to discuss and describe how the selected work process could be carried out in the best of all worlds with minimal physical exertion. In this phase, the participants were instructed not to take any barriers into account to facilitate creative resourcefulness. In the *realization phase*, the participants were asked to consider possibilities and barriers to reach the utopias. Furthermore, a plan of action was written.

In the second workshop, the participants were asked to recapture the focal points of the first workshop and to describe the progress concerning each of the selected topics in the first workshop. Then, they were encouraged to describe the barriers they had encountered in the process of reaching the goals of changing the working situations. Following this, the researchers described the current knowledge on organizational and social practices about WMSDs in the construction industry. The purpose of this was to nudge the participants to increased creativity and to challenge potentially frozen conceptions of how work should be done. Finally, the participants were encouraged to come up with further ideas on how to work toward the utopias or to aim for new utopias if they had reached their initial goals.

The third workshop had the purpose of anchoring initiatives. The researchers initially asked the participants to describe the status of the goals set earlier in the project. Then, the participants were invited to discuss whether the organization would be able to implement the initiatives of the project into long-term working practice and to come up with ideas for initiatives that could help secure this long-term anchorage.

Control Group

The control group received handouts about WMSDs [36] and lifting guidelines from The Danish Working Environment Authority [37]. These handouts described the association between WMSDs and the impact on working life, regulations for the prevention of WMSDs, and which precautions should be taken to limit WMDSs [36]. Furthermore, the handouts described the regulations for lifting, pushing and pulling, and the risk of injuries [37].

Technical Measurements

At baseline and 3 and 6 months' follow-up, the participants were equipped with sEMG, IMU, cameras, and heart rate monitors. The sEMG, IMU, and camera were synchronized [29], whereas the heart rate monitor was used to estimate the overall activity level during the working day.

The procedure for placement of sEMG electrodes is described elsewhere [29,32,38]. In short, sEMG electrodes (Blue Sensor N-00-S/25, Ambu A/S, Ballerup, Denmark) were placed bilaterally over the erector spinae and the upper trapezius muscles [39] according to the Surface Electromyography for the Non-Invasive Assessment of Muscles (SENIAM) recommendations [40]. A reference electrode was placed over the C7 vertebra. The sEMG signals were amplified 19.5 times using a 24-bit portable data-logger (Nexus10, Mind Media, Herten, Netherlands) and sampled at 1024 Hz.

IMU including triaxial accelerometer and gyroscopes (ActiGraph GT9X Link, ActiGraph, Pensacola, United States) were positioned on the upper back at T1-T2 level [41] and the thigh. The latter was used for obtaining the number of steps per day [42]. When positioned, the IMUs were calibrated in a standing neutral position (N-pose) for 15 seconds. Kinematics data were sampled at 100 Hz.

A body-worn video camera with a resolution of 848x480/30F (Reveal Media, RS2-X2L, Hampton Wick, Surrey, United Kingdom) was placed around the chest and recorded the area in front of the participant.

For heart rate monitoring electrodes (Ø: 68 mm; Blue Sensor VL-00-S/25, Ambu A/S, Ballerup, Denmark) were positioned just below the apex of the sternum and laterally under the left pectoralis major muscle [43,44], before connecting to the heart rate monitor (Actiheart, CamNtech Ltd, Cambridge, United Kingdom). Heart rate was sampled at 128 Hz and interpolated with a resolution of +/-1 ms.

The data from the IMU, sEMG, and cameras were synchronized using a custom-made device and a MatLab (2013a) program. A 2 mV trigger signal was sent to the EMG logger. At the same time, the IMUs were turned 95 degrees using a rotary solenoid (GDAX 050 X20 B71 24V, 100% ED). The synchronization with the cameras was obtained by having the cameras record a custom-made flashing device that flashed at the same time as the signal was sent to the sEMG logger box and the rotary solenoid. The synchronization was made before the equipment was positioned on the participant and repeated after the working day [29].

Test Protocol

The test protocol consisted of (1) maximal voluntary contractions (MVCs) for the lower back and shoulders, (2) reference lifting, and (3) calibration of the IMUs.

The MVCs for the upper trapezius was performed with a strap around the wrist in a standing position with 90-degree bilateral shoulder abduction. The participants performed maximal bilateral shoulder abduction. For the MVCs for the erector spinae, the participants were fixed with a strap around the shoulders with a slight flexion in their back leaning toward a

pillow at the height of the anterior iliac spine on the hip. The participants performed 3 repetitions for each MVC with a 30 seconds rest period between the trials. The participants ramped up the force to maximum in 2 to 3 seconds and held it for 3 seconds. The participants performed 10 reference lifts from floor to table (73 cm high) with a 20 kg box (width: 56 cm, length: 34 cm, and height: 20 cm) using a forearm horizontal distance. From a starting standing position, the participants descended without load and lifted the box from the floor onto the table. After a pause of 2 seconds, they lifted the box to the floor and returned to starting position. Following a break of 2 seconds, the lifting cycle was repeated. The IMUs were calibrated by having the participants standing in a neutral position (N-pose) for 15 seconds. After these preliminary steps, the participants started their planned work and the attached equipment continuously captured data.

Data Analysis-Event Detection

The analyses for detecting the events are described in detail elsewhere [29]. In short, the sEMG segments corresponding to the references lifts and the MVC trials were extracted. For each of the MVC trials, the sEMG root mean square (RMS) was calculated over 500 ms epochs with 20% overlap between successive epochs. Then the maximum of calculated RMS across the epochs was found, and out of the 3 MVC trials, the highest RMS value was considered as the maximum voluntary electrical activity. Similarly, the 90th percentile of calculated RMS during the reference lift was considered as the reference threshold. Subsequently, the recorded signals during the working time were analyzed over 10-second nonoverlapping epochs. Similar to what is described above, in each epoch, the 90th percentile of the calculated EMG RMS was derived, and the extent of forward and sideways inclination of the IMU concerning the N-pose position was calculated [45]. During the working time, each of the 10-second epochs were labeled an event if the calculated sEMG amplitude was higher than the event threshold (either the average of the reference lifts in the morning and afternoon or 50% of the average MVC [46] in the morning and afternoon) for at least two of the muscles. Furthermore, for the erector spinae muscle on both sides, the event threshold was linearly decayed based on the calculated forward and sideways inclination such that the threshold would be reduced to its half at 90 degrees forward or 30 degrees sideways inclinations and it would be fixed beyond that level of inclination. The minimum of the modified threshold for the forward and sideways inclination was utilized as the modified threshold. If any of the calculated sEMG RMS over the 10-second epochs for the erector spinae on both sides was greater than the modified threshold, the 10-second epoch was labeled as an event as well. Furthermore, as exploratory analyses, the number of events was calculated based on a higher reference value of 150% of the sEMG from the reference lifts and 50% of the sEMG from MVCs. The criteria for events from the analyses were that at least two muscles should exceed the limit.

Outcomes

Primary Outcome

The primary outcome was defined as the change in the number of events with excessive physical workload from baseline to follow-up. The reference lift of 20 kg used for normalization purposes was a deviation from the protocol study [32] as we planned to use 30 kg. However, because 30 kg exceeds the acceptable lifting limit of The Danish Working Environment Authority, we chose to decrease the load.

Other Outcomes

Other outcomes were obtained from previously established and validated questionnaires and included physical (Borg category ratio 10 scale [Borg CR10]) [47,48], psychosocial, and organizational conditions (AH2012 and COPSOQ) [49-51]. Furthermore, the pain intensity in the last week (WAS-scale) [32] was enquired.

Sample Size

The sample size was calculated based on the observed changes in the level of muscular activity during a working day in different occupational groups with pronounced lifting [19]. The power calculation showed that 17 participants in each group were needed to demonstrate a reduction of 20% in normalized sEMG assuming an SD of approximately 20% in normalized sEMG between individuals and a type 1 risk of 5% and power of 80%. Due to the cluster design and including an inflation factor of 1.5, 26 participants were required in each group [32]. For generalizability and risk of dropouts, we aimed to recruit 10 construction gangs of 3 to 5 individuals in each group, that is, a total of 80 participants.

Statistics

t tests assessed possible group differences at baseline. The difference from baseline to follow-up between the intervention and control groups was evaluated using a linear mixed model. The number of events was log-transformed because the residuals were not normally distributed. Factors included in the model were group (intervention and control), time (baseline, first follow-up, and second follow-up), and *group-by-time* interaction. The analysis was adjusted for the baseline value of the outcome, age, gender, duration of measuring time, mean heart rate, number of steps, and muscle strength. Individual nested in cluster was included as a random factor. Analyses were performed using SAS statistical software (Proc Mixed, SAS version 9.4) according to the *intention-to-treat* principle, including all participants (n=80) regardless of loss to follow-up. The estimation method was restricted maximum likelihood with degrees of freedom based on the Kenward-Roger approximation. *P* levels ≤.05 were accepted as statistically significant. Outcomes are reported as within- and between-group least square mean differences with 95% CIs. Furthermore, Fischer exact test was used to test for differences in questions with categorical response variables.

Results

Participant Characteristics

Table 1 shows the baseline characteristics of the participants. Age was higher in the control group compared with the intervention group (*P*=.02), which was controlled for in the statistics by including age as a covariate. At the first follow-up test, 12 participants dropped out, and 4 participants did not show

up for the test. At the second follow-up test, 12 participants dropped out, and 14 participants did not show up for the test. Hence, 42 participants completed the study (Figure 1). All dropouts were included in analyses.

Primary Outcome

The results showed no *group-by-time* interaction effect (P=.75) and (P=.51) for the number of events obtained using technical equipment in the unadjusted and adjusted analysis, respectively (Table 2). The results show a within-group difference (*time* effect) in the number of events from baseline to the first follow-up test (unadjusted, P=.002 and adjusted P=.05 and (unadjusted P<.001 and adjusted P<.001) for the intervention and control group, respectively. Furthermore, a within-group difference was observed from baseline to second follow-up in the unadjusted analysis for the intervention and control group (P<.01 and P<.001), respectively. The exploratory analyses confirmed the results of the primary outcome, that is, no significant *group-by-time* interaction (Table 2).

The analyses of heart rate and step count showed no *group-by-time* interaction or within-group difference. However a between-group difference was observed at baseline, first and second follow-up, and at baseline for heart rate (P<.001, P=.049, and P=.003, respectively). This between-group difference was also observed for the step count at baseline (P=.004) but not at the follow-ups. The mean heart rate was 100 (95% CI 96 to 104), 101 (95% CI 97 to 105), and 100 (95% CI 96 to 105) and 91 (95% CI 88 to 95), 95 (95% CI 92 to 99), and 91 (95% CI 87 to 95) bpm for the intervention and control group at baseline, first follow-up, and second follow-up, respectively. The mean number of steps adjusted for length of the working day were 5952 (95% CI 5517 to 6387), 5479 (95% CI 5023 to 5934), and 5980 (95% CI 5372 to 6588) and 5133 (95% CI 4791 to 5475), 5340 (95% CI 4958 to 5722), and 5320 (95% CI 4852 to 5788) steps per day for the intervention and control group at baseline, first follow-up, and second follow-up, respectively.

Other Outcomes

The results from the other outcomes are presented in Tables 3 and 4. In the intervention group, the results showed a within-group decrease in general fatigue after a typical working day (P=.001; Table 3) from baseline to second follow-up and in influence on own work from baseline to first follow-up (P=.04; Table 3). The remainder of the other outcomes showed no effect from the intervention (Tables 3 and 4).

Adverse Events

No adverse events were reported.

Table 1. Basic characteristics of the participants in the study.

Characteristics	Intervention group	Control group
Number of participants, n (all males)	32	48
Age in years, mean (SD)	34.2[a] (12.5)	41.2[a] (12.5)
Height in centimeters, mean (SD)	180.0 (6.2)	180.1 (7.2)
Weight in kilograms, mean (SD)	85.0 (12.2)	86.4 (14.6)
Weekly working hours, mean (SD)	39.1 (4.8)	38.1 (2.5)
Gang size, mean (SD)	4.5 (1.5)	6.0 (2.3)
Smokers, n (%)		
Yes	15 (47)	17 (36)
No	17 (53)	31 (64)
Current position, n (%)		
Concrete workers	25 (78)	27 (56)
Bricklayers	5 (16)	19 (40)
Others (eg, bricklayer´s assistant)	2 (6)	2 (4)
Term of employment, n (%)		
Hourly paid	13 (41)	35 (73)
Monthly paid	1 (3)	0 (0)
Paid according to performance	18 (56)	13 (27)
Experience in construction, n (%)		
<3 years	4 (12)	4 (8)
4-10 years	14 (44)	13 (27)
>11 years	14 (44)	31 (65)
How often can you take it easy and still reach your working tasks?, n (%)		
Always	0 (0)	1 (2)
Often	6 (19)	11 (23)
Sometimes	16 (50)	23 (48)
Rarely	9 (28)	12 (25)
Never	1 (3)	1 (2)
How exhausting do you find your regular work? (Borg CR10[b]), n (%)		
Light (0-2.5)	1 (3)	6 (12.5)
Moderate (3-5)	16 (50)	36 (75)
Hard (6-10)	15 (47)	6 (12.5)
How often do you feel pain in your body? n (%)		
Every day	10 (31)	16 (33)
A few times a week	10 (31)	8 (17)
A few times a month	8 (25)	18 (38)
Maximum a few times a year	4 (13)	6 (12)
Never	0 (0)	0 (0)
Degree of difficulty in the low back within the last week (0-10 VAS[c]), n (%)		
0-3	13 (42)	24 (52)
4-6	10 (32)	15 (33)
7-10	8 (26)	7 (15)

Characteristics	Intervention group	Control group
Degree of difficulty in the upper back within the last week (0-10 VAS[c]), n (%)		
0-3	19 (61)	35 (76)
4-6	8 (26)	9 (20)
7-10	4 (13)	2 (4)
Degree of difficulty in the shoulders within the last week (0-10 VAS[c]), n (%)		
0-3	23 (74)	32 (70)
4-6	6 (19)	11 (24)
7-10	2 (6)	3 (6)

[a]Difference between groups at baseline, *P*=.02.

[b]Borg CR10: Borg category ratio 10 scale.

[c]VAS: visual analog scale.

Table 2. Results of the primary outcome (change from baseline to follow-up in events with excessive physical workload during a working day) from the mixed-model analysis.

Group and outcome		Within-group difference				Between-group difference at follow-up			
Group	Baseline	First follow-up	*P* value	Second follow-up	*P* value	First follow-up	*P* value	Second follow-up	*P* value
Primary outcome									
100% sEMG[a] from reference lifts-unadjusted (95% CI)									
Intervention	5.2 (5 to 5.5)	5.8 (5.5 to 6.1)	.002	5.8 (5.4 to 6.1)	.01	−0.1 (−0.4 to 0.3)	.72	0 (−0.4 to 0.4)	.94
Control	5.1 (4.9 to 5.3)	5.9 (5.6 to 6.1)	<.001	5.8 (5.5 to 6)	<.001	—[b]	—	—	—
100% sEMG from reference lifts-adjusted (95% CI)									
Intervention	5.4 (5.1 to 5.7)	5.8 (5.5 to 6.1)	.05	5.7 (5.3 to 6.0)	.3	0 (−0.4 to 0.4)	.89	0.1 (−0.4 to 0.6)	.62
Control	5.1 (4.9 to 5.4)	5.8 (5.6 to 6.1)	<.001	5.5 (5.2 to 5.9)	.052	—	—	—	—
Explorative analysis									
150% sEMG from reference lifts (95% CI)									
Intervention	3.8 (3.5 to 4.2)	4.3 (3.9 to 4.8)	.06	3.8 (3.3 to 4.3)	.94	0.3 (−0.2 to 0.8)	.21	0.4 (−0.2 to 0.9)	.89
Control	3.6 (3.2 to 3.9)	4.0 (3.6 to 4.4)	.08	3.9 (3.4 to 4.3)	.3	—	—	—	—
50% sEMG from MVCs[c] (95% CI)									
Intervention	4.2 (3.7 to 4.8)	4.2 (3.6 to 4.9)	.97	4.8 (3.9 to 5.6)	.35	0.2 (−0.6 to 1.1)	.6	0.6 (−0.5 to 1.7)	.3
Control	4.0 (3.5 to 4.4)	4.0 (3.4 to 4.6)	.89	4.2 (3.4 to 4.9)	.62	—	—	—	—

[a]sEMG: surface electromyography.

[b]Not applicable.

[c]MVCs: maximal voluntary contraction.

Table 3. Results from the other outcome.

Group and scale	Baseline	First follow-up	Second follow-up	Between-group difference at follow-up		Time effect, *P* value
				First follow-up	Second follow-up	
Heaviest lift last week, 0-10 scale (95% CI)						
Intervention	6.7 (6.1 to 7.2)	6.8 (6.2 to 7.4)	6 (5.3 to 6.7)	0.2 (0.6 to 0.1)	−0.1 (−1 to 0.8)	.52
Control	6.2 (5.7 to 6.6)	6.5 (6 to 7)	6.1 (5.5 to 6.6)	—[a]	—	—
General fatigue after a typical working day, 5-point scale (not tired, a little tired, tired, very tired, and exhausted) converted to 0-100 (95% CI)						
Intervention	41.1 (37.2 to 45)	40.8 (36.6 to 45)	35.6 (31 to 40.3)	−6.1 (−11.7 to −0.5)	−11.2 (−17.4 to −5)	.001
Control	39.4 (36.2 to 42.5)	46.9 (43.4 to 50.4)	46.8 (42.9 to 50.7)	—	—	—
How physically strenuous do you usually perceive your current work?, Borg CR10 scale[b] (95% CI)						
Intervention	4.8 (4.2 to 5.3)	4.4 (3.9 to 5)	5 (4.4 to 5.7)	−0.4 (−1.1 to 0.4)	0.5 (−0.3 to 1.4)	.10
Control	4.3 (3.9 to 4.8)	4.8 (4.3 to 5.3)	4.5 (3.9 to 5)	—	—	—
How much influence do you have on your work, 5-point scale (very much, much, some, little, very little) converted to 0-100 (95% CI)						
Intervention	59.6 (56.6 to 62.5)	60.4 (57.3 to 63.6)	58.9 (55.4 to 62.5)	5.6 (1.5 to 9.8)	−0.1 (−4.8 to 4.6)	.04
Control	60.3 (58 to 62.7)	54.8 (52.2 to 57.4)	59 (56 to 62)	—	—	—
Do you wish more influence on your work, 2-point scale (yes or no) converted to 0-100 (95% CI)						
Intervention	39.3 (29.6 to 49)	49.4 (39 to 59.7)	48.7 (37.3 to 60.2)	−3.9 (−17.6 to 9.7)	0.1 (−15.3 to 15.4)	.85
Control	44.1 (36.3 to 51.9)	53.3 (44.7 to 61.9)	48.7 (38.9 to 58.5)	—	—	—

[a]Not applicable.

[b]Borg CR10: Borg category ratio 10 scale.

Table 4. Results from the other outcome. Numbers indicate the participants who answered the question (percent of the population who answered the question).

Group and question	Baseline				First follow-up				Second follow-up			
	Daily	Weekly	Hardly ever	*P* value[a]	Daily	Weekly	Hardly ever	*P* value[a]	Daily	Weekly	Hardly ever	*P* value[a]
How often do you perform heavy lifting?, n (%)												
Intervention	21 (66)	11 (34)	0 (0)	.2	14 (50)	13 (46)	1 (4)	.54	11 (50)	11 (50)	0 (0)	.22
Control	23 (48)	24 (50)	1 (2)	—[b]	19 (48)	16 (40)	5 (12)	—	10 (32)	18 (58)	3 (10)	—
How often do you feel pain in your body (eg, arms, hands, knees, shoulders, and back)?, n (%)												
Intervention	10 (31)	10 (31)	12 (38)	.32	8 (29)	7 (25)	13 (46)	.53	3 (14)	10 (45)	9 (41)	.59
Control	16 (33)	8 (17)	24 (50)	—	14 (34)	13 (33)	13 (33)	—	8 (26)	12 (39)	11 (35)	—
Do you take analgesics because of pain in your neck/shoulders or back?, n (%)												
Intervention	1 (3)	1 (3)	30 (94)	.65	1 (4)	2 (7)	25 (89)	.86	0 (0)	1 (5)	21 (95)	.37
Control	3 (6)	4 (8)	41 (86)	—	2 (5)	5 (13)	33 (82)	—	2 (6)	4 (13)	25 (81)	—

[a]Between group differences.

[b]Not applicable.

Discussion

Principal Findings

This study is the first to detect events with excessive physical workload using only technical measurements, individual thresholds, and applying these measurements in a PE intervention. The results of this cluster randomized controlled trial showed that a PE intervention did not decrease the number of events during a working day. Other outcomes showed positive effects on influence on own work and general fatigue after a typical working day, but not on pain, perceived workload, and how often heavy lifting was performed.

Interpretation of Results

Technical measurements have the advantage of being objective. Furthermore, it builds on from standardized analytical procedures of the raw data, rather than, for example, self-reports or visual observations [52,53]. We have recently shown that the intraday reliability for sEMG during lifting tasks is acceptable in laboratory conditions [38]. Furthermore, the method for detecting events of the lower back and shoulder based on sEMG and IMU from the upper back has shown high accuracy in a laboratory setting [29]. Thus, we are certain the measurement method per se was not cause of the nonsignificant findings.

There was no *group-by-time* interaction for the primary outcome, which was the change in the number of events from baseline to follow-up. However, the number of events increased over time in both groups, and because this technical detection was used for the first time in a field study, we performed exploratory analyses with the detection of events based on 150% of the sEMG obtained during the reference lifts and 50% MVCs (Table 2). This confirmed that there was no effect of the intervention on the primary outcome. However, the exploratory analysis did not show a within-group difference as seen in the preplanned analysis [32]. This could be related to the threshold, which might have been too low, and we might have seen a more stable normalization factor by using, for example, 30 kg as reference value. However, the within-group increase in both groups could also be related to the organization of construction work, which is characterized by a distinctive variation regarding work pace, work tasks, and work processes. This variation makes it challenging to conduct intervention studies with field measurements in the construction industry as the inherent variance will necessitate a larger sample size than anticipated based on laboratory measurements. As the participants were further into the process of their current construction project during the follow-up than at baseline, this may have increased the work pace due to incentive reasons. However, analyses of heart rate and step count did not support this speculation. Another possibility could be that the participants were more aware of the measurements at baseline and therefore acting more carefully to avoid heavy lifting. This effect may have diminished during follow-up.

The other outcomes showed effects on influence at work and general fatigue after a typical working day. The difference in influence on own work was only seen at first follow-up and was primarily driven by the control group experiencing a decrease in influence. This may indicate that the control group felt neglected compared with the intervention group who attended the workshops and had the opportunity to bring forward new ideas. The second follow-up was 3 months after the last workshop. Therefore, the feeling of being neglected might have eased off, likely because the tangible consequences of the intervention were often only “increased attention to physically stressful work” as shown in Multimedia Appendix 2, rather than real changes in the working process or technical assistive devices. The decrease in general fatigue after a typical working day in the intervention group indicates that some effect occurred in response to the intervention despite not being effective in reducing events. It can be speculated that the implemented solutions led to work that reduced light loads, repetitive work, or made the work processes more efficient in general and thus less physically fatiguing. A review has shown reductions in physical work demands and musculoskeletal symptoms if mechanical lifting devices are introduced at workplaces [54], and other studies have shown a decreased discomfort [55] or ergonomic improvement when introducing devices for raised bricklaying that may decrease the physical workload during construction work [56]. As the majority of the implemented suggestions concerned technical assistive devices, it can be speculated that the increased use of assistive devices may partially explain the decrease in general fatigue.

The majority of the suggestions were related to assistive devices (Multimedia Appendix 2) and are in accordance with previous findings, where the workers identified ergonomic solutions using assistive devices to reduce WMSDs, but the support from the contractors to implement these was lacking [57]. Accordingly, other studies suggest that support from the management is critical for providing changes in the construction industry [58,59]. In this study, the management was often not willing to support the suggestions if they involved increased costs. Hence, more support from the management might have had a positive effect [60]. The intervention might have failed in involving the management as we underestimated the challenge of obtaining economical and persistent commitment from the management. However, this seems to be a highly common but underaddressed issue in participatory research [27,61].

There may be several contributing factors to the high physical demands of construction work, of which work organization plays an important role. Construction work in Denmark is characterized by being organized in small working units, often on a group-based wage, which can be associated with an increased risk factor for WMSDs [62] and can induce a group pressure within the gang to get the work done at a certain time without taking pain into consideration [63]. Studies suggest that both structural and cultural changes are necessary to create changes in the construction industry [57,64]. The lack of effect from the intervention in this study might be related to the culture in the construction industry where WMSDs are an accepted part of being a construction worker [65,66].

Perspectives

With the rapid technological development, this method could be integrated into portable devices connected to, for example, mobile phones and thereby provide the worker with direct feedback to prevent work tasks with excessive physical workload.

Strengths and Limitations

A strength of this study is the cluster randomized design, making it possible to intervene at gang level, thus reducing the risk of a number of biases associated with nonrandomized studies. However, there are also known challenges of conducting behavioral randomized controlled trials, for example, blinding of participants or potential participants and supervisors who do not accept randomization [67]. Another strength is the use of technical measurements to quantify the workload rather than relying on self-reports or observations.

A limitation of this study is that the number of dropouts was higher than expected, resulting in reduced statistical power. The terms of employment in the construction industry are dominated by short-term contracts, which resulted in a relatively high turnover of workers in the participating gangs in this study. This affected the number of participants employed over the entire intervention period. In research involving randomized controlled trial, it is preferable to have a stable group of participants. However, to our knowledge, no participants drop out of the study due to a lack of willingness to participate but were missing at random, and all participant were included in the intention-to-treat analysis. To fully control a randomized controlled trial in the construction industry showed to be extremely difficult due to, for example, variation in work tasks performed during a working day and sudden changes based on unpredictable incidents at the construction site. On the other hand, a considerable strength of this study is that the measurements have been conducted during actual construction work. Furthermore, it is a limitation that the size of the individual clusters was larger than anticipated, which also reduced the statistical power due to intracluster correlation. In general, the variance in measurements was also higher than expected. Due to these factors, future studies would need to recruit a much larger sample size to be randomized. On the other hand, the results of this study do not indicate that a relevant between-group difference would be reached even with a larger sample size.

The loss of data from sEMG, especially from the erector spinae muscles at baseline, was also a limitation. This loss of sEMG data was primarily caused by electrodes that slipped off, and future studies should minimize this loss of data by securing the sEMG cables such that excessive sweating of the participants does not compromise the skin-electrode impedance. Larger band aids over the electrodes or performing the measurements during the cooler season of the year when sweating is not a big issue might also help.

The inherent variation in daily working tasks at the construction sites is a practical challenge because the necessary sample size can easily grow to a level that is not realistic to achieve. We tried to control this by having close contact with the construction sites and conducting measures on the workers during similar working tasks, but this was not always possible. However, we compensated for this by controlling for steps and heart rate in the analysis.

During the recruitment, we were in contact with many small-scale construction companies that were unable to participate because their job tasks did not permit the long follow-up time in this study. Hence, we only included large-scale construction companies; thus, one should be cautious about generalization of our results to small-scale construction companies. On the other hand, changes are often even more difficult to implement in smaller companies where resources are scarce. It is, therefore, unlikely that inclusion of smaller companies would have changed the main conclusion of the study.

Finally, the difference in age and, partly, experience between the intervention and control group could be limitations to the study. Therefore, we controlled for age in the statistical analysis, but it cannot be ruled out that a more experienced intervention group could have increased the implementation rate of the suggested solution and thus reduced the number of events with excessive physical workload following the intervention.

Conclusions

This PE intervention with 3 workshops did not reduce the number of events with excessive physical workload during construction work. An exploratory analysis using higher thresholds confirmed the results. The intervention group experienced a reduced general fatigue and an increased influence on own work following the intervention, compared with the control group.

Acknowledgments

This study was supported by grants from The Danish Council for Independent Research (DFF-4092-00320). The loss of earnings for the participants was covered by The Danish Construction Association and the Danish union 3 F's “Arbejdsmiljø & Samarbejdsfonden.” The funders had no role in the study design, data collection, data analysis, the decision to publish, or preparation of the manuscript. The authors would like to thank Klaus Hansen and Sebastian Skals from The National Research Centre for the Working Environment for their great help and support during data collection. The authors would like to thank all the construction firms and the participants at the construction sites.

Authors' Contributions

MB, LLA, and PM conceived the study design. MB, LLA, and PM obtained the funding. AS, PM, MDJ, and MB were responsible for developing the technical measurement and analyzing the signals. AS developed the MatLab script. MB and LLA performed the statistical analyses. MB, LLA, JZNA, and ES designed the questionnaires. MB, LLA, and JZNA developed the program for the workshops. MB, ES, and MDJ did the data collection. MB wrote the draft of the manuscript before all authors read, critically reviewed, and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 342KB - jmir_v20i12e10272_app1.pdf]

Multimedia Appendix 2

Descriptions of the workshops, including issues and solution from the participants. OHS: occupational health consultants, OHC: occupational health chief, int: internal, ext: external.

[PDF File (Adobe PDF File), 31KB - jmir_v20i12e10272_app3.pdf]

References

1. Punnett L, Wegman DH. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. J Electromyogr Kinesiol 2004 Feb;14(1):13-23. [doi: 10.1016/j.jelekin.2003.09.015] [Medline: 14759746]
2. Brooks PM. The burden of musculoskeletal disease--a global perspective. Clin Rheumatol 2006 Nov;25(6):778-781. [doi: 10.1007/s10067-006-0240-3] [Medline: 16609823]
3. Dunning KK, Davis KG, Cook C, Kotowski SE, Hamrick C, Jewell G, et al. Costs by industry and diagnosis among musculoskeletal claims in a state workers compensation system: 1999-2004. Am J Ind Med 2010 Mar;53(3):276-284. [doi: 10.1002/ajim.20774] [Medline: 19937981]
4. Andersen LL, Fallentin N, Thorsen SV, Holtermann A. Physical workload and risk of long-term sickness absence in the general working population and among blue-collar workers: prospective cohort study with register follow-up. Occup Environ Med 2016 Apr;73(4):246-253. [doi: 10.1136/oemed-2015-103314] [Medline: 26740688]
5. Saastamoinen P, Laaksonen M, Kääriä SM, Lallukka T, Leino-Arjas P, Rahkonen O, et al. Pain and disability retirement: a prospective cohort study. Pain 2012 Mar;153(3):526-531. [doi: 10.1016/j.pain.2011.11.005] [Medline: 22340946]
6. Reeuwijk KG, Robroek SJ, Hakkaart L, Burdorf A. How work impairments and reduced work ability are associated with health care use in workers with musculoskeletal disorders, cardiovascular disorders or mental disorders. J Occup Rehabil 2014 Dec;24(4):631-639 [FREE Full text] [doi: 10.1007/s10926-013-9492-3] [Medline: 24390780]
7. Alavinia SM, de Boer AG, van Duivenbooden JC, Frings-Dresen MH, Burdorf A. Determinants of work ability and its predictive value for disability. Occup Med (Lond) 2009 Jan;59(1):32-37. [doi: 10.1093/occmed/kqn148] [Medline: 19073989]
8. Badley EM, Rasooly I, Webster GK. Relative importance of musculoskeletal disorders as a cause of chronic health problems, disability, and health care utilization: findings from the 1990 Ontario Health Survey. J Rheumatol 1994 Mar;21(3):505-514. [Medline: 8006895]
9. da Costa BR, Vieira ER. Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies. Am J Ind Med 2010 Mar;53(3):285-323. [doi: 10.1002/ajim.20750] [Medline: 19753591]
10. Andersen LL, Fallentin N, Ajslev JZ, Jakobsen MD, Sundstrup E. Association between occupational lifting and day-to-day change in low-back pain intensity based on company records and text messages. Scand J Work Environ Health 2017 Jan 01;43(1):68-74. [doi: 10.5271/sjweh.3592] [Medline: 27611578]
11. Sterud T, Johannessen HA, Tynes T. Work-related psychosocial and mechanical risk factors for neck/shoulder pain: a 3-year follow-up study of the general working population in Norway. Int Arch Occup Environ Health 2014 Jul;87(5):471-481. [doi: 10.1007/s00420-013-0886-5] [Medline: 23708752]
12. Sterud T. Work-related mechanical risk factors for long-term sick leave: a prospective study of the general working population in Norway. Eur J Public Health 2014 Feb;24(1):111-116. [doi: 10.1093/eurpub/ckt072] [Medline: 23748849]
13. Andersen LL, Thorsen SV, Flyvholm MA, Holtermann A. Long-term sickness absence from combined factors related to physical work demands: prospective cohort study. Eur J Public Health 2018 May 05:824-829. [doi: 10.1093/eurpub/cky073] [Medline: 29741617]
14. Mattila M, Karwowski W, Vilkki M. Analysis of working postures in hammering tasks on building construction sites using the computerized OWAS method. Appl Ergon 1993 Dec;24(6):405-412. [Medline: 15676938]
15. Hartmann B, Fleischer AG. Physical load exposure at construction sites. Scand J Work Environ Health 2005;31 Suppl 2:88-95 [FREE Full text] [Medline: 16363451]
16. Järvholm B, Stattin M, Robroek SJ, Janlert U, Karlsson B, Burdorf A. Heavy work and disability pension - a long term follow-up of Swedish construction workers. Scand J Work Environ Health 2014 Jul;40(4):335-342 [FREE Full text] [doi: 10.5271/sjweh.3413] [Medline: 24385007]
17. Rinder MM, Genaidy AM, Salem S, Shell R, Karwowski W. Interventions in the construction industry: a systematic review and critical appraisal. Hum Factors Ergon Man 2008 Mar;18(2):212-229. [doi: 10.1002/hfm.20109]
18. Gagnon D, Plamondon A, Larivière C. A biomechanical comparison between expert and novice manual materials handlers using a multi-joint EMG-assisted optimization musculoskeletal model of the lumbar spine. J Biomech 2016 Dec 06;49(13):2938-2945. [doi: 10.1016/j.jbiomech.2016.07.009] [Medline: 27469898]

19. Jakobsen MD, Sundstrup E, Persson R, Andersen CH, Andersen LL. Is Borg's perceived exertion scale a useful indicator of muscular and cardiovascular load in blue-collar workers with lifting tasks? A cross-sectional workplace study. Eur J Appl Physiol 2014 Feb;114(2):425-434 [FREE Full text] [doi: 10.1007/s00421-013-2782-9] [Medline: 24337669]
20. Balaguier R, Madeleine P, Rose-Dulcina K, Vuillerme N. Trunk kinematics and low back pain during pruning among vineyard workers-A field study at the Chateau Larose-Trintaudon. PLoS One 2017;12(4):e0175126 [FREE Full text] [doi: 10.1371/journal.pone.0175126] [Medline: 28384277]
21. Lagersted-Olsen J, Thomsen BL, Holtermann A, Søgaard K, Jørgensen MB. Does objectively measured daily duration of forward bending predict development and aggravation of low-back pain? A prospective study. Scand J Work Environ Health 2016 Jun 01;42(6):528-537. [doi: 10.5271/sjweh.3591] [Medline: 27606607]
22. Robert-Lachaine X, Mecheri H, Larue C, Plamondon A. Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. Med Biol Eng Comput 2017 Apr;55(4):609-619. [doi: 10.1007/s11517-016-1537-2] [Medline: 27379397]
23. Wahlström J, Bergsten E, Trask C, Mathiassen SE, Jackson J, Forsman M. Full-shift trunk and upper arm postures and movements among aircraft baggage handlers. Ann Occup Hyg 2016 Oct;60(8):977-990. [doi: 10.1093/annhyg/mew043] [Medline: 27417186]
24. Samani A, Holtermann A, Søgaard K, Holtermann A, Madeleine P. Following ergonomics guidelines decreases physical and cardiovascular workload during cleaning tasks. Ergonomics 2012;55(3):295-307. [doi: 10.1080/00140139.2011.640945] [Medline: 22409167]
25. Rivilis I, Van Eerd D, Cullen K, Cole DC, Irvin E, Tyson J, et al. Effectiveness of participatory ergonomic interventions on health outcomes: a systematic review. Appl Ergon 2008 May;39(3):342-358. [doi: 10.1016/j.apergo.2007.08.006] [Medline: 17988646]
26. Hignett S, Wilson JR, Morris W. Finding ergonomic solutions--participatory approaches. Occup Med (Lond) 2005 May;55(3):200-207. [doi: 10.1093/occmed/kqi084] [Medline: 15857899]
27. van Eerd D, Cole D, Irvin E, Mahood Q, Keown K, Theberge N, et al. Process and implementation of participatory ergonomic interventions: a systematic review. Ergonomics 2010 Oct;53(10):1153-1166. [doi: 10.1080/00140139.2010.513452] [Medline: 20865600]
28. Driessen MT, Proper KI, van Tulder MW, Anema JR, Bongers PM, van der Beek AJ. The effectiveness of physical and organisational ergonomic interventions on low back pain and neck pain: a systematic review. Occup Environ Med 2010 Apr;67(4):277-285. [doi: 10.1136/oem.2009.047548] [Medline: 20360197]
29. Brandt M, Madeleine P, Samani A, Jakobsen MD, Skals S, Vinstrup J, et al. Accuracy of identification of low or high risk lifting during standardised lifting situations. Ergonomics 2018;61(5):710-719. [doi: 10.1080/00140139.2017.1408857] [Medline: 29171789]
30. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. Br Med J 2012;345:e5661. [Medline: 22951546]
31. Eysenbach G, CONSORT- EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
32. Brandt M, Madeleine P, Ajslev JZ, Jakobsen MD, Samani A, Sundstrup E, et al. Participatory intervention with objectively measured physical risk factors for musculoskeletal disorders in the construction industry: study protocol for a cluster randomized controlled trial. BMC Musculoskelet Disord 2015 Oct 16;16:302 [FREE Full text] [doi: 10.1186/s12891-015-0758-0] [Medline: 26474867]
33. Ajslev J, Brandt M, Møller JL, Skals S, Vinstrup J, Jakobsen MD, et al. Reducing physical risk factors in construction work through a participatory intervention: protocol for a mixed-methods process evaluation. JMIR Res Protoc 2016 May 26;5(2):e89 [FREE Full text] [doi: 10.2196/resprot.5648] [Medline: 27230696]
34. Lewin K. Field Theory in Social Science: Selected Theoretical Papers. 1st edition. New York: Harper & Brothers; 1964.
35. Nielsen KA, Nielsen BS. Methodologies in Action Research: Action Research and Critical Theory. Maastricht: Shaker Publishing; 2006:87.
36. Danish WEA – Arbejdstilsynet. Arbejdsrelateret muskel- og skeletbesvær: At-vejledning D.3.4 [Work-related muscle and skeletal problems] URL: https://arbejdstilsynet.dk/da/regler/at-vejledninger/a/d-3-4-arbejdsrelateret-muskel-og-skeletbesvaer [accessed 2018-10-25] [WebCite Cache ID 73Qn29Kg8]
37. Danish WEA – Arbejdstilsynet. At-vejledning D.3.1 - Løft, træk og skub [Lift, drag, and push] URL: https://arbejdstilsynet.dk/da/regler/at-vejledninger/l/d-3-1-loft-traek-og-skub [accessed 2018-10-25] [WebCite Cache ID 73QlVHtQI]
38. Brandt M, Andersen LL, Samani A, Jakobsen MD, Madeleine P. Inter-day reliability of surface electromyography recordings of the lumbar part of erector spinae longissimus and trapezius descendens during box lifting. BMC Musculoskelet Disord 2017 Dec 11;18(1):519 [FREE Full text] [doi: 10.1186/s12891-017-1872-y] [Medline: 29228936]
39. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of recommendations for SEMG sensors and sensor placement procedures. J Electromyogr Kinesiol 2000 Oct;10(5):361-374. [Medline: 11018445]
40. Seniam. URL: http://www.seniam.org/ [accessed 2018-10-28] [WebCite Cache ID 73VSjBlTu]

41. Korshøj M, Skotte JH, Christiansen CS, Mortensen P, Kristiansen J, Hanisch C, et al. Validity of the Acti4 software using ActiGraph GT3X+accelerometer for recording of arm and upper body inclination in simulated work tasks. Ergonomics 2014;57(2):247-253. [doi: 10.1080/00140139.2013.869358] [Medline: 24392673]
42. Feng Y, Wong CK, Janeja V, Kuber R, Mentis HM. Comparison of tri-axial accelerometers step-count accuracy in slow walking conditions. Gait Posture 2017 Dec;53:11-16. [doi: 10.1016/j.gaitpost.2016.12.014] [Medline: 28064084]
43. Brage S, Brage N, Ekelund U, Luan J, Franks PW, Froberg K, et al. Effect of combined movement and heart rate monitor placement on physical activity estimates during treadmill locomotion and free-living. Eur J Appl Physiol 2006 Mar;96(5):517-524. [doi: 10.1007/s00421-005-0112-6] [Medline: 16344938]
44. Kristiansen J, Korshøj M, Skotte JH, Jespersen T, Søgaard K, Mortensen OS, et al. Comparison of two systems for long-term heart rate variability monitoring in free-living conditions--a pilot study. Biomed Eng Online 2011 Apr 12;10:27 [FREE Full text] [doi: 10.1186/1475-925X-10-27] [Medline: 21481282]
45. Madgwick SO, Harrison AJ, Vaidyanathan A. Estimation of IMU and MARG orientation using a gradient descent algorithm. IEEE Int Conf Rehabil Robot 2011;2011:5975346. [doi: 10.1109/ICORR.2011.5975346] [Medline: 22275550]
46. Jonsson B. Measurement and evaluation of local muscular strain in the shoulder during constrained work. J Hum Ergol (Tokyo) 1982 Sep;11(1):73-88. [Medline: 7186516]
47. Borg G. Borg's Perceived Exertion and Pain Scales. Champaign, IL: Human Kinetics; 1998.
48. Andersen LL, Clausen T, Persson R, Holtermann A. Dose-response relation between perceived physical exertion during healthcare work and risk of long-term sickness absence. Scand J Work Environ Health 2012 Nov;38(6):582-589 [FREE Full text] [doi: 10.5271/sjweh.3310] [Medline: 22714069]
49. Det Nationale Forskningscenter for Arbejdsmiljø. Arbejdsmiljø og helbred i Danmark. Copenhagen: Det Nationale Forskningscenter for Arbejdsmiljø; 2013.
50. Clausen T, Nielsen K, Carneiro IG, Borg V. Job demands, job resources and long-term sickness absence in the Danish eldercare services: a prospective analysis of register-based outcomes. J Adv Nurs 2012 Jan;68(1):127-136. [doi: 10.1111/j.1365-2648.2011.05724.x] [Medline: 21658095]
51. Pejtersen JH, Kristensen TS, Borg V, Bjorner JB. The second version of the Copenhagen Psychosocial Questionnaire. Scand J Public Health 2010 Feb;38(3 Suppl):8-24. [doi: 10.1177/1403494809349858] [Medline: 21172767]
52. Takala EP, Pehkonen I, Forsman M, Hansson GA, Mathiassen SE, Neumann WP, et al. Systematic evaluation of observational methods assessing biomechanical exposures at work. Scand J Work Environ Health 2010 Jan;36(1):3-24 [FREE Full text] [Medline: 19953213]
53. Eliasson K, Palm P, Nyman T, Forsman M. Inter- and intra- observer reliability of risk assessment of repetitive work without an explicit method. Appl Ergon 2017 Jul;62:1-8 [FREE Full text] [doi: 10.1016/j.apergo.2017.02.004] [Medline: 28411720]
54. van der Molen HF, Sluiter JK, Hulshof CT, Vink P, Frings-Dresen MH. Effectiveness of measures and implementation strategies in reducing physical work demands due to manual handling at work. Scand J Work Environ Health 2005;31 Suppl 2:75-87 [FREE Full text] [Medline: 16363450]
55. van der Molen HF, Grouwstra R, Kuijer PP, Sluiter JK, Frings-Dresen MH. Efficacy of adjusting working height and mechanizing of transport on physical work demands and local discomfort in construction work. Ergonomics 2004 Jun 10;47(7):772-783. [doi: 10.1080/00140130420001933 09] [Medline: 15204287]
56. Luijsterburg PA, Bongers PM, de Vroome EM. A new bricklayers' method for use in the construction industry. Scand J Work Environ Health 2005 Oct;31(5):394-400 [FREE Full text] [Medline: 16273966]
57. Dale AM, Jaegers L, Welch L, Gardner BT, Buchholz B, Weaver N, et al. Evaluation of a participatory ergonomics intervention in small commercial construction firms. Am J Ind Med 2016 Jun;59(6):465-475 [FREE Full text] [doi: 10.1002/ajim.22586] [Medline: 27094450]
58. Ajslev JZ, Persson R, Andersen LL. Contradictory individualized self-blaming: a cross-sectional study of associations between expectations to managers, coworkers, one-self and risk factors for musculoskeletal disorders among construction workers. BMC Musculoskelet Disord 2017 Jan 10;18(1):13 [FREE Full text] [doi: 10.1186/s12891-016-1368-1] [Medline: 28068961]
59. Grill M, Pousette A, Nielsen K, Grytnes R, Törner M. Safety leadership at construction sites: the importance of rule-oriented and participative leadership. Scand J Work Environ Health 2017 Jul 01;43(4):375-384. [doi: 10.5271/sjweh.3650] [Medline: 28558110]
60. van der Beek AJ, Dennerlein JT, Huysmans MA, Mathiassen SE, Burdorf A, van Mechelen W, et al. A research framework for the development and implementation of interventions preventing work-related musculoskeletal disorders. Scand J Work Environ Health 2017 Nov 01;43(6):526-539. [doi: 10.5271/sjweh.3671] [Medline: 28945263]
61. Gilmore B, Vallières F, McAuliffe E, Tumwesigye NM, Muyambi G. The last one heard: the importance of an early-stage participatory evaluation for programme implementation. Implement Sci 2014 Sep 26;9:137 [FREE Full text] [doi: 10.1186/s13012-014-0137-5] [Medline: 25253191]
62. Ajslev JZ, Persson R, Andersen LL. Associations between wage system and risk factors for musculoskeletal disorders among construction workers. Pain Res Treat 2015;2015:513903 [FREE Full text] [doi: 10.1155/2015/513903] [Medline: 26605083]

63. Ajslev JZ, Møller JL, Persson R, Andersen LL. Trading health for money: agential struggles in the (re)configuration of subjectivity, the body and pain among construction workers. Work Employ Soc 2016 Oct;31(6):887-903. [doi: 10.1177/0950017016668141]
64. Kramer DM, Bigelow PL, Carlan N, Wells RP, Garritano E, Vi P, et al. Searching for needles in a haystack: identifying innovations to prevent MSDs in the construction sector. Appl Ergon 2010 Jul;41(4):577-584. [doi: 10.1016/j.apergo.2009.12.003] [Medline: 20170903]
65. Ajslev J, Lund H, Møller J, Persson R, Andersen L. Habituating pain: questioning pain and physical strain as inextricable conditions in the construction industry. Nord J Work Life Stud 2013;3(3):195-218.
66. Boatman L, Chaplan D, Teran S, Welch LS. Creating a climate for ergonomic changes in the construction industry. Am J Ind Med 2015 Aug;58(8):858-869. [doi: 10.1002/ajim.22499] [Medline: 26147396]
67. Burdorf A, van der Beek AJ. To RCT or not to RCT: evidence on effectiveness of return-to-work interventions. Scand J Work Environ Health 2016 Jul 01;42(4):257-259. [doi: 10.5271/sjweh.3577] [Medline: 27271241]

Abbreviations

IMU: inertial measurement units
MVC: maximal voluntary contraction
PE: participatory ergonomics
RMS: root mean square
sEMG: surface electromyography
WMSD: work-related musculoskeletal disorder

Edited by G Eysenbach; submitted 01.03.18; peer-reviewed by K Walker-Bone, M Alshehri; comments to author 23.07.18; revised version received 04.09.18; accepted 10.09.18; published 19.12.18.

Please cite as:
Brandt M, Madeleine P, Samani A, Ajslev JZN, Jakobsen MD, Sundstrup E, Andersen LL
Effects of a Participatory Ergonomics Intervention With Wearable Technical Measurements of Physical Workload in the Construction Industry: Cluster Randomized Controlled Trial
J Med Internet Res 2018;20(12):e10272
URL: https://www.jmir.org/2018/12/e10272/
doi:10.2196/10272
PMID:30567694

Original Paper

Digital Weight Loss Intervention for Parents of Children Being Treated for Obesity: A Prospective Cohort Feasibility Trial

Melissa C Kay[1], MPH, MS, RD, PhD; Jasmine Burroughs[2], RD; Sandy Askew[2], MPH; Gary G Bennett[3], PhD; Sarah Armstrong[4], MD; Dori M Steinberg[5], MS, RD, PhD

[1]Duke Global Digital Health Science Center, Duke Center for Childhood Obesity Research, Duke University, Durham, NC, United States

[2]Duke Global Digital Health Science Center, Duke University, Durham, NC, United States

[3]Duke Global Digital Health Science Center, Duke Department of Psychology and Neuroscience, Duke University, Durham, NC, United States

[4]Duke Center for Childhood Obesity Research, Duke Department of Pediatrics, Duke University, Durham, NC, United States

[5]Duke Global Digital Health Science Center, Duke School of Nursing, Duke University, Durham, NC, United States

Corresponding Author:
Melissa C Kay, MPH, MS, RD, PhD
Duke Global Digital Health Science Center
Duke Center for Childhood Obesity Research
Duke University
310 Trent Drive
Durham, NC, 27708
United States
Phone: 1 781 249 3062
Email: melissa.kay@duke.edu

Abstract

Background: The prevalence of childhood obesity continues to increase, and clinic-based treatment options have failed to demonstrate effectiveness. One of the strongest predictors of child weight is parent weight. Parental treatment for weight loss may indirectly reduce obesity in the child. We have previously demonstrated the effectiveness among adults of a fully automated, evidence-based digital weight loss intervention (Track). However, it is unknown if it is feasible to deliver such a treatment directly to parents with obesity who bring their child with obesity to a weight management clinic for treatment.

Objective: The objective of our study was to evaluate the feasibility of and engagement with a digital weight loss intervention among parents of children receiving treatment for obesity.

Methods: We conducted a 6-month pre-post feasibility trial among parents or guardians and their children aged 4-16 years presenting for tertiary care obesity treatment. Along with the standard family-based treatment protocol, parents received a 6-month digital weight loss intervention, which included weekly monitoring of personalized behavior change goals via mobile technologies. We examined levels of engagement by tracking completed weeks of self-monitoring and feasibility by assessing change in weight.

Results: Participants (N=48) were on average 39 years old, mostly female (35/42, 82%), non-Hispanic Black individuals (21/41, 51%) with obesity (36/48, 75%). Over a quarter had a yearly household income of <US $25,000, and about a third had the equivalent of a high school education. Children were on average 10 years old and had a body mass index of 29.8 kg/m^2. The median percentage of weeks participants tracked their behaviors was 77% (18.5/24 total weeks; interquartile range [IQR] 6.3 to 100). The median number of attempts via phone or text message (short message service) required to complete one tracking week was 3.3 (IQR 2.6 to 4.9). Nearly half (23/48, 48%) had high levels of engagement, completing 80% (19/24) or more weeks of tracking. Of the 26 participants with weight measurements reported at 6 months, of which 81% (21/26) were self-reported, there was a median 2.44 kg (IQR −6.5 to 1.0) decrease in weight.

Conclusions: It is feasible to deliver an evidence-based digital weight loss intervention to parents or guardians whose children are enrolled in a weight management program. Given the feasibility of this approach, future studies should investigate the effectiveness of digital weight loss interventions for parents on child weight and health outcomes.

(J Med Internet Res 2018;20(12):e11093) doi:10.2196/11093

KEYWORDS

obesity; parent; child; digital weight loss intervention; self-monitoring; weight; text

Introduction

The prevalence of obesity among children has increased since 1999, and rates among non-Hispanic black and Hispanic children are consistently higher compared with non-Hispanic White children [1-3]. Children with obesity are at increased risk of developing chronic conditions during childhood and during adulthood if obesity persists [4-8]. Children of racial or ethnic minority are disproportionately affected both in terms of obesity and chronic disease.

Recent recommendations from the US Preventive Services Task Force suggest children aged 6 and older with obesity be referred to intensive lifestyle-based weight loss programs [9]. These require 26 or more hours of provider contact with greater effectiveness demonstrated with more contact hours and the incorporation of behavior change techniques such as goal setting and self-monitoring [10]. Although children and parents report positive experiences in behavioral weight loss programs, logistical issues such as clinic hours and location and required time commitment lead to discontinuation of care [11,12]. These high levels of attrition have resulted in poor efficacy [13]. Thus, innovative approaches to pediatric weight management are necessary.

Obesity is highly comorbid in families [14,15]. Although family-based interventions are effective in reducing child body mass index (BMI) [16], they can be time intensive and costly [17]. Yet, parent-only interventions have been effective in the treatment of childhood overweight and obesity [18-20]. Indeed, parent weight change is a strong predictor of child weight change [21,22], in that a 1-unit reduction in parent BMI is associated with a 0.26 reduction in child BMI after participation in a behavioral weight loss program [23]. Because the child weight status is associated with the parent weight status [24-26], parental treatment for weight loss may indirectly reduce obesity in the child by impacting the family's shared environment and through parental role modeling of healthy behaviors. Although pediatric obesity management programs include discussion on changing family behaviors, most programs do not directly and independently treat the parent's obesity. Innovative strategies are needed to consider how best to treat parental obesity while treating children with obesity. Digital health interventions may be well suited to achieve this goal [27].

Digital health approaches capitalize on the ubiquitous utilization of mobile technologies [28], and they have great potential to be scalable and integrated into the existing clinical infrastructure (eg, electronic health records). Digital approaches overcome barriers to parental involvement in weight management programs, such as the time required for attendance and childcare, because they can be asynchronous with care (ie, delivered without requiring real-time interaction). Prior work demonstrates that using mobile technologies to administer weight loss treatment can be successful in the clinic setting [29,30]. We recently demonstrated the effectiveness of "Track," a fully automated, evidence-based digital weight loss program, among adults in a clinic setting [31,32]. In a similar intervention, participants who engaged more, as measured by self-monitoring of behaviors associated with weight loss, lost more weight [33]. Others have demonstrated the importance of user engagement leading to optimal behavior change [34-38]. Measuring engagement is an important measure of fidelity, ensuring that treatments are delivered in the dose intended [39]. Thus, the primary aim of this feasibility study was to measure user engagement, as measured by self-monitoring, after delivering Track to parents or guardians of children with obesity who are presenting for weight management. Assessing feasibility and engagement will aid in determining how best to design future intervention studies.

Methods

Study Design

We delivered a 6-month pre-post feasibility trial called Families on Track to parents or guardians of children seeking treatment for weight management. We recruited participants from the Duke Healthy Lifestyles clinic. Healthy Lifestyles is a referral-based pediatric weight management program located in Durham, NC, which serves a population that is racially and ethnically diverse; 57% are female, 61% are black individuals, 29% are Hispanic individuals, and 70% of patients have public health insurance. The Healthy Lifestyles clinical protocol, patient demographics, and outcomes have been previously described [11,13], and the program represents the current standard of care for obesity treatment. All participants received the Healthy Lifestyles intended clinical treatment protocol. The Duke Medical Center Institutional Review Board approved all procedures.

Participants

Participants included parents or guardians of children aged 4-16 years with an age- and gender-specific BMI of ≥95th percentile presenting for obesity treatment to the Duke Healthy Lifestyles clinic. Eligibility criteria included parents or guardians aged 18-60 years with BMI between 25 and 50 kg/m^2. We required that participants have English fluency, own a mobile phone and be willing to send and receive multiple short message service (SMS) text messages per day, and reside in the same household as the patient attending Healthy Lifestyles. We excluded participants who were pregnant or lactating; had prior or planned bariatric surgery; were participating in other obesity trials; had a history of heart attack, stroke, bipolar disorder, schizophrenia or recent cancer diagnosis; or had plans to relocate within 1 year. We recruited a convenience sample of 50 participants; 2 were excluded (1 did not meet BMI criteria and 1 declined). A total of 48 participants were consented and enrolled by a trained research assistant, who then collected baseline data.

Intervention

The Families on Track intervention included the Healthy Lifestyles program plus a 6-month modified version of Track, a digital weight loss intervention for adults that was conducted in the primary care setting. The Healthy Lifestyles program has been described in detail elsewhere [13]. Briefly, the Healthy Lifestyles program uses best-practice pediatric weight management strategies, which involves monthly visits for patients and their families with medical, dietary, and exercise specialists all certified in Motivational Interviewing. Patients

set dietary and activity behavioral goals aimed to improve the severity of overweight or obesity and obesity-related comorbidities. The Track intervention, summarized elsewhere [31], was modified to contain 4 components: tailored behavioral goals (eg, no sugary drinks, watch less than 2 hours of television per day, and walk 10,000 steps per day); self-monitoring of these goals via interactive voice response (IVR) phone calls or SMS text messages; skills training videos; and an analog bathroom scale and a pedometer to self-monitor daily weights and steps.

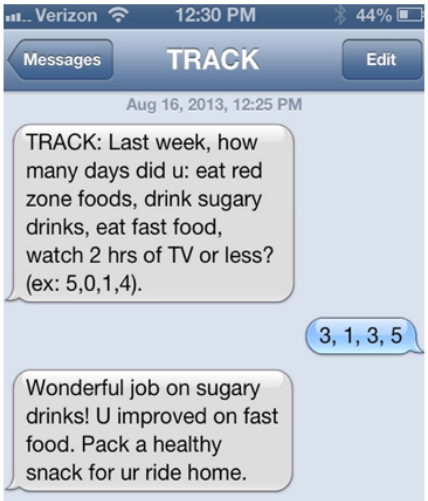
Behavior Change Goals

The intervention utilized the Interactive Obesity Treatment Approach (iOTA), which results in weight loss through the modification of everyday obesogenic behaviors [29,30,40,41]. At baseline, each intervention participant completed a short, self-administered survey to assess the level of engagement in various dietary, physical activity, and other weight control behaviors. A computer algorithm used participant responses to assign personalized behavioral goals from a vast library of goals known to create an energy deficit (eg, *no sugary drinks, no fast food consumption*) based on each participant's need to change each behavior, readiness and self-efficacy, and the potential caloric deficit promoted by the specific behavior change. The algorithm rank orders the goals, and participants are asked to self-monitor their adherence to the top 3 goals for the first 8 weeks of the study. Goals changed every 8 weeks throughout the 24-week intervention period to maintain motivation and facilitate goal mastery. Participants also received a universal 4th goal. We assigned a "*no red zone foods* " goal for the first 8 weeks. To determine the "red zone foods," we asked participants to select the foods they consume regularly (at least 3 days per week) from a list of commonly eaten, high-calorie foods and beverages (eg, sodas, sweet tea, desserts, potato chips, pizza, and hamburgers). The other universal goals were to "*practice portion control*" and "*walk 7-10,000 steps per day.*" We provided all intervention participants with Web links to a study-specific YouTube channel that included descriptive and skills training videos specific to each Track goal. We reminded participants to refer to the videos for additional skills training and behavior change tips, specifically when goals changed every 8 weeks.

User Engagement

We measured user engagement with the intervention both quantitatively and qualitatively. Using quantitative measures, we tracked the frequency of weekly self-monitoring across the 24-week intervention. Participants were expected to self-monitor daily via paper log and weekly using the IVR system or through SMS text message. Each week, participants received an automated prompt from the Families on Track intervention system to track adherence to behaviors goals. These prompts were delivered either via IVR or SMS text message. The IVR system called intervention participants weekly to request self-monitoring data and provided automated tailored feedback on the 4 goals. Participants who did not respond to IVR attempts were sent a SMS text message prompting them to communicate their weekly tracking data via SMS text messages (Figure 1). Participants who provided self-monitoring data via SMS text messaging also received tailored feedback. We have a robust retry protocol that attempts to reach participants if the first IVR call or SMS text message goes unanswered. Tracking was considered complete if the participant completed the entire weekly IVR call or responded to the weekly SMS text message. User engagement with the intervention was assessed by totaling the number of weeks each participant responded to prompts to track behavior across the 24 weeks. In addition, we created a dichotomous outcome variable to compare those who were high versus low engagers using an established cutoff of 80% or more engagement in weekly self-monitoring [33,42]. We also tracked the mean number of prompts required to elicit a response for each participant as an additional measure of user engagement. For a qualitative measure of user engagement, participants were asked to complete satisfaction surveys upon study completion to assess the acceptability of the message frequency, timing, content, and perceived usefulness. Prior to their 6-month follow-up, participants were prompted to complete the satisfaction survey. Attempts were made via phone, email, and SMS text messages to complete the survey even if a follow-up appointment could not be scheduled.

Figure 1. Example of a self-monitoring SMS text message sent weekly to participants in Families on Track (Interactive Obesity Treatment Approach [iOTA], Duke Global Digital Health Science Center).



Weight

A trained nurse in the Healthy Lifestyles clinic collected parent and guardian height and weight at baseline and at 6 months; we measured height using a stadiometer (Model: Healthometer Professional CE No 92977) and weight using a digital scale (Model: Seca CE No 96990). A high percentage of participants did not return to the Healthy Lifestyles clinic for follow-up appointments despite email, phone, and SMS text message reminders. Therefore, we experienced difficulties in scheduling the 6-month visits. Thus, we also collected weights via self-report. Self-reported weights were sent to the study staff via email, SMS text message, or phone from participants who were unable to complete their in-clinic study visit. To verify self-reported weights, participants were asked to SMS text message or email the study staff a photo of their feet on their study-issued or personal scale with a visible weight reading.

Other Measurements

Sociodemographic variables were measured using standard questionnaires completed by the parent or guardian at the baseline clinic visit.

Analysis

We used descriptive statistics to characterize participants and examine tracking completion rate and weight change over the 6-month period. Characteristics were summarized using frequencies and proportions for categorical variables and mean (SD) for continuous variables. We used medians and interquartile ranges (IQR) to summarize intervention engagement and weight change owing to its highly skewed distribution. We conducted bivariate analyses to examine potential predictors of intervention engagement using Wilcoxon-Mann-Whitney and Kruskal Wallis tests for continuous data and chi-square and Fisher's exact tests for categorical data. We used Poisson regression with a robust variance to examine sociodemographic differences among those with higher levels of tracking engagement (80% or more weeks of tracking) and estimate risk ratios (RRs) and 95% CIs. To assess intervention feasibility, we assessed differences in weight change among high and low engagers using the Wilcoxon-Mann-Whitney test. We conducted all analyses using Stata 14 for Mac (StataCorp, College Station, TX) with an alpha value of <.05 to assess statistical significance.

Results

Baseline Characteristics

At baseline (N=48), participants were on average 39.4 years old (SD 7.3) with a mean BMI of 36.5 kg/m^2 (Table 1). Half (21/41, 51%) were non-Hispanic black individuals. Most were female (35/42, 83%) and the mother of the child (34/41, 83%) and many were employed (33/41, 81%). Over a quarter, 26% (10/38) had an income of <US $25,000, and the highest level of education for over a third of participants (13/37, 35%) was a high school equivalent.

User Engagement

At least half of the participants engaged in tracking their behaviors in each study week, as measured by a complete IVR call or SMS text message (Figure 2). The median engagement rate was 77% (IQR 6.3 to 100) across all study weeks. A fifth of participants (10/48, 21%) did not track any of their behaviors, and 27% (13/48) completed all tracking weeks. Nearly half (23/48, 48%) of the participants were considered high engagers (based on a median split), tracking their behaviors for at least 80% (19/24) of study weeks. The median number of prompts required to get participants to complete 1 tracking week (either through IVR or SMS text messages) was 3.3 (IQR 2.6 to 4.9). Most of the tracking was completed via SMS text messages (87%) than with IVR. The average duration in minutes for those who did complete IVR calls was 0.5 (SD 0.9).

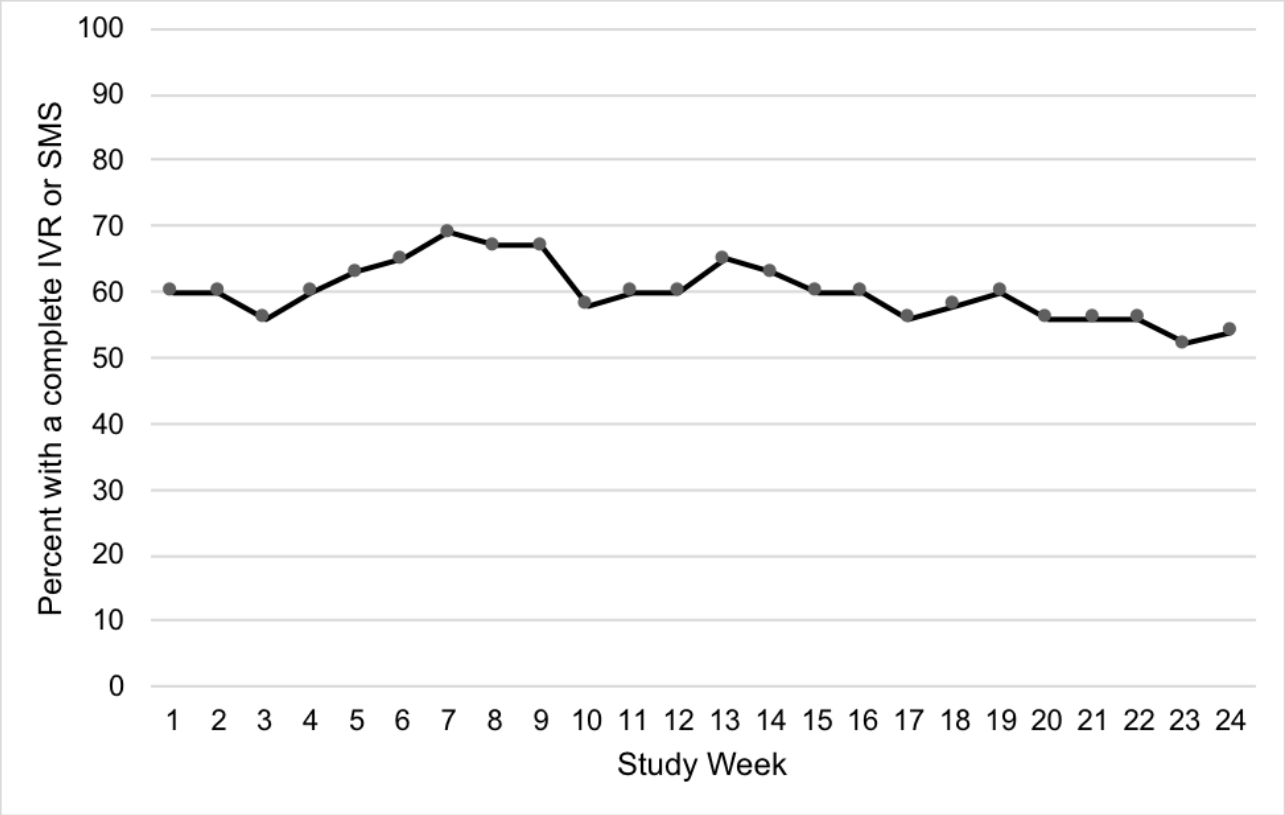
Among the included participants, 40% (19/48) completed the post intervention satisfaction survey. Those with complete surveys completed more weeks of tracking, 21.3 (SD 3.8) versus 9.9 (SD 10.2) with P<.001, and were more likely to be high engagers (P=.005). All respondents (19/19, 100%) felt it was easy to understand their 4 Track goals, among whom 89% (16/18) felt confident they could follow them and 89% (17/19) felt they were what they needed to work on to lose weight. Most (16/19, 84%) liked being able to choose each week whether they responded to tracking requests via SMS text messages or IVR. A few felt the weekly automated calls (5/19, 26%) and texts (2/19, 11%) were difficult to understand, but most (16/19, 84%) felt the feedback received on them was helpful. About a quarter of the respondents (5/19, 26%) found getting started in Track was hard. Of those receiving SMS text messages (10/19, 53%), most reported the number of texts was just enough.

Table 1. Baseline characteristics of parents or guardians participating in a digital weight loss intervention.

Characteristics at enrollment	Value
Parent or guardian characteristics	
Race or ethnicity (N=41), n (%)	
Non-Hispanic black	21 (51)
Hispanic	5 (12)
Non-Hispanic white	11 (27)
Other	2 (5)
Declined	2 (5)
Age (N=42), mean (SD)	39.4 (7.9)
Gender (N=42), n (%)	
Female	35 (83)
Relation to the child (N=41), n (%)	
Mother	34 (83)
Father	6 (15)
Grandmother	1 (2)
Employed (N=41), n (%)	33 (81)
Income level (N=38), n (%)	
<US $25,000	10 (26)
US $25,000-34,999	13 (34)
≥ US $35,000	15 (40)
Education (N=37), n (%)	
High School equivalent	13 (35)
Tech or community college	9 (24)
College degree or more	15 (41)
Married (N=41), n (%)	21 (51)
Household size (N=42), mean (SD)	4.2 (1.3)
BMI[a] (N=48), mean (SD)	36.5 (8.0)
Child characteristics	
Race or ethnicity (N=48), n (%)	
Non-Hispanic black	25 (52)
Hispanic	6 (13)
Non-Hispanic white	9 (19)
Pacific Islander	1 (2)
Other	5 (10)
Declined	2 (4)
Age (N=48), mean (SD)	10.0 (3.4)
Gender (N=48), n (%)	
Female	27 (56)
BMI (N=48), mean (SD)	29.8 (7.9)

[a]BMI: body mass index.

Figure 2. Proportion of participants with a complete tracking week as measured by a completed IVR call or SMS text message, by study week (N=48). IVR: interactive voice response; SMS: short message service.



Predictors of User Engagement

The percent of tracking weeks completed was positively associated with education and income (P=.01 and P<.001, respectively; Figures 3 and 4). Income and parent race or ethnicity were associated with level of engagement. Participants with incomes >US $35,000 per year were 4 times as likely to be high engagers (ie, completed >80% of tracking weeks) compared with those with incomes <US $25,000 (RR 4.0; 95% CI 1.1-14.4; P=.03); this relationship was attenuated after controlling for parent race or ethnicity, though remaining significant (RR 3.5; 95% CI 1.1-11.4; P=.04). Non-Hispanic White individuals were twice as likely to be high engagers (RR 2.1; 95% CI 1.2-4.0; P=.02) compared with non-Hispanic black individuals, though this relationship was not significant when controlling for income (RR 1.5; 95% CI 0.8-2.6; P=.19). As child age increased, participants were less likely to be high engagers (RR 0.9; 95% CI 0.8-1.0; P=.04).

Figure 3. Proportion of participants with a complete tracking week as measured by a completed IVR call or SMS text messages, by study week and income level (N=38).

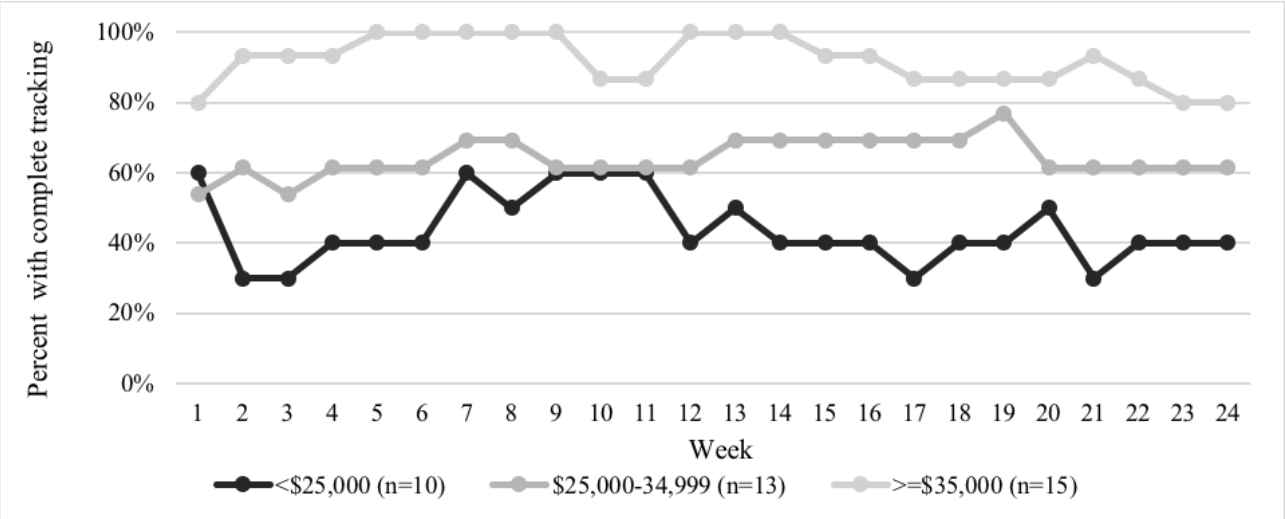
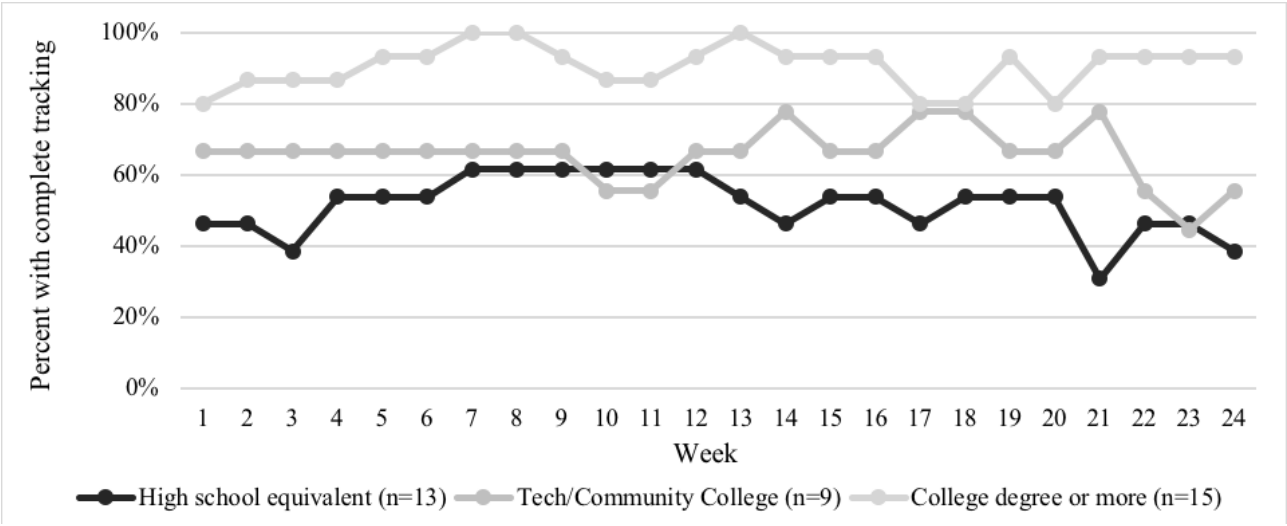


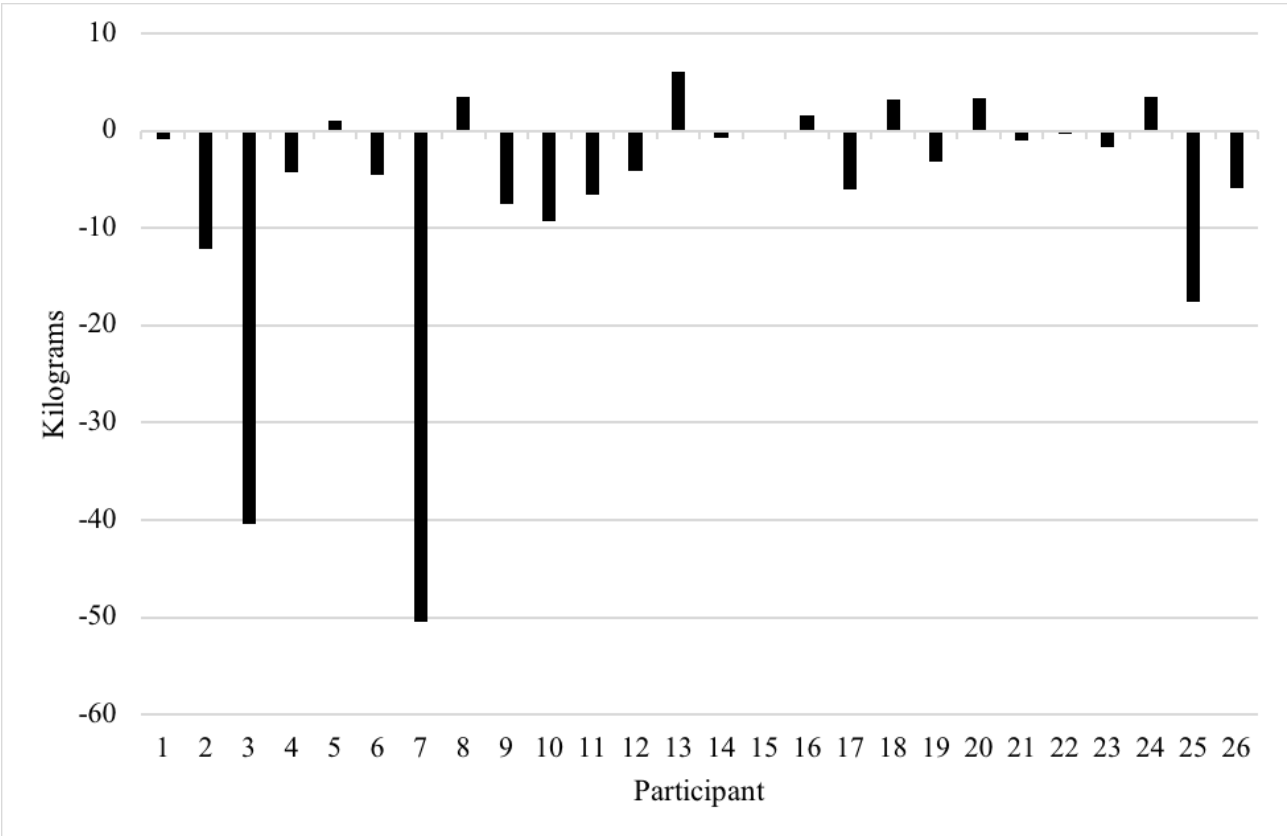
Figure 4. Proportion of participants with a complete tracking week as measured by a completed IVR call or SMS text message by study week and education level (N=37).



Weight Change

At the 6-month intervention completion point, weight was recorded from 54% (26/48) of the participants. Of those, 81% (21/26) were self-reported. There were no significant sociodemographic differences among those with a self-reported weight versus those who were missing weight measurements at 6 months. Those who reported weight at 6 months tracked significantly more weeks than those who did not report weight, 17.3 (SD 8.7) versus 11.0 (SD 10.4), *P*=.03; however, they were not more likely to be high engagers (*P*=.14). Of the 26 participants with complete pre and post intervention weight data, there was significant median weight loss of 2.44 kg (IQR −6.5 to 1.0; *P*=.01; Figure 5). Many (18/26, 69%) had a net weight loss, whereas few (7/26, 27%) had a net weight gain. There was no difference in weight change among high and low engagers. We conducted a sensitivity analysis to exclude those with self-reported weight loss of >40 kg (N=2); the results remained significant with a median weight loss of 1.3 kg (IQR −6.0 to 1.3; *P*=.04).

Figure 5. Change in weight among participants with complete pre and post weight data participating in a digital behavioral weight loss intervention (N=26).



Discussion

Principal Findings

These findings suggest that we can feasibly recruit and engage parents or guardians who are attending a pediatric weight management program for their child with obesity in a digital weight loss intervention. Almost half of the participants (23/48, 48%) were high engagers, tracking behaviors for 80% (19/24) or more study weeks. What is most notable about this study is the parent focus. We aimed to recruit a sample of parents who were interested in obesity treatment for their children, but what we found is that some parents did not engage in self-monitoring of behaviors that result in weight loss for themselves, despite presenting for treatment for their children. We were able to recruit and engage parents, but we had difficulties retaining them and asking them to complete study assessments for evaluation. However, our study did demonstrate favorable behavioral outcomes. Most of the participants who reported weight upon study completion experienced weight loss and found the intervention easy to participate in with accurate goals and helpful feedback. Although this study was not designed to establish efficacy, it is promising that among this group of participants with relatively high engagement, there was significant weight loss.

The results from the Families on Track study are similar to what we found in the Shape Plan trial, which aimed to test the feasibility of delivering daily SMS text messages tracking behavioral goals [41]. In that study, we found that 85% tracked at least 2 days per week and the average weight loss was around 2 kg after 6 months [29]. Finding similar feasibility and weight loss findings suggest that a standalone approach to weight loss that focuses primarily on tracking behavioral goals through mobile technologies can be effective for parents of children presenting for obesity treatment or other adult populations.

Our findings are comparable to reported engagement outcomes from other similar behavioral interventions using SMS text messaging or other digital health modalities. Among breast cancer survivors who were overweight and participated in a 10-week multifaceted mobile health study, engagement with SMS text messaging was 86% [43]. In a year-long childhood obesity reduction intervention targeting parents and their children, 66% of parents were considered high completers for SMS text message response rates [44]. A unique aspect to the Track system, which likely contributed to high engagement, is its ability to provide fully automated, tailored feedback based on participant behaviors [45]. Many intervention studies have relied on one-way SMS text messaging delivering less personalized, more static content. Engagement and effectiveness can be increased by adding other components, such as the provision of human support, but requires greater cost and intensity [46]. That greater cost and intensity may not be feasible to deliver to parents with children presenting for obesity treatment.

Studies show that mobile phones can be an effective tool in weight loss interventions, given the increased ease in self-monitoring behaviors compared with using typical paper logs [47]. Participant engagement in Families on Track was largely completed via SMS text messages than with IVR, which was contrary to similar studies in which IVR was the preferred modality [48]. Parents might find it easier to engage in SMS text messages given they can respond at a time that suits them and are provided visual feedback, which they can retain and refer to.

Involving parents in weight-related behavior change interventions has demonstrated effectiveness in reducing child overweight or obesity [18-20]. However, the best way to support parents of children with obesity is not well known. Few pediatric weight management clinics or organizations have the resources to provide a parent-only approach in addition to childhood obesity treatment. Most pediatric weight management clinics are not well equipped to care for adult health. Additionally, parents are not uniformly engaged in their own weight management when they bring their child to weight management programs, making it difficult to determine the most generalizable way to engage parents. Future studies are needed to determine the best way to engage parents of children with obesity in a way that meets various levels of motivation without high burden. Targeting parents based on characteristics associated with higher levels of engagement and feasibility may be the best approach. We found that parents of older children were less likely to be highly engaged, meaning effective treatment strategies can vary by the age of the child. Family-based interventions may be more effective for parents of older adolescents because these children are more autonomous and make many of their own decisions regarding food choice. They are also able to help when it comes to cooking and meal planning.

Our study was not immune to the disparities or inequalities in engagement seen in other digital health interventions [33,48]. Although overall participants were of lower socioeconomic status, parents or guardians with higher levels of education and income demonstrated higher engagement. This speaks to the importance of designing digital weight loss interventions that are adapted to the needs and habits of various social groups, particularly those most vulnerable and at risk for obesity. Studies show that it is possible to reach and engage more socioeconomically disadvantaged populations with a lower intensity digital health intervention [49], but more work needs to be done to ensure a broader reach and consistent engagement concurrent with positive behavioral outcomes.

Limitations

Limitations to our feasibility study include a small sample size and lack of a control group. We felt it would be difficult to withhold treatment from parents who are already presenting for their child's treatment. An additional limitation is that parents or guardians that attend tertiary care clinics for their children may have different motivations, especially given they are presenting for their children and not themselves. Future studies are needed to assess true generalizability among parents within the general population and also within primary care clinics. Although the results of this study show promise having achieved high engagement, more research is needed to assess behavioral changes as a result of engagement in this population. A large limitation is that our weight change and qualitative user engagement data are not complete given the lack of returning

clinic visits. We also collected postintervention weight primarily through self-report. As mentioned earlier, we provided this option because it was difficult to have participants return for assessment visits. This is likely because of the way the Healthy Lifestyle program is structured—children do not present for treatment often after the initial treatment is provided in the first month. It may be that parents were not interested in attending if their children were not attending for their own treatment. As a result, it is possible our results are biased toward a larger effect. However, previous evidence does suggest that self-reported weights provide a reasonably accurate measurement among adults [50].

Conclusion

In this feasibility study, we demonstrate that it is possible to engage parents or guardians of children with obesity in a digital weight loss intervention aimed at reducing parent weight. The digital intervention engaged a population of parents who are hard to reach through in-person visits and shows promise for reaching and engaging parents in future family-based obesity treatment interventions, an important aspect of intervention fidelity.

Acknowledgments

The authors acknowledge the Healthy Lifestyles clinic staff and research assistants, specifically Jacob Christy, for their help in completing this project.

Conflicts of Interest

GGB is on the Nutrisystem advisory board and has equity in Coeus Health, LLC.

References

1. Ogden C, Carroll M, Fryar C, Flegal K. Prevalence of Obesity Among Adults and Youth: United States, 2011-2014. NCHS Data Brief 2015 Nov(219):1-8 [FREE Full text] [Medline: 26633046]
2. Skinner AC, Perrin EM, Skelton JA. Prevalence of obesity and severe obesity in US children, 1999-2014. Obesity (Silver Spring) 2016 Dec;24(5):1116-1123 [FREE Full text] [doi: 10.1002/oby.21497] [Medline: 27112068]
3. Skinner A, Ravanbakht S, Skelton J, Perrin E, Armstrong S. Prevalence of Obesity and Severe Obesity in US Children, 1999-2016. Pediatrics 2018 Mar;141(3):1-11 [FREE Full text] [doi: 10.1542/peds.2017-3459] [Medline: 29483202]
4. Skinner AC, Perrin EM, Moss LA, Skelton JA. Cardiometabolic Risks and Severity of Obesity in Children and Young Adults. N Engl J Med 2015 Oct;373(14):1307-1317. [doi: 10.1056/NEJMoa1502821] [Medline: 26422721]
5. Freedman D, Mei Z, Srinivasan S, Berenson G, Dietz W. Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa Heart Study. J Pediatr 2007 Jan;150(1):12-17.e2. [doi: 10.1016/j.jpeds.2006.08.042] [Medline: 17188605]
6. Lloyd LJ, Langley-Evans SC, McMullen S. Childhood obesity and risk of the adult metabolic syndrome: a systematic review. Int J Obes (Lond) 2012 Jan;36(1):1-11 [FREE Full text] [doi: 10.1038/ijo.2011.186] [Medline: 22041985]
7. Mohanan S, Tapp H, McWilliams A, Dulin M. Obesity and asthma: pathophysiology and implications for diagnosis and management in primary care. Exp Biol Med (Maywood) 2014 Nov;239(11):1531-1540 [FREE Full text] [doi: 10.1177/1535370214525302] [Medline: 24719380]
8. Morrison K, Shin S, Tarnopolsky M, Taylor V. Association of depression & health related quality of life with body composition in children and youth with obesity. J Affect Disord 2015 Feb 01;172:18-23. [doi: 10.1016/j.jad.2014.09.014] [Medline: 25451390]
9. US Preventive Services Task Force, Grossman DC, Bibbins-Domingo K, Curry SJ, Barry MJ, Davidson KW, et al. Screening for Obesity in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. JAMA 2017 Jun 20;317(23):2417-2426. [doi: 10.1001/jama.2017.6803] [Medline: 28632874]
10. O'Connor EA, Evans CV, Burda BU, Walsh ES, Eder M, Lozano P. Screening for Obesity and Intervention for Weight Management in Children and Adolescents: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 2017 Jun 20;317(23):2427-2444. [doi: 10.1001/jama.2017.0332] [Medline: 28632873]
11. Dolinsky D, Armstrong S, Østbye T. Predictors of attrition from a clinical pediatric obesity treatment program. Clin Pediatr (Phila) 2012 Dec;51(12):1168-1174. [doi: 10.1177/0009922812458355] [Medline: 22935220]
12. Skelton JA, Martin S, Irby MB. Satisfaction and attrition in paediatric weight management. Clin Obes 2016 Apr;6(2):143-153. [doi: 10.1111/cob.12138] [Medline: 27008068]
13. Dolinsky D, Armstrong S, Walter E, Kemper A. The effectiveness of a primary care-based pediatric obesity program. Clin Pediatr (Phila) 2012 Apr;51(4):345-353. [doi: 10.1177/0009922811425232] [Medline: 22013148]
14. Lake JK, Power C, Cole TJ. Child to adult body mass index in the 1958 British birth cohort: associations with parental obesity. Arch Dis Child 1997 Nov;77(5):376-381 [FREE Full text] [Medline: 9487953]
15. Garn SM, Sullivan TV, Hawthorne VM. Fatness and obesity of the parents of obese individuals. Am J Clin Nutr 1989 Dec;50(6):1308-1313. [doi: 10.1093/ajcn/50.6.1308] [Medline: 2596422]

16. Berge JM, Everts JC. Family-Based Interventions Targeting Childhood Obesity: A Meta-Analysis. Child Obes 2011 Apr;7(2):110-121 [FREE Full text] [doi: 10.1089/chi.2011.07.02.1004.berge] [Medline: 26182126]
17. Janicke DM, Steele RG, Gayes LA, Lim CS, Clifford LM, Schneider EM, et al. Systematic review and meta-analysis of comprehensive behavioral family lifestyle interventions addressing pediatric obesity. J Pediatr Psychol 2014 Sep;39(8):809-825. [doi: 10.1093/jpepsy/jsu023] [Medline: 24824614]
18. Ewald H, Kirby J, Rees K, Robertson W. Parent-only interventions in the treatment of childhood obesity: a systematic review of randomized controlled trials. J Public Health (Oxf) 2014 Sep;36(3):476-489. [doi: 10.1093/pubmed/fdt108] [Medline: 24273229]
19. Yavuz HM, van Ijzendoorn MH, Mesman J, van der Veek S. Interventions aimed at reducing obesity in early childhood: a meta-analysis of programs that involve parents. J Child Psychol Psychiatry 2014 Oct 8:677-692. [doi: 10.1111/jcpp.12330] [Medline: 25292319]
20. Niemeier B, Hektner J, Enger K. Parent participation in weight-related health interventions for children and adolescents: a systematic review and meta-analysis. Prev Med 2012 Jul;55(1):3-13. [doi: 10.1016/j.ypmed.2012.04.021] [Medline: 22575353]
21. Wrotniak B, Epstein L, Paluch R, Roemmich J. Parent weight change as a predictor of child weight change in family-based behavioral obesity treatment. Arch Pediatr Adolesc Med 2004 Apr;158(4):342-347. [doi: 10.1001/archpedi.158.4.342] [Medline: 15066873]
22. Andriani H, Liao C, Kuo H. Parental weight changes as key predictors of child weight changes. BMC Public Health 2015 Jul 12;15:645 [FREE Full text] [doi: 10.1186/s12889-015-2005-x] [Medline: 26164227]
23. Boutelle KN, Cafri G, Crow SJ. Parent predictors of child weight change in family based behavioral obesity treatment. Obesity (Silver Spring) 2012 Jul;20(7):1539-1543 [FREE Full text] [doi: 10.1038/oby.2012.48] [Medline: 22421896]
24. Davis M, McGonagle K, Schoeni R, Stafford F. Grandparental and parental obesity influences on childhood overweight: implications for primary care practice. J Am Board Fam Med 2008;21(6):549-554 [FREE Full text] [doi: 10.3122/jabfm.2008.06.070140] [Medline: 18988722]
25. Liu Y, Chen H, Liang L, Wang Y. Parent-child resemblance in weight status and its correlates in the United States. PLoS One 2013;8(6):e65361 [FREE Full text] [doi: 10.1371/journal.pone.0065361] [Medline: 23762352]
26. Bushnik T, Garriguet D, Colley R. Parent-Child association in body weight status. Health Rep 2017 Jun 21;28(6):12-19 [FREE Full text] [Medline: 28636069]
27. Hammersley ML, Jones RA, Okely AD. Parent-Focused Childhood and Adolescent Overweight and Obesity eHealth Interventions: A Systematic Review and Meta-Analysis. J Med Internet Res 2016 Jul 21;18(7):e203 [FREE Full text] [doi: 10.2196/jmir.5893] [Medline: 27443862]
28. Pew Research Center. 2015. The Smartphone Difference URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2018-11-01] [WebCite Cache ID 73bwdu0kf]
29. Bennett GG, Foley P, Levine E, Whiteley J, Askew S, Steinberg DM, et al. Behavioral treatment for weight gain prevention among black women in primary care practice: a randomized clinical trial. JAMA Intern Med 2013 Oct 28;173(19):1770-1777 [FREE Full text] [doi: 10.1001/jamainternmed.2013.9263] [Medline: 23979005]
30. Bennett GG, Warner ET, Glasgow RE, Askew S, Goldman J, Ritzwoller DP, et al. Obesity treatment for socioeconomically disadvantaged patients in primary care practice. Arch Intern Med 2012 Apr 9;172(7):565-574 [FREE Full text] [doi: 10.1001/archinternmed.2012.1] [Medline: 22412073]
31. Foley P, Steinberg D, Levine E, Askew S, Batch BC, Puleo EM, et al. Track: A randomized controlled trial of a digital health obesity treatment intervention for medically vulnerable primary care patients. Contemp Clin Trials 2016 Dec;48:12-20 [FREE Full text] [doi: 10.1016/j.cct.2016.03.006] [Medline: 26995281]
32. Bennett G, Steinberg D, Askew S, Levine E, Foley P, Batch B, et al. Effectiveness of an App and Provider Counseling for Obesity Treatment in Primary Care. Am J Prev Med 2018 Oct 17:a. [doi: 10.1016/j.amepre.2018.07.005] [Medline: 30361140]
33. Steinberg DM, Levine EL, Lane I, Askew S, Foley PB, Puleo E, et al. Adherence to self-monitoring via interactive voice response technology in an eHealth intervention targeting weight gain prevention among Black women: randomized controlled trial. J Med Internet Res 2014;16(4):e114 [FREE Full text] [doi: 10.2196/jmir.2996] [Medline: 24780934]
34. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
35. Funk KL, Stevens VJ, Appel LJ, Bauck A, Brantley PJ, Champagne CM, et al. Associations of internet website use with weight change in a long-term weight loss maintenance program. J Med Internet Res 2010;12(3):e29 [FREE Full text] [doi: 10.2196/jmir.1504] [Medline: 20663751]
36. Couper MP, Alexander GL, Zhang N, Little RJA, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]

37. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
38. Sharpe EE, Karasouli E, Meyer C. Examining Factors of Engagement With Digital Interventions for Weight Management: Rapid Review. JMIR Res Protoc 2017 Oct 23;6(10):e205 [FREE Full text] [doi: 10.2196/resprot.6059] [Medline: 29061557]
39. Bellg A, Borrelli B, Resnick B, Hecht J, Minicucci D, Ory M, Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004 Sep;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]
40. Steinberg DM, Bennett GG, Askew S, Tate DF. Weighing every day matters: daily weighing improves weight loss and adoption of weight control behaviors. J Acad Nutr Diet 2015 Apr;115(4):511-518. [doi: 10.1016/j.jand.2014.12.011] [Medline: 25683820]
41. Foley P, Levine E, Askew S, Puleo E, Whiteley J, Batch B, et al. Weight gain prevention among black women in the rural community health center setting: the Shape Program. BMC Public Health 2012 Jun 15;12:305 [FREE Full text] [doi: 10.1186/1471-2458-12-305] [Medline: 22537222]
42. Vincze G, Barner JC, Lopez D. Factors associated with adherence to self-monitoring of blood glucose among persons with diabetes. Diabetes Educ 2004;30(1):112-125. [doi: 10.1177/014572170403000119] [Medline: 14999899]
43. Quintiliani LM, Mann DM, Puputti M, Quinn E, Bowen DJ. Pilot and Feasibility Test of a Mobile Health-Supported Behavioral Counseling Intervention for Weight Management Among Breast Cancer Survivors. JMIR Cancer 2016;2(1):e4 [FREE Full text] [doi: 10.2196/cancer.5305] [Medline: 27761518]
44. Price S, Ferisin S, Sharifi M, Steinberg D, Bennett G, Wolin KY, et al. Development and Implementation of an Interactive Text Messaging Campaign to Support Behavior Change in a Childhood Obesity Randomized Controlled Trial. J Health Commun 2015;20(7):843-850. [doi: 10.1080/10810730.2015.1018582] [Medline: 25996181]
45. Lyzwinski LN, Caffery LJ, Bambling M, Edirippulige S. Consumer perspectives on mHealth for weight loss: a review of qualitative studies. J Telemed Telecare 2017 Jan 01:1357633X17692722. [doi: 10.1177/1357633X17692722] [Medline: 28181859]
46. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions - A systematic review. Internet Interventions 2014 Oct;1(4):205-215 [FREE Full text]
47. Liu F, Kong X, Cao J, Chen S, Li C, Huang J, et al. Mobile phone intervention and weight loss among overweight and obese adults: a meta-analysis of randomized controlled trials. Am J Epidemiol 2015 Mar 1;181(5):337-348. [doi: 10.1093/aje/kwu260] [Medline: 25673817]
48. Wolin K, Steinberg D, Lane I, Askew S, Greaney M, Colditz G, et al. Engagement with eHealth Self-Monitoring in a Primary Care-Based Weight Management Intervention. PLoS One 2015;10(10):e0140455 [FREE Full text] [doi: 10.1371/journal.pone.0140455] [Medline: 26469065]
49. Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS. The efficacy of a daily self-weighing weight loss intervention using smart scales and e-mail. Obesity (Silver Spring) 2013 Sep;21(9):1789-1797 [FREE Full text] [doi: 10.1002/oby.20396] [Medline: 23512320]
50. Spencer E, Appleby P, Davey G, Key T. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. Public Health Nutr 2002 Aug;5(4):561-565. [doi: 10.1079/PHN2001322] [Medline: 12186665]

Abbreviations

BMI: body mass index
iOTA: Interactive Obesity Treatment Approach
IQR: interquartile range
IVR: interactive voice response
RR: risk ratios
SMS: short message service

Edited by G Eysenbach; submitted 18.05.18; peer-reviewed by MA Styn, Y Huang; comments to author 26.07.18; revised version received 06.09.18; accepted 10.09.18; published 20.12.18.

Please cite as:
Kay MC, Burroughs J, Askew S, Bennett GG, Armstrong S, Steinberg DM
Digital Weight Loss Intervention for Parents of Children Being Treated for Obesity: A Prospective Cohort Feasibility Trial
J Med Internet Res 2018;20(12):e11093
URL: http://www.jmir.org/2018/12/e11093/
doi:10.2196/11093
PMID:30573449

Original Paper

Using a Mobile Social Networking App to Promote Physical Activity: A Qualitative Study of Users' Perspectives

Huong Ly Tong[1], B Health; Enrico Coiera[1], MBBS, PhD; Liliana Laranjo[1], MD, MPH, PhD

Centre for Health Informatics, Australian Institute of Health Innovation, Sydney, Australia

Corresponding Author:
Huong Ly Tong, B Health
Centre for Health Informatics
Australian Institute of Health Innovation
Level 6
75 Talavera Road
Sydney, 2109
Australia
Phone: 61 029850 ext 2475
Email: huong-ly.tong@students.mq.edu.au

Abstract

Background: Despite many health benefits of physical activity, nearly a third of the world's adult population is insufficiently active. Technological interventions, such as mobile apps, wearable trackers, and Web-based social networks, offer great promise in promoting physical activity, but little is known about users' acceptability and long-term engagement with these interventions.

Objective: The aim of this study was to understand users' perspectives regarding a mobile social networking intervention to promote physical activity.

Methods: Participants, mostly university students and staff, were recruited using purposive sampling techniques. Participants were enrolled in a 6-month feasibility study where they were provided with a wearable physical activity tracker (Fitbit Flex 2) and a wireless scale (Fitbit Aria) integrated with a social networking mobile app (named "fit.healthy.me"). We conducted semistructured, in-depth qualitative interviews and focus groups pre- and postintervention, which were recorded and transcribed verbatim. The data were analyzed in Nvivo 11 using thematic analysis techniques.

Results: In this study, 55 participants were enrolled; 51% (28/55) were females, and the mean age was 23.6 (SD 4.6) years. The following 3 types of factors emerged from the data as influencing engagement with the intervention and physical activity: individual (self-monitoring of behavior, goal setting, and feedback on behavior), social (social comparison, similarity and familiarity between users, and participation from other users in the network), and technological. In addition, automation and personalization were observed as enhancing the delivery of both individual and social aspects. Technological limitations were mentioned as potential barriers to long-term usage.

Conclusions: Self-regulatory techniques and social factors are important to consider when designing a physical activity intervention, but a one-size-fits-all approach is unlikely to satisfy different users' preferences. Future research should adopt innovative research designs to test interventions that can adapt and respond to users' needs and preferences throughout time.

(J Med Internet Res 2018;20(12):e11439) doi:10.2196/11439

KEYWORDS

exercise; fitness trackers; mobile apps; mobile phone; social networking

Introduction

Physical inactivity has been identified by the World Health Organization as a global public health problem, emerging as the fourth leading risk factor for global mortality [1]. Research has shown that physical inactivity increases the risk of many chronic diseases—most notably, type 2 diabetes, coronary heart disease, and colon cancer [2]. Nearly a third of adults worldwide are insufficiently active [3], highlighting the need for effective health interventions to change behavior and promote physical activity.

It is widely acknowledged that behavior change is a challenging process. The success of behavior change depends not only on an individual but also on social and environmental factors [4,5].

Behavior change interventions are usually complex (ie, involving several interacting components), which makes it hard to identify what is effective in changing a particular behavior, for whom, and in what context [6-8]. Several taxonomies for behavior change techniques (ie, the active components in health behavior change interventions) have been developed [9,10] in an attempt to isolate and identify the most effective components of interventions. For physical activity promotion, some behavior change techniques seem to be particularly relevant such as self-monitoring of behavior, goal setting, and social support [11,12]. In addition, the mode of delivery of the intervention is equally important, as it can influence its acceptance, dissemination, and long-term use [8,13].

The use of technology in the delivery of behavior change interventions has potential in promoting their success and diffusion. Notably, mobile health (mHealth) interventions, involving mobile apps and wearable devices, can reach individuals continuously, enabling the self-monitoring of health and physical activity data [14] and the tailoring of intervention components in real time [15]. In addition, Web-based social networks seem to hold great promise, as they can help address social processes related to behavior change such as social support and social comparison [16,17]. Given their potential, interventions combining mHealth technologies and Web-based social networks might be particularly effective in promoting physical activity.

To date, a few qualitative studies have sought users' attitudes and views on the use of mHealth technologies and Web-based social networks for physical activity promotion [18-22], with most focusing on just one of these technologies. This limits the ability of researchers and developers to assess whether these 2 technologies can work in synergy. In addition, it remains unclear which behavior change components are most effective and which are considered more engaging by consumers [23]. The aim of this study was to explore individuals' perspectives before and after using a mobile social networking app for physical activity promotion. Specifically, we were interested in exploring potential barriers and facilitators to engagement with the intervention, as well as the behavior change techniques and delivery features considered important by users to promote physical activity. This research will help guide the future development of interventions and public health initiatives that could be more effective in influencing physical activity.

Methods

Study Overview

This study is part of a larger mixed-methods feasibility study on the use of a social networking mobile app to promote physical activity and weight management [24]. Given the importance of physical activity and its impact on weight management [1-3], this paper focused specifically on factors influencing physical activity. This study adheres to the COnsolidated criteria for REporting Qualitative research checklist for reporting qualitative research (Multimedia Appendix 1) [25]. This study protocol was approved by the Macquarie University's Human Research Ethics Committee for Medical Sciences (reference number: 5201600716). The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files.

Study Setting and Participants

This study was conducted at Macquarie University (Sydney, Australia). We recruited 55 participants, mostly university staff and students, using purposive sampling techniques through several channels, including posters around campus, website information, social media, and an email newsletter. Eligible participants were healthy adults with sufficient English to understand and participate in the study; aged between 19 and 35 years; who planned to be living in Sydney for the duration of the study; and owned a mobile phone (iOS or Android) with internet access. The exclusion criteria included pregnancy; body mass index (BMI) <17; prior history of eating disorders; or having diabetes or other comorbid conditions that could impact the study participation (eg, severe mental illness and end-stage disease).

For a 6-month period, participants were asked to use an intervention bundle (detailed below). Interviews were conducted pre- and postintervention, with the aim of assessing participants' perspectives on the use of social networking and mHealth interventions to promote physical activity. Of 55 initial participants, 45 returned for the final interviews.

Intervention Description

The intervention bundle was composed of a mobile social networking app (named "fit.healthy.me"), a fitness tracker (Fitbit Flex 2), and short message service text messages and emails [24]. The mobile app "fit.healthy.me" consisted of several features—"My measures," "My team," "Social forum," and "Private message"—which directly supported different behavior changes techniques (self-monitoring, social support, and social comparison). Specifically, "My measures" provided a summary of the number of steps, weight, and BMI. "My team" was a platform for participants to visualize and compare their steps with others. "Social forum" and "Private message" were designed for individuals to network with other users and provide and receive social support.

To enable the automation of self-monitoring, the app was integrated with the Fitbit Flex 2 fitness tracker, through the Fitbit Application Programming Interface. Reminders to wear the trackers and check the app were sent to participants every 2 weeks in the form of short message service text messages and emails. Table 1 provides a detailed description of the modes of delivery and features of the intervention, and Multimedia Appendix 2 shows the screenshots of the "fit.healthy.me" app.

Table 1. Intervention description.

Modes of delivery	Features	Behavior change techniques[a]
fit.healthy.me app	My measures	Self-monitoring of behavior (ie, physical activity)
	My team	Social comparison
	Social forum	Social support Social comparison
	Private message	Social support Social comparison
	My journey	Instruction on how to perform the behavior
Fitbit Flex 2	Fitness tracker	Self-monitoring of behavior (ie, physical activity)
Texts and emails	Reminders	Prompts or cues

[a]Classified according to the behavior change technique taxonomy developed by Michie et al [26].

Interview Procedure

Prior to study commencement, an interview guide (Multimedia Appendix 3) was developed and pilot-tested. Participants were invited to attend the initial study session at the research center, where they received information about the purpose of the study, signed the consent form, and filled in a questionnaire about their demographic characteristics and smartphone usage (eg, the type of smartphone used and hours per day spent using the smartphone).

In the preintervention session, 55 participants attended a brief individual interview (10-15 minutes) in which they were asked about perceived facilitators and barriers to physical activity and their views on the potential advantages and disadvantages of the mobile app and wireless devices (fitness tracker and scale). The content of the preintervention interviews was summarized and used as prompts for discussion in the postintervention sessions.

In the postintervention session, we conducted 32 individual interviews and 5 focus groups with 13 participants (20-45 minutes); data saturation was reached. While the interviews allowed us to understand individual perspectives, the focus groups enabled us to explore group differences and similarities [27,28].

At the postintervention sessions, participants talked about their experiences regarding the use of the intervention and provided suggestions on the devices and the intervention. Furthermore, semistructured interviews were conducted by 2 researchers with expertise in qualitative methods. Field notes were taken throughout the interviews.

Data Management and Analysis

With participants' consent, the interviews were recorded and transcribed verbatim, and transcripts were analyzed in Nvivo 11 (QRS International Pty Ltd., Melbourne, Australia). The data were analyzed using thematic analysis techniques [29]. Specifically, the transcripts were explored using the inductive analysis to identify themes and patterns [29]. First, we open-coded the transcripts to identify all important aspects related to the research questions. Subsequently, by scrutinizing and comparing different data and codes (ie, constant comparison), we pinpointed concepts that seemed to cluster together [30]. Informed by engagement with the literature, we identified the similarities, differences, and general patterns in the open codes, to fill in underdeveloped categories, narrow excess ones, and organize them into major themes [30,31].

Results

Sample Characteristics

Table 2 summarizes participants' demographic characteristics. At baseline, 51% (28/55) participants were females; the mean age was 23.6 years. On average, participants spent 5.6 hours daily using smartphones, and 89% (49/55) participants stated that they frequently used social media. Of all, 76% (42/55) participants were university students.

Summary of Results

We found the following 3 types of factors emerging from the data as influencing user engagement with the intervention and physical activity levels: individual, social, and technological. At the individual level, participants mentioned that goal setting, self-monitoring, and feedback were important for their physical activity. At the social level, social comparison and the connection with other users in terms of familiarity and similarity were considered motivating. Finally, at the technological level, automation and personalization were considered to be facilitators, while technological limitations were observed as reducing user engagement. The following sections discuss each of these themes in detail, with illustrative quotations (Textboxes 1-3).

Table 2. Baseline sample characteristics (N=55).

Characteristics	Value
Age, mean (SD)	23.6 (4.6)
Female gender, n (%)	28 (51)
Weight, mean (SD)	78.1 (22.3)
BMI[a] (kg/m^2), mean (SD)	26.5 (6.8)
BMI categories[b], n (%)	
17-18.49	3 (6)
18.5-24.9	24 (44)
25-29.9	15 (27)
≥30	13 (24)
Steps/day, mean (SD)	9937 (3527)
Marital status, n (%)	
Single	27 (49)
In a relationship	22 (40)
Married or de facto	6 (11)
Daily smartphone use (hours), mean (SD)	5.6 (3.4)
Most used apps[c], n (%)	
Social media	49 (89)
Fitness apps	6 (10)
Occupation, n (%)	
Student	42 (76)
Other	13 (24)
Smartphone, n (%)	
iPhone	36 (66)
Samsung	6 (11)
Other	13 (24)

[a]BMI: body mass index.

[b]According to the World Health Organization, a BMI of <18.5 is classified as underweight, 18.5-24.9 as normal, 25-29.9 as preobese, and ≥30 as obese [32].

[c]Most used apps—options are not mutually exclusive.

Individual-Level Factors Influencing Physical Activity

Self-Monitoring

Self-monitoring was deemed important by many users, as it increased their awareness of activity levels and performance, as well as enabled them to review their progress over time and better plan their exercise (Textbox 1, quotes 1 and 2). Some users indicated that even though self-monitoring was important, knowing the daily number of steps was not sufficient, as they were doing other types of exercise. Thus, they would prefer to measure parameters that were relevant to the type of activity they did (Textbox 1, quotes 3 and 4).

Other than physical activity, users also expressed the desire to monitor a wide range of health-related information (eg, sleep). By having multiple types of information about themselves, users felt they could get an overall view of their daily patterns, and how external factors (eg, family, jobs, and study) affected their health and well-being (Textbox 1, quote 5).

Goal Setting

Many participants expressed that they benefited from goal setting. They believed that setting a goal (eg, 10,000 steps daily) kept them accountable for their physical activity performance and motivated them to reach that goal. Participants indicated that goal setting and self-monitoring complemented each other because, without self-monitoring, they would have no way of knowing whether their goals had been achieved (Textbox 1, quote 6). In addition, many participants expressed the desire to be able to personalize their goals to fit with their ability and daily routines, rather than having a standard goal (Textbox 1, quote 7).

Feedback on Behavior

For many users, the feedback on progress toward goals was particularly encouraging; knowing that they were close to reaching their goals would motivate users to do more physical activity, while being notified of goal achievement gave them positive emotions (Textbox 1, quotes 8 and 9). Nevertheless, some participants mentioned that knowing they had not achieved their goals also brought on some negative feelings such as disappointment or guilt (Textbox 1, quote 10).

Social-Level Factors Influencing Physical Activity

Social Comparison

Participants mentioned that comparing themselves with other users encouraged them to be more engaged with the intervention, as well as to be more physically active (Textbox 2, quotes 1 and 2). One interesting aspect was that comparisons with higher, lower, or similar standards of physical activity (upward, downward, and lateral comparisons in accordance to [33]) had different effects on performance, according to participants. Most users said that they preferred to compare themselves against higher performers because that motivated them to try to learn their strategies and be more physically active, to beat the top level (Textbox 2, quote 3). Other users mentioned that they would like to compare themselves to both similar and higher standards (Textbox 2, quotes 4 and 5). On the other hand, some participants mentioned that comparison to higher standards could be rather demotivating and confronting, especially when they failed to achieve as many steps as others. Instead, those users preferred comparing themselves with lower standards, which gave them a sense of confidence and assurance that they were on the right track (Textbox 2, quotes 6 and 7).

Familiarity With Other Users

For many participants, social comparison and providing social support did not hold much meaning if they did not personally know other users. Many suggested that they were more likely to be engaged if they were "familiar" with other users (eg, if other users were their real-life social connections; Textbox 2, quotes 8 and 9). On the other hand, some participants mentioned that they did not necessarily need to know other users in real life; however, they needed to have some information about other users such as their lifestyle, fitness goals, or the types of activity they did, which could form the basis for social comparison (Textbox 2, quotes 10 and 11).

Similarity With Other Users (Homophily)

Other users did not stress the importance of "familiarity"; instead, they described a preference to share data within a social network of people who shared *similar* attributes or goals to them (a phenomenon known as "homophily" [34]). Particularly, some participants preferred to connect with users who had similar BMI or were doing the same type of physical activities (Textbox 2, quotes 12 and 13). In addition, a lot of participants emphasized the importance of having a similar goal, as it might facilitate more meaningful comparison and discussion on PA strategies (Textbox 2, quotes 14 and 15).

Textbox 1. Illustrative quotations for individual-level factors that influence participant engagement and physical activity.

Self-monitoring of behavior

- Quote 1: The important part for me is [keeping track] – I know I'm going beyond the average, like the normal number of steps for a person […] - it makes me more motivated. (Female, 24)
- Quote 2: I could use the data, so I know how [many] steps for one run, or how long I take for one run. It helps me to evaluate how [many] runs I could actually do, or what should be my targets for next day. (Male, 24)
- Quote 3: I climb now […] I'm actually looking for a watch or something that can measure altitude, it will be more interesting because I'd get to see how far I've climbed. (Female, 20)
- Quote 4: [I do] martial arts, so it's not so much running and movement. I want to have heartrate, it'd probably be a bit more useful. (Male, 20)
- Quote 5: I realized because of work pressure, in fact, I'm doing two jobs right now […] my average sleep has gone down. (Male, 27)

Goal setting

- Quote 6: There was a goal to reach every day. It kept me motivated […]. I would feel bad if I'm not wearing the [Fitbit]. It was like an additional limb in my body sort of thing." (Male, 27)
- Quote 7: I want to set my own goals each day […]. Some days I'm more active than other days. On those days, I'll automatically reach 10,000 steps in […] one session alone. But if I changed [the goal] to 20,000 steps then […] it would not really [be] achievable on the days that I don't do that much physical activity. If you could tailor the steps per day, then the motivation would be continuous. Because the motivation only works if I get close up to the end. (Female, 20)

Feedback on behavior

- Quote 8: Because I work long hours, I would reach 10,000 steps at like 10am. It always made me feel good when it vibrated and all the colors everywhere. I was like, yes! (Female, 20)
- Quote 9: When I […] got 80% of my goal, [I would just] go aimlessly for a walk. So that was getting me to walk more. Solely because I was on 80% and I wanted that 100%. (Female, 20)
- Quote 10: It sorts of guilt-tripped a bit. When I'd see it and I was like oh, I'd only done so many steps today. (Female, 19)

Participation From Other Users

Participation from other users was important for people to engage with the social network component of the intervention. Many users described attrition as a "domino effect"—once a certain number of people stopped using the app or the wearable tracker, other users subsequently felt less motivated to use the technology (Textbox 2, quotes 16 and 17).

Textbox 2. Illustrative quotations for social-level factors that influence participant engagement and physical activity.

Social comparison

- Quote 1: It gives me positive reinforcement at the same time because…I'm at the top chart of the steps. It kind of motivates me to stay on that level of rank and in general it motivates me because I can see if I'm doing well or not. I compare myself with the others. (Male, 24)
- Quote 2: I find competition helps me to regularly exercise often by going for runs with friends or family or competing in team sports.…Other people can see [your effort] and keep you accountable to your fitness goals. There's also that element of showing off…and also being able to see how other people exercise and then try to match them. (Male, 23)
- Quote 3: I probably look up more.…A lot of my days, I get up to 17,000 steps. So, I don't look down, I'd look up and be like, "Oh, why are those people getting 21,000 steps? I need to get 21,000 steps." (Female, 24)
- Quote 4: I would obviously want my comparison to be done with somebody who is exactly like me, or similar in certain ways. It gives me some kind of happiness that I'm achieving my goals in comparison to this person. It's like a competition. It's like scoring 87 and the other person is scoring 84.…Then I would also want to know the person who has got a 96 and why did he get a 96?…If you want to achieve 100, you want to know where you went wrong and what did you do right. But I don't want to compare with a person who got a 40. (Male, 27)
- Quote 5: I was probably competing to the person closest in terms of kilometers that we were doing. It was interesting to see what they were doing and how they progress…. I tried to beat them every day. (Male, 21)
- Quote 6: Being compared to other people was a bit shocking—I was [at] the end of the group, so it was a bit demotivating. (Female, 20)
- Quote 7: If I'm having more steps than others, I feel motivated, and know that at least I keep myself healthy. (Female, 24)

Familiarity with other users

- Quote 8: It's like, I don't really know anyone [in the study] and then…the fad of comparing yourself against people wears off; I did try and use it a little bit more, but it was just like because you don't know anyone, you forget about it.…If it was in a group of my friends, we probably would've been checking it weekly. (Female, 24)
- Quote 9: I guess not knowing what they do…—whether they worked or whether they were students— not knowing that, it's a bit hard to…compare because there's all these variables. Also, because I really didn't know them, I didn't feel obliged to try to motivate them at all in any way. I guess with friends—and if I got to know them at all— yeah, I might have done that. (Male, 30)
- Quote 10: [I'd like to see] more information about the kind of fitness people are doing. For example, someone has done 20,000 steps in a day, which is a huge amount, then give me a basic idea of what that person has done to get to that goal. (Female, 19)
- Quote 11: If everybody [had a] profile, maybe it [would be easier] to make friends. At the beginning I thought "Maybe I can [make a] friend and we can train together to lose some weight." (Female, 34)

Similarity with other users

- Quote 12: I think it would help if you had people…with a similar body type doing similar things that would suit you more. (Female, 23)
- Quote 13: I like that you could go through and track people who were similar to you…. I went and found people with similar BMI. I'm happy to track myself against similar people and see how many steps [they've done]. (Female, 24)
- Quote 14: Everyone's goal might be different. So, you need to group people with similar goals together. …I would want to compare myself to somebody who [has similar goals] and is using it on a daily basis like me." (Male, 27)
- Quote 15: Having a goal section where people say whether they want to gain or lose weight would be good. Then all people who want to lose weight can get together and talk about it. (Male, 20)

Participation from other users

- Quote 16: It was a bit like a domino effect, so after about two months you could see that 20 to 30 per cent had zero [steps]. It felt like people weren't using the app, so there was no reason for me to use it as well. (Male, 22)
- Quote 17: There's no number of steps [from some people] sometimes. It can be a little demotivating when you see a lot of zeros…It's like are they taking this seriously? (Male, 24)

Technology-Level Factors Influencing Physical Activity

Technological Facilitators of Engagement and Behavior Change

Automation

Many participants found that using the wireless tracker and scale in combination with a mobile app offered many advantages. Specifically, wireless devices provided an automatic way for users to collect and self-monitor personal measurements, and their integration with the mobile app provided a user interface platform for participants to visualize those data and to review progress (Textbox 3, quotes 1 and 2).

Personalization

Many users mentioned that having personalized information and services would also support long-term engagement, as they could offer the advantage of providing relevant information tailored to each specific user, thus eliminating the cognitive burden of dealing with information overload. Many users described that personalization should go beyond the content generated by the system and extend to the provision of relevant services (eg, suggestion of exercise routines; Textbox 3, quotes 3-5).

Technological Barriers to Continued Usage

Additional Workload

As time went on, many users described the feeling that the novelty of the technology had worn off, and they started to think of it as a chore. Even apparently simple tasks like charging the devices were seen by participants as an extra burden in their already busy daily routines (Textbox 3, quotes 6 and 7).

Technical Problems and User Experience

Technical problems were often described as a common cause for attrition (Textbox 3, quote 8). In addition, user experience factors, such as the design aspects of the interface and its usability, were reported as important aspects of engagement and continued use (Textbox 3, quotes 9 and 10).

Textbox 3. Illustrative quotations for technological-level factors that influence participant engagement and physical activity.

Technological facilitators of engagement and behavior change

- Quote 1: I enjoyed how [the wearable tracker] linked with the app, and then on the app you could track how many steps you [did]. […] With the scale as well, the scale was able to track my weight and then it gives you a trend line to show how you're doing, so I enjoyed that as well. Having the combination of the tracker, the scale and the app was really good. (Male, 22)
- Quote 2: I like the [Fitbit] app. It integrates so well, so you wear your [tracker] and then [the app] tells you [how many] exercises you've done in a week, your steps, sleep. (Female, 31)
- Quote 3: [Having health information] would be good, but it has to be personalized or customized to me, (…) my body type, […] not like a general advice like [what is] BMI etc. […] A lot of people can read about general information; but if it's personalized to you or customized to your needs, it's going to be more interesting and more reliable […]. (Male, 24)
- Quote 4: I liked that at the end [of a fitness video], you can put a smiley face on how difficult it was. Based on my reaction, I want the app to give me recommendations on what types of exercises I should do. So, it was tailored to me, according on my reaction. (Female, 20)
- Quote 5:
  - Male: Whether to have one or multiple buddies, the choice depends on what works for the person. Maybe you can personalize it in some way. Maybe you can elect I want only one partner, or I want to be put in a group. (Male, 20)
  - Female: It is like gym training session, you can have private sessions, you can have small group sessions, or you can have a class session and you choose which one is best for you. The same with the app and your buddy. (Female, 20)

Technological barriers to continued usage

- Quote 6: The charge lasted three days, and because I had such a busy schedule, charging it again [was] such a big chore. So, it would then just sit for another week and I'd get a [reminder] email and then I would plug it in […]. I was doing so many things, so remembering to charge it became a challenge. (Male, 33)
- Quote 7: After a first couple of months, it started to feel more like a chore to do. I got into the thinking "I had to [check the app] everyday" as opposed to "I want to do this every day to keep track of my weight". Then university started, and things started getting busy. (Male, 22)
- Quote 8: The battery was discharging very quickly. In the morning it was telling me that I had achieved my goals when I just started the day. (Female, 20)
- Quote 9: I liked the social comparison feature in fit.healthy.me, but it's hidden in several menus. I liked the Fitbit app better—the design is certainly more elegant. (Female, 26)
- Quote 10: I checked the Fitbit app more than the fit.healthy.me app. I think the reason was because the Fitbit app was much sleeker, looks nicer and more inviting and easier to use. (Female, 20)

Discussion

Principal Findings

This study explored users' perspectives regarding facilitators and barriers in using mobile social networking interventions to promote physical activity. The following 3 categories of influencing factors emerged: individual, social, and technological. At the individual level, behavior change techniques, such as goal setting, self-monitoring, and feedback, were suggested as important for user engagement in physical activity. At the social level, social comparison, familiarity, and similarity with other users were mentioned as motivating aspects. Finally, automation and personalization were highlighted as technological facilitators, enhancing the delivery of both individual and social aspects of the intervention. However, some technological limitations were also found to be barriers to user engagement.

Comparison With Previous Literature

Our findings suggest that the success of a behavior change depends on a range of factors, including both individual and social aspects. These findings are in line with other behavior change theories, namely the social cognitive theory [4], and the Capability Opportunity Motivation—Behavior model [5]. Both theories suggest that even though several behavioral factors (eg, self-regulation [35], capability, and motivation [5]) are largely dependent on individuals, external factors (eg, peer modeling [4] and environmental structure [5]) can arise from the physical or social environments to prompt behavior. Hence, it seems sensible to integrate both individual and social aspects of behavior change in physical activity interventions to increase their long-term success.

In line with our results, behavioral informatics interventions (eg, a mobile social networking app, connected with a fitness tracker) can facilitate the delivery of both individual and social aspects in physical activity interventions [8]. Specifically, fitness trackers can automate the self-monitoring of behavior and connect to mobile apps with social features, allowing users to not only view their progress but also continuously benefit from social support [23,36]. To date, one qualitative study has examined how wearable trackers, mobile apps, and Web-based social networks may interact, finding that social support from Web-based networks can be effective in increasing users' adherence and engagement with the wearable trackers [37]. However, this study had a couple of limitations—it included a small number of users, as well as nonusers of wearable trackers; and it examined Web-based social networks as a stand-alone feature, not integrated with the trackers. In contrast, our study provided participants with an integrated intervention, including mHealth and social networking components, which allowed us to explore the informed perspectives of participants who used these technologies for 6 months.

Individual-Level Behavior Change Techniques

Our users indicated that goal setting, self-monitoring of behavior, and feedback on behavior could encourage them to engage in physical activity, which is in line with previous qualitative studies [18,19]. Indeed, these 3 self-regulatory techniques have demonstrated the effectiveness in physical activity interventions [11] and may work in synergy—to maximize the effects of goal setting, people may need to self-monitor and receive feedback, which allows them to see their progress in relation to their goals and change their strategies if necessary [38].

In addition, previous research has suggested the need to examine which type of goal is best for motivating individuals to be more active and how technologies can best support monitoring those goals and providing feedback. The literature seems to suggest that small goals (described as "graded tasks" in the Coventry, Aberdeen, and London—Refined taxonomy [10]) are more effective for long-term engagement compared with larger and harder to achieve goals [39]. For example, Fitbit provides users with small goals of taking 250 steps per hour, which then facilitates the achievement of the daily goal of 10,000 steps [23]. It is worth noting the importance of real-time self-monitoring and consistent feedback for the success of this "small goals" approach [23], underlining implications for the design of mobile apps and wearable trackers.

Social Networks and Social Features

This study emphasized the role of social comparison, familiarity, and similarity with other users in a social networking intervention. First, our participants revealed different preferences regarding social comparison. This finding is in line with previous research, where it has been demonstrated that individual preferences might depend on their tendency to make upward or downward comparisons [40]. Specifically, previous studies have illustrated that some people seek social comparison to self-improve [33], and thus, upward comparison may reinforce positive fitness behavior by making it seem normative or even rewarding [41,42]. For others, instead of seeking feedback about themselves, they want to create and maintain a positive self-image, and thus, prefer to make a downward comparison [33,42]. Taken as a whole, this finding suggests that a one-size-fits-all approach to social comparison is unlikely to suit all users, and thus, social comparison needs to be tailored to each individual.

Second, *familiarity* and *similarity* were found to be important factors in a social networking intervention for physical activity. The importance of familiarity seems to be in line with previous literature, where researchers have demonstrated that existing social networks can greatly influence individual health behaviors [43,44], leveraging social support and potentially increasing the intervention effectiveness [17,40,45-47]. Research has shown that strategies involving new networks might not be as effective as ones capitalizing on existing connections [46,47], which suggests that fitness technology may be most effective when groups of people who know one another have access to the same device or app [23]. Thus, allowing study participants to invite friends and family to join an app may increase the real-world effectiveness of these interventions [40], despite potential problems of contamination.

Furthermore, this study showed that similarity is important for motivation and engagement, highlighting the role of homophily (ie, the tendency of people to bond with alike individuals) [34]. Notably, previous research has indicated that social networks

structured on the basis of homophily lead to higher adoption of healthy behaviors [48]. Moreover, it has been suggested that when people with similar interests interact to achieve a shared goal, they can provide each other with support and companionship in the activity, and thus, reduce the perceived costs of adopting a new exercise routine [46,49]. Taken together, these findings highlight the benefits of leveraging homophily to foster collective efficacy and improve physical activity.

Technology As a Platform to Bring Together Individual and Social Levels

Through automation and personalization, multiple modes and features of technology can work synergistically to deliver a physical activity intervention with both individual and social factors [37,50,51]. Thus, the integration of multiple mHealth technologies can automate several aspects of health management, reducing the burden on users. Furthermore, many users suggested the importance of personalized features within the intervention. Indeed, a one-size-fits-all approach is unlikely to satisfy many needs and wants of users [52], which emphasizes the need to consider individual lifestyles and preferences when designing interventions.

Strengths and Limitations

This study has several strengths. We interviewed users after 6 months of experiencing the intervention, ensuring that our sample had an informed perspective. The combination of individual interviews and focus groups enabled us to capture both individual perspectives and social dynamics in a group setting, which are essential aspects to understand in a social networking intervention. The findings of this paper must be interpreted in light of some limitations. First, study recruitment was limited to a university setting with a young age group. Though the main purpose of qualitative studies is not to make generalizable claims [53], future research with a diverse sample could explore other contextual factors related to behavioral informatics interventions (eg, an older age group might encounter different barriers and facilitators of a mobile social networking app). Second, as this was part of a feasibility study, the technology used was at a prototype stage and not yet extensively tested. Finally, despite our engagement efforts, we were not able to interview participants who dropped out of the study—they might have different perspectives on the facilitators and barriers of the intervention.

Implications for Future Research

This study highlights several important implications, including suggestions on the intervention design and new research avenues. Interventions for physical activity promotion should consider offering goal setting, self-monitoring, and feedback as a bundle, as these techniques have been shown to be both effective and acceptable to end users. Consequently, the design of mobile apps and wearable trackers need to effectively assist with real-time self-monitoring and provide consistent feedback to enable the achievement of goals [23]. In addition, the potential of social behavior change techniques (eg, social comparison) should be further explored, and aspects of leveraging existing social ties and homophily could be considered in constructing a social network intervention for physical activity. Questions remain about the cost-effectiveness of wearable trackers and mobile apps as a public health initiative, opening up new possibilities for future health economics research and public health programs [23,54].

Furthermore, this study highlights the importance of *personalization*. By identifying users' behavioral patterns and preferences, researchers can design and deliver interventions at the right time, using the right channel and tone, and the most relevant content or services [55,56]. Future studies should use innovative study designs to determine which intervention components are effective, what is the optimal sequence for delivering these components, and which tailoring variables should be used [23,57].

Conclusions

This study provides insights into the individual, social, and technological factors that influence user engagement with a mobile social networking app for physical activity promotion. Our findings reveal that self-regulatory behavior change techniques seem to be a necessary element in these interventions, and that aspects related to social comparison, existing social ties, and homophily should be considered in the development of the social network component. Future research should adopt innovative research designs to evaluate the effectiveness of these different components, as well as investigate the delivery of personalized interventions.

Acknowledgments

The authors would like to thank Ms Paige Martin for her contribution to the mobile app design and development and data collection. HLT is supported by the International Macquarie University Research Training Pathway Master of Research (iMQRTPMRES) Scholarship (2016302). This research is supported by a grant received from the National Health and Medical Research Council Centre of Research Excellence in Informatics and E-Health (1032664). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

HLT, EC, and LL conceptualized the study. HLT developed and pilot-tested the interview guide, conducted the interviews and focus groups, performed data analysis, and wrote the first draft of the manuscript. LL pilot-tested the interview guide, conducted some data collection and analysis, and provided guidance on data analysis and critical feedback on the manuscript. EC critically revised the manuscript.

Conflicts of Interest

EC could benefit from commercialization of fit.healthy.me.

Multimedia Appendix 1

COnsolidated criteria for REporting Qualitative research checklist.

[PDF File (Adobe PDF File), 132KB - jmir_v20i12e11439_app1.pdf]

Multimedia Appendix 2

Screenshots of the fit.healthy.me mobile app.

[PDF File (Adobe PDF File), 177KB - jmir_v20i12e11439_app2.pdf]

Multimedia Appendix 3

Interview guides.

[PDF File (Adobe PDF File), 66KB - jmir_v20i12e11439_app3.pdf]

References

1. World Health Organization. World Health Organization. 2010. Global recommendations on physical activity for health URL: https://www.who.int/dietphysicalactivity/factsheet_recommendations/en/ [accessed 2018-10-15] [WebCite Cache ID 73CLuQkUQ]
2. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. The Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
3. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, et al. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
4. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, New Jersey: Prentice Hall; 1985.
5. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
6. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696 [FREE Full text] [Medline: 10987780]
7. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci 2009 Jul 16;4:40 [FREE Full text] [doi: 10.1186/1748-5908-4-40] [Medline: 19607700]
8. Laranjo L, Lau A, Coiera E. Design and Implementation of Behavioral Informatics Interventions. In: Patel VL, Arocha JF, Ancker JS, editors. Cognitive Informatics in Health and Biomedicine: Understanding and Modeling Health Behaviors. New York: Springer International Publishing; 2017:13-42.
9. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
10. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
11. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
12. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
13. Davidson KW, Goldstein M, Kaplan RM, Kaufmann PG, Knatterud GL, Orleans CT, et al. Evidence-based behavioral medicine: what is it and how do we achieve it? Ann Behav Med 2003 Dec;26(3):161-171. [Medline: 14644692]
14. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015;12(1):159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]

15. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
16. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256. [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
17. Coiera E. Social networks, social media, and social diseases. BMJ 2013;346:f3007. [Medline: 23697672]
18. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Dec 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
19. 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] [Medline: 23598614]
20. Fritz T, Huang E, Murphy G, Zimmermann T. Persuasive technology in the real world: A study of long-term use of activity sensing devices for fitness. : ACM; 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 26 - May 01, 2014; Toronto.
21. Shih P, Han K, Poole E, Rosson M, Carroll J. Use and adoption challenges of wearable activity trackers. 2015 Presented at: iConference; March 24-27, 2015; California.
22. Van Kessel G, Kavanagh M, Maher C. A Qualitative Study to Examine Feasibility and Design of an Online Social Networking Intervention to Increase Physical Activity in Teenage Girls. PLoS One 2016;11(3):e0150817 [FREE Full text] [doi: 10.1371/journal.pone.0150817] [Medline: 26934191]
23. Sullivan AN, Lachman ME. Behavior Change with Fitness Technology in Sedentary Adults: A Review of the Evidence for Increasing Physical Activity. Front Public Health 2016;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
24. Laranjo L, Lau AYS, Martin P, Tong HL, Coiera E. Use of a mobile social networking intervention for weight management: a mixed-methods study protocol. BMJ Open 2017 Jul 12;7(7):e016665 [FREE Full text] [doi: 10.1136/bmjopen-2017-016665] [Medline: 28706104]
25. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
26. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
27. Kitzinger J. The methodology of focus groups: the importance of interaction between research participants. Sociology of Health & Illness 1994;16(1):103-121 [FREE Full text]
28. Minichiello V, Aroni R, Timewell E, Alexander L. In-depth Interviewing: Researching People. London: Routledge; 1990.
29. Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. In: Willig C, editor. The Sage Handbook of Qualitative Research in Psychology. London: Sage Publication; 2013:17-36.
30. Bowen G. Naturalistic inquiry and the saturation concept: a research note. Qualitative research 2008;8(1):137-152.
31. Saldaña J. The coding manual for qualitative researchers. London: Sage Publications; 2015.
32. World Health Organization. World Health Organization. Body mass index - BMI URL: http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi [accessed 2018-10-15] [WebCite Cache ID 73CORIBSj]
33. Corcoran K, Crusius J, Mussweiler T. Social comparison: Motives, standards,mechanisms. In: Chadee D, editor. Theories in Social Psychology. Hoboken, New Jersey: John Wiley & Sons; 2011:119-139.
34. McPherson M, Smith-Lovin L, Cook J. Birds of a feather: Homophily in social networks. Annual Review of Sociology 2001;27(1):415-444.
35. Bandura A. Social cognitive theory of self-regulation. Organizational Behavior and Human Decision Processes 1991;50(2):248-287.
36. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
37. Chang RC, Lu H, Yang P, Luarn P. Reciprocal Reinforcement Between Wearable Activity Trackers and Social Network Services in Influencing Physical Activity Behaviors. JMIR Mhealth Uhealth 2016 Jul 05;4(3):e84 [FREE Full text] [doi: 10.2196/mhealth.5637] [Medline: 27380798]
38. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [Medline: 12237980]
39. Fogg B. A behavior model for persuasive design. : ACM; 2009 Presented at: The 4th International Conference on Persuasive Technology; 26 - 29 April; Claremont, California p. 40.

40. Klein MCA, Manzoor A, Mollee JS. Active2Gether: A Personalized m-Health Intervention to Encourage Physical Activity. Sensors (Basel) 2017 Jun 19;17(6):--- [FREE Full text] [doi: 10.3390/s17061436] [Medline: 28629178]
41. Arigo D, Schumacher LM, Pinkasavage E, Butryn ML. Addressing barriers to physical activity among women: A feasibility study using social networking-enabled technology. Digit Health 2015;1:2055207615583564 [FREE Full text] [doi: 10.1177/2055207615583564] [Medline: 29942539]
42. Burke TJ, Rains SA. The Paradoxical Outcomes of Observing Others' Exercise Behavior on Social Network Sites: Friends' Exercise Posts, Exercise Attitudes, and Weight Concern. Health Commun 2018 Jan 24:1-9. [doi: 10.1080/10410236.2018.1428404] [Medline: 29364740]
43. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007 Jul 26;357(4):370-379. [doi: 10.1056/NEJMsa066082] [Medline: 17652652]
44. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008 May 22;358(21):2249-2258 [FREE Full text] [doi: 10.1056/NEJMsa0706154] [Medline: 18499567]
45. Caperchione CM, Kolt GS, Savage TN, Rosenkranz RR, Maeder AJ, Vandelanotte C, et al. WALK 2.0: examining the effectiveness of Web 2.0 features to increase physical activity in a 'real world' setting: an ecological trial. BMJ Open 2014 Oct 10;4(10):e006374 [FREE Full text] [doi: 10.1136/bmjopen-2014-006374] [Medline: 25304191]
46. Cavallo DN, Tate DF, Ward DS, DeVellis RF, Thayer LM, Ammerman AS. Social support for physical activity-role of Facebook with and without structured intervention. Transl Behav Med 2014 Dec;4(4):346-354 [FREE Full text] [doi: 10.1007/s13142-014-0269-9] [Medline: 25584083]
47. Rovniak LS, Kong L, Hovell MF, Ding D, Sallis JF, Ray CA, et al. Engineering Online and In-Person Social Networks for Physical Activity: A Randomized Trial. Ann Behav Med 2016 Dec;50(6):885-897 [FREE Full text] [doi: 10.1007/s12160-016-9814-8] [Medline: 27405724]
48. Centola D. An experimental study of homophily in the adoption of health behavior. Science 2011 Dec 2;334(6060):1269-1272 [FREE Full text] [doi: 10.1126/science.1207055] [Medline: 22144624]
49. Uchino BN. Social support and physical health: Understanding the health consequences of relationships. New Haven, Connecticut: Yale University Press; 2004.
50. Rovniak LS, Sallis JF, Kraschnewski JL, Sciamanna CN, Kiser EJ, Ray CA, et al. Engineering online and in-person social networks to sustain physical activity: application of a conceptual model. BMC Public Health 2013 Aug 14;13:753 [FREE Full text] [doi: 10.1186/1471-2458-13-753] [Medline: 23945138]
51. Eapen ZJ, Peterson ED. Can Mobile Health Applications Facilitate Meaningful Behavior Change?: Time for Answers. JAMA 2015;314(12):1236-1237. [doi: 10.1001/jama.2015.11067] [Medline: 26393844]
52. Ohno-Machado L. Tailoring informatics interventions to patients and healthcare providers. J Am Med Inform Assoc 2015 Jul;22(4):747. [doi: 10.1093/jamia/ocv075] [Medline: 26116707]
53. Holloway I. Qualitative Research in Health Care. Berkshire: Open University Press; 2005.
54. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
55. Fadhil A. Cornell University Library. 2018. Towards Automatic and Personalised Mobile Health Interventions: An Interactive Machine Learning Perspective URL: https://arxiv.org/abs/1803.01842 [accessed 2018-10-15] [WebCite Cache ID 73CPboP5L]
56. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
57. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]

Abbreviations

BMI: body mass index
mHealth: mobile health

Edited by G Eysenbach; submitted 28.06.18; peer-reviewed by A Sullivan, A Henriksen; comments to author 29.08.18; revised version received 06.09.18; accepted 08.09.18; published 21.12.18.

Please cite as:
Tong HL, Coiera E, Laranjo L
Using a Mobile Social Networking App to Promote Physical Activity: A Qualitative Study of Users' Perspectives
J Med Internet Res 2018;20(12):e11439
URL: https://www.jmir.org/2018/12/e11439/
doi:10.2196/11439
PMID:30578201

Original Paper

Creating Engaging Health Promotion Campaigns on Social Media: Observations and Lessons From Fitbit and Garmin

Sarah Edney[1], BA; Svetlana Bogomolova[2], PhD; Jillian Ryan[3], PhD; Tim Olds[1], PhD; Ilea Sanders[1]; Carol Maher[1], PhD

[1]Alliance for Research in Exercise, Nutrition and Activity, University of South Australia, Adelaide, Australia

[2]Ehrenberg-Bass Institute for Marketing Science, University of South Australia, Adelaide, Australia

[3]Commonwealth Scientific and Industrial Research Organisation, Adelaide, Australia

Corresponding Author:
Sarah Edney, BA
Alliance for Research in Exercise, Nutrition and Activity
University of South Australia
GPO Box 2471
Adelaide, 5000
Australia
Phone: 61 +61402371792
Email: sarah.edney@mymail.unisa.edu.au

Abstract

Background: The popularity and reach of social media make it an ideal delivery platform for interventions targeting health behaviors, such as physical inactivity. Research has identified a dose-response relationship whereby greater engagement and exposure are positively associated with intervention effects, hence enhancing engagement will maximize the potential of these interventions.

Objective: This study examined the social media activity of successful commercial activity tracker brands to understand which creative elements (message content and design) they use in their communication to their audience, which social media platforms attract the most engagement, and which creative elements prompted the most engagement.

Methods: Posts (n=509) made by Fitbit and Garmin on Facebook, Twitter, and Instagram over a 3-month period were coded for the presence of creative elements. User engagement regarding the total number of likes, comments, or shares per post was recorded. Negative binomial regression analyses were used to identify creative elements associated with higher engagement.

Results: Engagement on Instagram was 30-200 times higher than on Facebook, or Twitter. Fitbit and Garmin tended to use different creative elements from one another. A higher engagement was achieved by posts featuring an image of the product, highlighting new product features and with themes of self-improvement (P<.01).

Conclusions: Findings suggest that Instagram may be a particularly promising platform for delivering engaging health messaging. Health messages which incorporate inspirational imagery and focus on a tangible product appear to achieve the highest engagement. Fitbit and Garmin employed difference creative elements, which is likely to reflect differences in their target markets. This underscores the importance of market segmentation in health messaging campaigns.

(J Med Internet Res 2018;20(12):e10911) doi:10.2196/10911

KEYWORDS

social media; engagement; physical activity

Introduction

Background

Social media platforms such as Facebook, Twitter, and Instagram are increasingly becoming a central part of daily life. Today, social media is used for everything from staying in touch with friends and family and accessing news media coverage to keeping up to date with brands and celebrities. Each platform has a large global user base. Facebook has more than two billion active users while Twitter and Instagram have 700 and 328 million each, respectively [1]. Most demographics are well represented on one or more platforms [2], and most aspects of these platforms are free to use. Facebook and Instagram appear to have a relatively equal balance of male and female users,

with 50% of Instagram users and 52% of Facebook users being female [3]. Twitter, on the other hand, appears to have more male users (64%) than females (36%). Regarding age demographics, Instagram is more popular among young people with 71% of users being 34 years old or younger, compared to 22% of Facebook users and 40% of Twitter users [3]. It is unsurprising then that these platforms are attracting attention as potential vehicles for the delivery of health promotion and behavior change interventions [4].

To date, most social media-based health behavior intervention and promotion efforts have been based on Facebook or Twitter [4], although they are starting to appear on Instagram [5]. The emerging body of research examining electronic health (eHealth) and mobile health (mHealth) interventions has identified a dose-response relationship, where greater engagement from participants and therefore exposure to an intervention is positively associated with both retention and positive intervention effects [6-8]. However, when subjected to empirical evaluation these interventions often report low rates of engagement which potentially limits their effectiveness [4,9]. The seminal hierarchical behavior change model awareness-interest-desire-action (AIDA), suggests consumer likes, reactions, and responses to commercial posts could serve as proxies for awareness and interest, which are precursors to intentions and eventual joining of health promotion programs and interventions [10]. The potential for such engagement to translate into positive intervention effects highlights a need to understand and enhance user engagement to maximize potential benefits.

The field of engagement science has expanded in recent years. Empirical work undertaken to understand, quantify, and make recommendations to enhance engagement with social media-based health intervention and promotion efforts varies in methodology. Several studies have used subgroup analyses of the intervention arms of randomized controlled trials intervening on weight loss [6] and physical activity [11], within Facebook settings. Content analysis approaches have been used to examine engagement with health-related social media content. For example, Guidry and colleagues [12] used this approach to understand how prominent public health organizations use Twitter and Instagram to disseminate information relating to infectious disease outbreaks, using the 2013-2014 Ebola outbreak as a case study. The authors examined the content of social media posts and the responses (ie, engagement) from users, concluding that Instagram holds particular promise as it elicited significantly higher rates of engagement from users when compared to Twitter.

The study by Rus and Cameron [13] recently examined how health topics are communicated and engaged with in online settings. They analyzed social media posts from diabetes-related Facebook groups to identify which post features elicited different forms of engagement (ie, "likes," "comments," or "shares") from users. Their content analysis approach categorized the post content and then used regression analyses to determine which message features were predictive of engagement. The authors were able to make recommendations for the design of future health-related social media content. Namely, that the use of imagery resulted in a post receiving more "likes" and "shares," and information relating to the consequences of having diabetes or positive self-identify resulted in more sharing of the post. Whereas posts containing negative affect or social support resulted in more comments.

The popularity of personal health and activity tracker devices (hereafter referred to as wearables) shows little signs of waning [14]. A plethora of wearable brands and models are currently available to consumers at various price points and with a range of features and styles. Two of the largest brands, Fitbit and Garmin, have well established social media profiles across Facebook, Twitter, and Instagram, all with large numbers of real users engaging with their branded content. These commercial wearables may act as a health behavior intervention, although they have less of a focus on determining which of their components are affecting positive behavior change, have far greater resources for software development and marketing, and a primary focus on product sales or revenue [15-17]. Wearables typically contain features that are found in health behavior interventions, for example, self-monitoring capabilities, which are well-established to have a potent influence on changing health behavior [18,19] are often a key feature. Wearables are typically marketed as devices to help users improve their health, and this means they potentially attract the same demographic of the user as health behavior change interventions. While the field of research is in its infancy, there is some emerging evidence to suggest that wearables may be efficacious in changing physical activity behaviors and that they may be able to act as a health intervention [20-25].

The popularity of social media platforms and wearable devices continues to grow while social media-based health interventions continue to report lower than intended rates of engagement [4,9]. Fitbit and Garmin are likely to use particular creative elements in their social media posts. Creative elements seek to translate the content of intended and targeted messages (eg, social media posts) into specific communication elements which include design and content features such as imagery, typeface (ie, within traditional print media), the content of the text, and any interactive features [26]. Examination of the social media activity of commercial wearables and the creative elements used in their posts may provide insights into the type of content that is appealing to current or prospective users of wearables. These insights can then be used to inform the development of appealing content that can be integrated into social media-based health interventions, thereby increasing the appeal and encouraging participants to engage with such interventions. In turn, increased engagement may boost adherence and positive intervention effects [6-8].

Research Aims

The aims of this study were:

- To examine the social media activity of commercially available wearable activity tracker brands to understand how they engaged social media users
- To determine which platform attracted the most engagement from users
- To examine which creative elements (message content and design elements) elicited higher engagement from users

Methods

This cross-disciplinary study is a content analysis of publicly-available social media posts made by successful commercial wearable activity trackers brands, Fitbit and Garmin, on their company Facebook, Twitter, and Instagram profiles. Approval for this study was granted by the University of South Australia's Human Research Ethics Committee (protocol no. 0000036513).

Sample Selection and Data Collection

Fitbit and Garmin were selected for inclusion based on their 2016 third-quarter worldwide market share figures of 23.0% and 5.7%, respectively and making them the first and third leading wearable activity tracker brands, globally [9]. The brand Xiaomi was second with 16.5% of market share for the same quarter [14]. However, the brand was not included in the sample here as their consumer base is heavily skewed toward a domestic market of Chinese users, the brand has a presence on Facebook and Twitter only and with many posts written in Chinese. Although originally included within the scope of the study, the social media profiles of Jawbone were inactive from early January 2017, and the company has since commenced liquidation [27] and was therefore removed from the analysis here.

All social media posts (n=509) made by the corporate social media accounts of Fitbit and Garmin on Facebook, Twitter, and Instagram over a three-month period (December 2016 to February 2017) were retrospectively collected at the end of March 2017. This three-month period is longer than that used in comparable studies which range from one week through to just under two months [13,28,29]. Posts made on these pages by social media users were not included. A screenshot of each social media post capturing the image, caption, and number of "likes," "comments," and "shares" was taken, and assigned an identification number.

Before statistical analysis, the engagement by social media users, regarding the total number of "likes," "comments," and "shares" per post was manually extracted from each screenshot and recorded in a Microsoft Excel file. Preliminary exploratory data collection determined that across all three platforms (ie, Facebook, Twitter, and Instagram) almost all user engagement (ie, "likes," "comments," and "shares") with each post occurred within seven days of posting. By 21 days, our exploratory work determined that the amount of engagement had increased by just 2.9% (range 0%-11%) per post, indicating that collecting all engagement data at a single time point would provide an accurate indication of engagement for each post. The number of followers of each brand on each platform at the time of data collection was also recorded.

Developing the Codebook

A standardized codebook of creative elements was developed to classify the content of the social media posts. Creative elements refer to message content and execution factors used to design communication with the greatest chance of eliciting the desired response from the target audience [30-32] in this case, user engagement. The creative elements that receive the most engagement can then be used in future social media-based health promotion and intervention efforts to aid participant engagement. Development of the codebook was guided by Stewart and Furse [31] who examined the influence of television advertising execution techniques on sales effectiveness. This codebook and study design have been replicated in full [30,32], partially [33,34], and partially for application to examining interactive television advertising [35]. An iterative process was used to condense the original 160-item codebook down to 34 items relevant to social media and health behaviors to accommodate the differences in advertising media. An initial review of the original 160 items identified 101 that were outside of the scope of the current study. For example, items related to "mechanical" advertisement devices such as the length of time until the product was shown were unlikely to be present in a sample comprised mainly of static imagery and text and were removed. Following the initial reduction, the research team piloted the codebook and identified a further 20 items that were unlikely to contribute to addressing the research aims. For example, "demonstration of the product in use" was removed as simply wearing the device in any scenario would constitute a demonstration of use. Following this, the item "nutrition and health" was expanded into 5 items that capture the overall theme of each of the posts (ie, whether the post primarily featured exercise or physical activity, incidental activity, weight loss, food or nutrition, or sleep information). The final version of the codebook with 34 creative elements can be found in Multimedia Appendix 1.

Applying the Codebook

The screenshots of each social media post were then cropped to display only the image and caption. These screenshots and the associated identification numbers were then entered into an Excel spreadsheet containing columns for each of the codebook categories. Each coder (SE, SB, JR, TO, IS, or CM) received an Excel file containing approximately 170 social media posts. This meant that each of the 509 social media posts was coded for the presence or absence of each of the 26 dichotomous and 8 categorical creative elements, by 2 independent coders. In addition to the Excel file, coders received a codebook containing a description and relevant example of each item, and training in how to administer the codebook. Total percentage pairwise agreement was acceptable, ranging from 82% to 91% for each pair of coders [36]. Disagreements between coders were assessed and resolved by a third independent coder.

Statistical Analyses

All statistical analyses were conducted using IBM SPSS version 23 [37]. Given that "likes" were the most common form of engagement in our sample and in similar studies (eg, see [13,38]) the authors made the decision to combine the number of "likes," "comments," and "shares" per post into "total engagement", to encompass all interactions with each post within a single metric. Chi-square tests of homogeneity were used to examine differences in use of creative elements between brands. Following a similar methodology to Rus and Cameron [13], the relationships between creative elements and engagement were examined using multivariate regression analyses. Poisson and negative binomial regression models were selected to account

for the nonnormal distribution of count data on the dependent variable. Overdispersion of the count data, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) indicated negative binomial regression as providing the best fit for the data. Creative elements with a significant univariate association (P<.25) were included in purposeful selection modeling [39], and those that had a P<.1 and changed the main effects by more than 10% were retained in the final models. Intercoder reliability for the 12 items retained for the final model was assessed using the Cohen kappa [40] and ranged from κ=.357 to .913 which is considered fair to almost perfect agreement beyond that of chance [41].

Results

Engagement by Platform

Fitbit had more followers than Garmin on each of the three platforms. Fitbit and Garmin both had the most followers on Facebook, then Twitter, and then Instagram. Despite the lower numbers of followers, both Fitbit and Garmin received the most engagement on Instagram. Conversely, although both brands have the largest number of followers on Facebook, the platform was the overall worst performer regarding "total engagement" per post. Table 1 shows the number of posts and followers, and mean number of "likes," "comments," "shares," and "total engagement" on each platform and for each brand on each platform.

Instagram posts received the most engagement by far with a mean of 4181.6 (SD 1413.3) "likes" per post, compared to just 47.1 (SD 118.9) and 65.7 (SD 50.3) for Facebook and Twitter, respectively. Instagram posts also received considerably more comments with a mean of 63.1 (SD 68.2) per post, compared to just 5.3 (SD 12.5) for Facebook and 3.0 (SD 9.7) comments per post for Twitter. Twitter posts were more likely to be shared, with a mean of 26.7 (SD 20.9) per post compared to just 3.3 (SD 20.4) per post on Facebook. Instagram does not offer a share function yet, despite this, "total engagement" for Instagram was still considerably higher at a mean of 4244.8 compared to a mean of 55.6 and 95.4 for Facebook and Twitter, respectively.

Brands' Use of Creative Elements

Chi-square tests of homogeneity indicated that there were differences in the use of these devices between the 2 brands. Fitbit was more likely to feature females (χ^2_3=49.4, P<.01) and indoor settings (χ^2_4=139.9, P<.01) and their posts emphasized social approval (χ^2_1=15.7, P<.01) and self-improvement (χ^2_1=35.2, P<.01) while delivering both positive (χ^2_2=77.6, P<.01) and rational (χ^2_2=19.7, P<.01) messages about the components or contents (χ^2_1=3.8, P<.05), or aesthetics (χ^2_1=8.1, P<.01) of the product. Fitbit was more likely to feature a nonwhite person (χ^2_2=19.9, P<.01) than Garmin, although for both brands most people in their images were white (84% for each). Unlike Garmin, Fitbit was also found to encompass a full suite of "lifestyle" factors into their posts by often featuring exercise or physical activity (χ^2_1=21.5, P<.01), incidental activity (χ^2_1=26.7, P<.01), weight loss (χ^2_1=19.4, P<.01), food and nutrition (χ^2_1=42.1, P<.01) and information related to sleep (χ^2_1=8.6, P<.01).

Garmin was more likely to feature males (χ^2_3=49.4, P<.01), celebrities (χ^2_1=46.2, P<.01), children (χ^2_1=12.1, P<.01), and animals (χ^2_1=23.4, P<.01). Garmin also featured exciting activities (χ^2_1=85.7, P<.01), scenic (χ^2_1=92.3, P<.01), and outdoor wilderness settings (χ^2_4=139.9, P<.01). Their posts often used rough and rugged themes (χ^2_1=86.2, P<.01) concerning setting or choice of activity. Garmin also made more emotional appeals (χ^2_2=19.7, P<.01), featured new or improved product features (χ^2_1=11.6, P<.01) and mentioned the product in their text (χ^2_1=27.7, P<.01) more often.

Table 1. Number of followers, mean engagement per platform and by platform and brand.

Platform brand	Followers	Posts, n[b]	Likes, mean (SD)	Comments/replies, mean (SD)	Shares/retweets, mean (SD)	Engagement[a], mean
Facebook						
Fitbit	1,846,974	62	18.3 (14.7)	6.1 (5.8)	0.5 (2.4)	24.9
Garmin	1,470,340	84	68.4 (153.1)	4.7 (15.7)	5.3 (26.7)	78.4
Total	—	146	47.1 (118.9)	5.3 (12.5)	3.3 (20.4)	55.6
Twitter						
Fitbit	313,000	156	84.2 (44.3)	3.8 (11.5)	35.6 (17.9)	123.6
Garmin	130,000	79	29.1 (40.8)	1.4 (3.4)	9.2 (14.2)	39.7
Total	—	235	65.7 (50.3)	3.0 (9.7)	26.7 (20.9)	95.4
Instagram						
Fitbit	415,581	58	4537.4 (1408.4)	113.4 (72.4)	—	4650.8
Garmin	260,864	70	3886.9 (1357.6)	21.5 (18.6)	—	3908.3
Total	—	128	4181.6 (1413.3)	63.1 (68.2)	—	4244.8

[a]Total engagement is the sum of the mean number of "likes," "comments," and "shares" per post.

Figure 1 illustrates the differences in creative elements used by each brand and the results of chi-square tests of homogeneity that were used to compare differences in frequency of use between brands. Full details of the frequency of use of each creative element by each brand, and the results of chi-square tests of homogeneity to compare differences in frequency of use between brands can be found in Multimedia Appendix 2.

Creative Elements Associated With Engagement

After controlling for brand and platform, the devices that were most influential on engagement were the mention of new or improved features (P<.01), displaying the product in the image (P<.01), or themes of self-improvement (P<.01). The inclusion of these devices was associated with engagement rates that were 90%, 30%, and 20% higher respectively compared to posts that did not contain these devices. In contrast, engagement rates were found to be between 16% to 45% lower when aesthetic claims (P<.01), specific product components or contents (P<.01), an outdoor setting (P<.01), the mention of a special offer or event (P<.01), having text overlaying an image (P<.01), using close-up images (P<.01), or mentioning a user's experience (P<.05) of the wearable were present in the post, when compared to posts that did not contain these devices. Table 2 presents the results of the multivariate negative binomial regression analyses of creative elements as predictors of "total engagement."

Figure 1. Use of creative elements by Fitbit and Garmin across Facebook, Twitter, and Instagram social media posts. Creative elements appearing in <5% for both Fitbit and Garmin are omitted. Significant differences between Fitbit and Garmin in use of a creative element are included in bold. Creative elements below the identity line are characteristic of Fitbit, while those above the identity line are characteristic of Garmin. The dotted lines connect creative elements that are polar opposites.

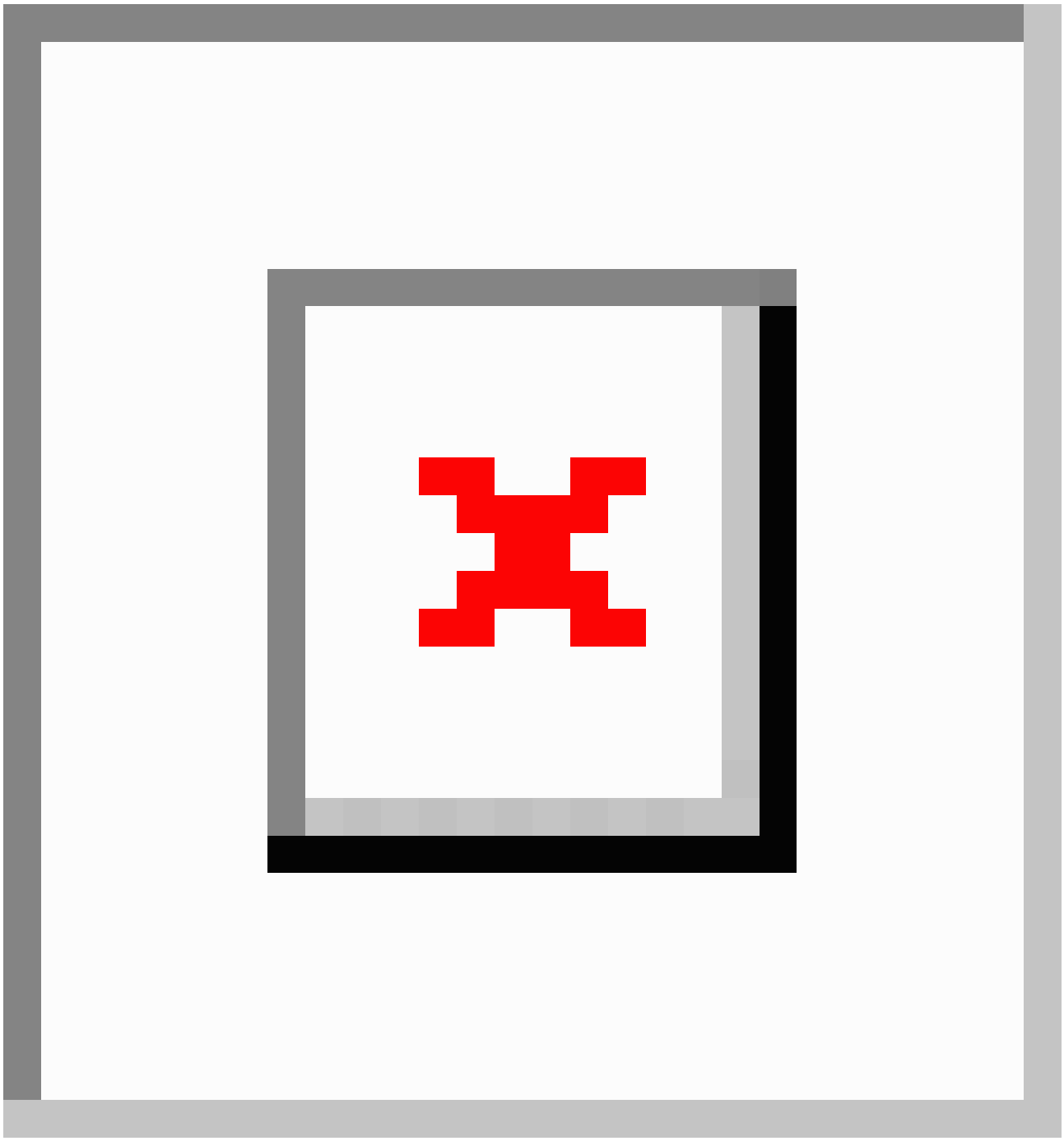


Table 2. Multivariate negative binomial regression analyses of creative elements as predictors of "total engagement."

Creative elements[a]	B[b]	SE	*P* value	IRR[c]
Intercept[a]	4.371	.117	.000	79.107
New or improved	0.618	.094	.000	1.856
Product in image	0.238	.075	.001	1.269
Self-improvement	0.179	.070	.010	1.196
Aesthetic claims	–0.441	.115	.000	0.643
Components or contents	–0.595	.074	.000	0.552
Setting[d]				
Outdoor wilderness	–0.378	.108	.000	0.685
Outdoor nature	–0.469	.121	.000	0.626
Outdoor cityscape	–0.175	.105	.096	0.840
Indoor setting	–0.109	.092	.234	0.897
Special offer or event	–0.247	.079	.002	0.781
Text over image	–0.312	.100	.002	0.732
Audience in image[e]				
Close up image	–0.261	.081	.001	0.770
Image view through own eyes	0.033	.100	.745	1.033
User experience	–0.242	.106	.022	0.785
Children present	–0.332	.218	.128	0.718
Exercise or physical activity	0.111	.068	.104	1.118

[a]Platform (Facebook, Twitter, and Instagram) included in the final model as a covariate, only creative elements with a $P<.1$ and that changed the main effects by 10% were retained in the final model.

[b]B: beta coefficient.

[c]IRR: incidence rate ratio (is the exponentiation of the regression coefficient, which equates to the odds ratio).

[d]Reference category: no setting.

[e]Reference category: audience not present.

Discussion

Principal Findings

This study examined the social media activity of leading brands of wearable activity trackers to make platform and content recommendations for future social media-based health promotion and intervention efforts. The study found that while wearables attracted their largest following on Facebook compared with Twitter or Instagram, engagement with posts was markedly higher on Instagram. Differences in the types of creative elements used were apparent between Fitbit and Garmin. In particular, Fitbit posts were characterized by featuring females and having an upbeat or lighthearted tone whereas Garmin posts featured men, with adventurous and outdoorsy themes. Featuring themes of self-improvement, new product features, or the product in the image were each associated with higher rates of engagement when compared to posts that did not contain these creative elements.

Both Fitbit and Garmin attracted the greatest number of followers on Facebook, which is the largest platform with more than two billion active users, compared to 700 and 328 million for Twitter and Instagram, respectively [1]. However, despite the smaller number of users and followers, it was Instagram that attracted rates of engagement that were from 30 to almost 200 times higher. This is consistent with previous research examining Instagram, where a study comparing dissemination of disease-outbreak information on Twitter and Instagram found higher rates of engagement on Instagram [12]. Furthermore, outside the health research domain, market research has suggested brand advertising receives better audience engagement on Instagram compared with Facebook [42]. Instagram typically has a younger and female audience known to be some of the most prolific social media users [43], and this may also contribute to the higher rates of engagement. It is also possible that engagement is influenced by differences between social media platforms and their respective complex predictive algorithms that control the percentage of followers who view a post, which we were not able to account for in this study.

Creative elements that were associated with the highest engagement were themes of self-improvement, highlighting "new" products or features, and featuring the wearable device in the image. In general, these findings are consistent with previous literature. For example, the scoping study by Van Kessel et al [44] that focused on the development of a social

media physical activity intervention for adolescent girls, found that the girls reported a desire for the content of an inspirational, self-improvement nature. Similarly, inspirational imagery taken from Instagram has been found to have a positive effect on motivation to pursue healthy goals in young women, although with the caveat that this needs to be balanced against the potential for Instagram to have a negative effect on body image [45]. In commercial marketing practice, featuring "new" and "improved" products are very common creative elements. Evidence has shown that consumers pay more attention and respond more positively to an advertisement with the word "new" [46] tapping into the novelty effect. As for featuring the product in the image, direct evidence for the effectiveness of such creative elements is mixed [31,32]. However, there is evidence to suggest that featuring the product aids brand awareness and recall [47]. Findings of the current study support this, suggesting that featuring the device is the more effective approach for social media posts. This is likely to be because social media posts need to stand out against the cluttered background, where users are presented with large volumes of posts vying for their attention [48].

The clear differences between how Fitbit and Garmin presented and promoted their products on social media are likely to reflect market segmentation and differences in how Fitbit and Garmin are positioning their respective products. Principles of market segmentation (eg, [32]) have been integrated with social marketing efforts over the past decade [49]. Health promotion and intervention efforts are increasingly mirroring this targeted approach with increasing interest in the use of tailored eHealth and mHealth interventions in recent years and an associated expanding body of evidence suggesting that this tailoring may increase their effectiveness in changing behavior [50-53].

Strengths and Limitations

This study considered three social media platforms and took a rigorous duplicate coding approach, both of which serve to strengthen the findings. A similar methodology to that presented in the Rus and Cameron [13] analysis of Facebook-based diabetes support groups is followed here, and both studies seek to extend the emerging field of research examining user engagement in order to increase the efficacy of social media-based health interventions and communication. A limitation of the current study was that engagement was operationalized in simple count terms (ie, no. of "likes," "comments," and "shares"). The terms "likes," "comments," and "shares" are qualitatively different, involving different levels of effort, and endorsement or enjoyment of content. A further limitation of our study is our use of the Stewart and Furse [31] codebook that was developed to assess the use of television execution techniques on sales effectiveness. Here, we cannot comment on the relationship between sales effectiveness and social media engagement, or if such a relationship does exist. Also, our sample of posts covered a specific period (December to February), and we cannot comment on whether our findings are generalizable to alternative three-month periods. The observational design used here means that differences in engagement on each platform due to user demographics cannot be explored. It should also be noted that users of social media are not necessarily representative of the broader population and our findings must be generalized with caution. Further research examining qualitative aspects (eg, content analysis of comments) may provide further useful insights for future posts, as would detailed examination of the creative elements used to advertise different models of wearable devices (which are likely to offer varying features and benefits to users). Our results indicate that brands can leverage user participation by encouraging "sharing" of content, although to a much lesser extent than engaging through one-click "likes" or writing comments, future research should seek to explore this to understand how to increase user input to maximize engagement. Also, it is important to acknowledge that engagement with social media posts does not necessarily reflect real-life behavior such as purchasing of wearable products, or adherence to the healthy lifestyle elements promoted in the Fitbit and Garmin posts.

Implications

This research offers several insights that may be useful for researchers developing social media-based health promotion campaigns and interventions in the future. First, it appears that Instagram is a promising platform for health promotion. In the literature to date, most efforts have been focused on Facebook, presumably as one of the earliest platforms with the largest user base. However, our study and background literature review suggest that Instagram achieves better reach to its audience, and vastly better engagement, highlighting the promise of this platform into the future. Key creative elements associated with highest "engagement" were the use of self-improvement themes, featuring "new" products and featuring the product in the image. Nonetheless, unlike wearables, most health promotion efforts do not have a concrete product to promote. However, intangible notions of improved health can be represented by tangibles, for example by focusing on the endpoint of related health benefits such as improved mood, vitality, and sleep quality, rather than the process of achieving these benefits. This finding also suggests that simple, clear, and direct messages may be best suited to social media. Just as wearable brands frequently post details of their new products and new features of their products, health promotion efforts should seek to refresh, rotate, and renew their health messages on a regular basis to attract engagement from social media users. Finally, differences in the Fitbit and Garmin approach underscore the need for health promotion efforts to clearly define their target population segment and tailor the messaging toward them. Clear examples of this from our study were targeting of gender, setting (indoors, urban, wilderness), and use of celebrities. That being said, there is evidence to suggest that people who engage with brands on social media are often existing users of the brand [33] and future work is required to determine how much social media engagement does reflect past or predict future behavior. Appealing content is more likely to receive attention and engagement from study participants. Several studies [6-8] have now demonstrated that higher engagement with the online component of an intervention translates to greater adherence to the behavior change aspects of the intervention, and is related to efficacious outcomes. This study offers important insights for researchers that will aid in the development of social media-based health promotion and intervention efforts that are

appealing, and are therefore able to maximize the potential of these approaches.

Conclusion

This study examined the social media activity of two wearable brands across three key social media platforms and provided novel insights into enhancing engagement. Future work should consider Instagram as a delivery platform and incorporate principles of market segmentation, or tailoring. Health messages on social media should be clear, direct, refreshed regularly, incorporate inspirational messages and imagery and be focused on tangible end products of health in order to maximize engagement and therefore the potential of this approach for positive behavior change.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Creative elements and descriptions.

[PDF File (Adobe PDF File), 33KB - jmir_v20i12e10911_app1.pdf]

Multimedia Appendix 2

Frequency of each creative element and comparison of use between brands.

[PDF File (Adobe PDF File), 33KB - jmir_v20i12e10911_app2.pdf]

References

1. Statistica. Statistica. 2017. Most popular social networks worldwide as of October 2018, ranked by number of active users (in millions) URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [accessed 2018-10-18] [WebCite Cache ID 73FxllW09]
2. Greenwood S, Perrin A, Duggan M. Pew Research Centre. 2016. Social media update 2016, in Facebook usage and engagement is on the rise, while adoption of other platforms holds steady URL: http://www.pewinternet.org/2016/11/11/social-media-update-2016/ [accessed 2018-10-18] [WebCite Cache ID 73FxqDpW7]
3. Statistica. Statistica. Social media and user-generated content URL: https://www.statista.com/markets/424/topic/540/social-media-user-generated-content/ [accessed 2018-10-18] [WebCite Cache ID 73FydNDmh]
4. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
5. Al-Eisa E, Al-Rushud A, Alghadir A, Anwer S, Al-Harbi B, Al-Sughaier N, et al. Effect of Motivation by “Instagram” on Adherence to Physical Activity among Female College Students. Biomed Res Int 2016;2016:1546013 [FREE Full text] [doi: 10.1155/2016/1546013] [Medline: 27034927]
6. Hales SB, Davidson C, Turner-McGrievy GM. Varying social media post types differentially impacts engagement in a behavioral weight loss intervention. Transl Behav Med 2014 Dec;4(4):355-362 [FREE Full text] [doi: 10.1007/s13142-014-0274-z] [Medline: 25584084]
7. Leahey T, Rosen J. DietBet: A Web-Based Program that Uses Social Gaming and Financial Incentives to Promote Weight Loss. JMIR Serious Games 2014 Feb 07;2(1):e2 [FREE Full text] [doi: 10.2196/games.2987] [Medline: 25658966]
8. Turner-McGrievy GM, Tate D. Weight loss social support in 140 characters or less: use of an online social network in a remotely delivered weight loss intervention. Transl Behav Med 2013 Sep;3(3):287-294 [FREE Full text] [doi: 10.1007/s13142-012-0183-y] [Medline: 24073180]
9. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
10. Lewis E. Financial advertising: The history of advertising. In: Financial Advertising for Commercial and Savings Banks, Trust Title Insurance, and Safe Deposit Companies, Investment Houses. New York: Taylor & Francis; 1985.
11. Ryan J, Edney S, Maher C. Engagement, compliance and retention with a gamified online social networking physical activity intervention. Transl Behav Med 2017 Dec;7(4):702-708 [FREE Full text] [doi: 10.1007/s13142-017-0499-8] [Medline: 28523603]
12. Guidry J, Jin Y, Orr C, Messner M, Meganck S. Ebola on Instagram and Twitter: How health organizations address the health crisis in their social media engagement. Public Relations Review 2017 Sep;43(3):477-486. [doi: 10.1016/j.pubrev.2017.04.009]

13. Rus HM, Cameron L. Health Communication in Social Media: Message Features Predicting User Engagement on Diabetes-Related Facebook Pages. Ann Behav Med 2016 Dec;50(5):678-689. [doi: 10.1007/s12160-016-9793-9] [Medline: 27059761]
14. International Data Corporation. International Data Corporation (IDC). 2016. Fitness trackers in the lead as wearables market grows 3.1% in the third quarter, according to IDC URL: http://www.idc.com/getdoc.jsp?containerId=prUS41996116 [accessed 2018-10-18] [WebCite Cache ID 73FzEYYTS]
15. Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The Pace of Technologic Change: Implications for Digital Health Behavior Intervention Research. Am J Prev Med 2016 Dec;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]
16. Garmin Ltd. 2016. Garmin Ltd. Annual Report URL: https://www8.garmin.com/aboutGarmin/invRelations/reports/2016_Annual_Report.pdf [accessed 2018-10-18] [WebCite Cache ID 73FzLMkSm]
17. Fitbit Inc. Fitbit Inc. Annual Report URL: http://d1lge852tjjqow.cloudfront.net/CIK-0001447599/f7a564df-cf62-4681-90a4-a3ef813b0d83.pdf [accessed 2018-10-18] [WebCite Cache ID 73FzRrx7T]
18. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13(1):35-39 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
19. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
20. Hartman SJ, Nelson S, Weiner L. Patterns of Fitbit Use and Activity Levels Throughout a Physical Activity Intervention: Exploratory Analysis from a Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 Feb 05;6(2):e29 [FREE Full text] [doi: 10.2196/mhealth.8503] [Medline: 29402761]
21. Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the Fitbit to Measure Adherence to a Physical Activity Intervention Among Overweight or Obese, Postmenopausal Women: Self-Monitoring Trajectory During 16 Weeks. JMIR Mhealth Uhealth 2015 Nov 19;3(4):e96 [FREE Full text] [doi: 10.2196/mhealth.4229] [Medline: 26586418]
22. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women. Am J Prev Med 2015 Sep;49(3):414-418 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
23. Ridgers ND, McNarry M, Mackintosh K. Feasibility and Effectiveness of Using Wearable Activity Trackers in Youth: A Systematic Review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
24. Patel MS, Asch D, Volpp K. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 03;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
25. Piwek L, Ellis DA, Andrews S, Joinson A. The Rise of Consumer Health Wearables: Promises and Barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
26. Cheng H, Kotler P, Lee N. Social Marketing for Public Health: Global Trends and Success Stories. Sudbury: Jones & Barlett Learning; 2009.
27. Bloomberg Technology. Jawbone liquidation inflicts losses, prompts CEO move to health URL: https://bloom.bg/2EX9bPB [WebCite Cache ID 73Fznarcd]
28. Ashley C, Tuten T. Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. Psychol. Mark 2014 Dec 11;32(1):15-27. [doi: 10.1002/mar.20761]
29. Carboni JL, Maxwell SP. Effective Social Media Engagement for Nonprofits: What Matters? J Pub Nonprof Aff 2015 Mar 02;1(1):18-28. [doi: 10.20899/jpna.1.1.18-28]
30. Stewart DW, Koslow S. Executional Factors and Advertising Effectiveness: A Replication. Journal of Advertising 1989 Oct;18(3):21-32. [doi: 10.1080/00913367.1989.10673158]
31. Stewart DW, Furse D. Effective Television Advertising: A Study of 1000 Commercials. Lexington: Prentice Hall & IBD; 1985.
32. Hartnett N, Kennedy R, Sharp B, Greenacre L. Creative That Sells: How Advertising Execution Affects Sales. Journal of Advertising 2015 Sep 29;45(1):102-112. [doi: 10.1080/00913367.2015.1077491]
33. Stanton J. Comparative Effectiveness of Executional Elements in TV Food Advertising on Recall and Persuasion. Journal of International Food & Agribusiness Marketing 2004 Mar 23;15(1-2):137-152. [doi: 10.1300/J047v15n01_07]
34. Phillips D, Stanton J. Age-related differences in advertising: Recall and persuasion. J Target Meas Anal Mark 2004 Sep 1;13(1):7-20. [doi: 10.1057/palgrave.jt.5740128]
35. Bellman S, Schweda A, Varan D. Interactive TV advertising: iTV ad executional factors. Journal of Business Research 2012 Jun;65(6):831-839. [doi: 10.1016/j.jbusres.2011.01.003]
36. Rust R, Cooil B. Reliability Measures for Qualitative Data: Theory and Implications. Journal of Marketing Research 1994 Feb;31(1):1. [doi: 10.2307/3151942]
37. IBM. Armonk, NY; 2015. SPSS Statistics 23 for Windows URL: https://www.ibm.com/analytics/spss-statistics-software [accessed 2018-10-18] [WebCite Cache ID 73G0UpRjA]

38. Australian Government. 2010. Australian burden of disease study: fatal burden of disease 2010 URL: https://www.aihw.gov.au/reports/burden-of-disease/australian-bod-study-fatal-burden-of-disease-2010/contents/table-of-contents [accessed 2018-10-18] [WebCite Cache ID 73G0bnv0b]
39. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. Source Code Biol Med 2008 Dec 16;3:17 [FREE Full text] [doi: 10.1186/1751-0473-3-17] [Medline: 19087314]
40. Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
41. Landis J, Koch G. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310]
42. Coelho R, Oliveira D, Almeida M. Does social media matter for post typology? Impact of post content on Facebook and Instagram metrics. Online Information Review 2016 Aug 08;40(4):458-471. [doi: 10.1108/OIR-06-2015-0176]
43. Duggan M. Pew Research Center. 2015. The demographics of social media users URL: http://www.pewinternet.org/2015/08/19/the-demographics-of-social-media-users/ [accessed 2018-10-18] [WebCite Cache ID 73G0oYwAH]
44. Van Kessel G, Kavanagh M, Maher C. A Qualitative Study to Examine Feasibility and Design of an Online Social Networking Intervention to Increase Physical Activity in Teenage Girls. PLoS One 2016;11(3):e0150817 [FREE Full text] [doi: 10.1371/journal.pone.0150817] [Medline: 26934191]
45. Tiggemann M, Zaccardo M. Exercise to be fit, not skinny: The effect of fitspiration imagery on women's body image. Body Image 2015 Sep;15(Supplement C):61-67. [doi: 10.1016/j.bodyim.2015.06.003] [Medline: 26176993]
46. Sung B, Hartley N, Vanman E, Phau I. How can the word “NEW” evoke consumers' experiences of novelty and interest? Journal of Retailing and Consumer Services 2016 Jul;31:166-173. [doi: 10.1016/j.jretconser.2016.02.010]
47. Romaniuk J, Sharp B. Conceptualizing and measuring brand salience. Marketing Theory 2016 Aug 11;4(4):327-342. [doi: 10.1177/1470593104047643]
48. Nelson-Field K, Riebe E, Sharp B. What's Not to “Like?”. JAR 2012 Jun 01;52(2):262-269. [doi: 10.2501/JAR-52-2-262-269]
49. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. Patient Educ Couns 1999 Feb;36(2):107-117. [Medline: 10223016]
50. Kubacki K, Ronto R, Lahtinen V, Pang B, Rundle-Thiele S. Social marketing interventions aiming to increase physical activity among adults. Health Education 2017 Jan 03;117(1):69-89. [doi: 10.1108/HE-02-2016-0008] [Medline: 4283]
51. Neville LM, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. Int J Behav Nutr Phys Act 2009 Jun 03;6(1):30-12 [FREE Full text] [doi: 10.1186/1479-5868-6-30] [Medline: 19490649]
52. Pellegrini CA, Steglitz J, Hoffman S. e-Health intervention development: a synopsis and comment on “What Design Features are Used in Effective e-Health Interventions? A Review Using Techniques From Critical Interpretive Synthesis”. Transl Behav Med 2014 Dec;4(4):342-345 [FREE Full text] [doi: 10.1007/s13142-014-0283-y] [Medline: 25584082]
53. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228.e1. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]

Abbreviations

AIC: Akaike information criterion
AIDA: awareness-interest-desire-action
BIC: Bayesian information criterion
eHealth: electronic health
mHealth: mobile health

Edited by G Eysenbach; submitted 30.04.18; peer-reviewed by K Winfree, A Sullivan, P Bondaronek, D Li; comments to author 18.08.18; revised version received 05.09.18; accepted 10.09.18; published 10.12.18.

Please cite as:
Edney S, Bogomolova S, Ryan J, Olds T, Sanders I, Maher C
Creating Engaging Health Promotion Campaigns on Social Media: Observations and Lessons From Fitbit and Garmin
J Med Internet Res 2018;20(12):e10911
URL: https://www.jmir.org/2018/12/e10911/
doi:10.2196/10911
PMID:30530449

Original Paper

Identifying Information Needs for Hirschsprung Disease Through Caregiver Involvement via Social Media: A Prioritization Study and Literature Review

Kristy DM Wittmeier[1,2], BMRPT, MSc, PhD; Kendall Hobbs-Murison[2], RN, BN; Cindy Holland[1,2,3], RN, MN, NP; Elizabeth Crawford[4]; Hal Loewen[5], MLIS; Melanie Morris[1,2,3], MD, FRCSC; Suyin Lum Min[1,2,3], MD, FRCSC; Ahmed Abou-Setta[6], MD, PhD; Richard Keijzer[1,2,3], MD, MSc, PhD

[1]Department of Pediatrics and Child Health, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

[2]Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada

[3]Department of Surgery, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

[4]Swish Productions Ltd, Winnipeg, MB, Canada

[5]Neil John McLean Library, University of Manitoba, Winnipeg, MB, Canada

[6]George and Fay Yee Centre for Healthcare Innovation, Winnipeg, MB, Canada

Corresponding Author:
Richard Keijzer, MD, MSc, PhD
Department of Surgery
Max Rady College of Medicine, Rady Faculty of Health Sciences
University of Manitoba
AE402-840 Sherbrook Street
Winnipeg, MB, R3A 1S1
Canada
Phone: 1 2047871246
Fax: 1 2047874618
Email: richardkeijzer@gmail.com

Abstract

Background: Patient and public involvement in health research is important to produce relevant and impactful results.

Objective: This paper aimed to prioritize and summarize Hirschsprung disease (HD)–related information needs among caregivers of children with HD and pediatric surgeons through partnership with a parent-initiated social media campaign.

Methods: We conducted a Web-based survey with the 2 stakeholder groups to identify information needs. The caregiver survey was conducted through a global Web-based community, and the surgeon survey was distributed to members of the Canadian Association of Paediatric Surgeons (CAPS). We conducted a literature review to identify evidence on the prioritized topics.

Results: Our findings showed that 54.9% (89/162) of the individuals completed the caregiver survey and 23.8% (52/218 listed members) of the pediatric surgeons completed the survey distributed through CAPS. Only 20% (18/89) of the caregivers reported being very satisfied or satisfied with the current HD-related resources. A final prioritized list of information needs included bowel management, nutrition and growth, infection, perianal irritation, gastrointestinal pain, surgical diagnostics, and surgical complications. In total, 87 studies were included in the literature review, which included the following: 8 reviews, 2 randomized controlled trials, 74 cohort studies, and 3 practice guidelines. Two priority issues identified by caregivers had only a single study that met the inclusion criteria, whereas 1 topic had none.

Conclusions: With caregiver and surgeon input, we identified 7 information priority areas related to HD. A review of the literature on the priorities found little evidence to support the development of high-quality guidelines. More research is necessary to meet the information needs related to HD as identified by stakeholders.

(J Med Internet Res 2018;20(12):e297) doi:10.2196/jmir.9701

KEYWORDS

Hirschsprung disease; caregivers; social media; research; surgeons; surveys and questionnaires

Introduction

> *For a family affected by Hirschsprung's disease (HD), the first few years can be a roller coaster of total normalcy and repeated hospitalizations. To learn more about this illness I created an online community for families and patients living with HD.* [Parent of child with HD]

There is an increasing focus on patient-centered or patient-oriented research to improve the relevance and impact of research. Programs and strategies such as INVOLVE [1], the Patient-Centered Outcomes Research Institute [2], and the Strategy for Patient-Oriented Research [3] are collectively functioning to promote and support patient and public involvement in health research [4]. Patient-centered or patient-oriented research emphasizes the involvement of patients and members of the public throughout all phases of research in a way that reflects the principles of inclusiveness, support, mutual respect, and cobuilding [3,5]. There is potential for tokenistic involvement if the public is merely informed of research without the ability to contribute to decision making within a project [4]. Instead, a combined method is recommended where patients or members of the public are involved as decision-making members of the research team, and broader consultation is conducted to help ensure representative input [4,6]. The highest level of involvement is patient- or public-led research.

We have previously conducted a parent-partnered study that examined the reach and responsiveness of a Web-based community and social media campaign developed to connect families affected by HD [7]. This initial work demonstrated that this community was highly responsive and the campaign helped to connect families across the globe. We then jointly developed this study to complete a needs assessment via social media to prioritize information needs related to HD. A secondary aim was to identify and summarize the best available evidence for each of the identified priorities.

Methods

Study Design

Our research team included a parent partner, 3 pediatric surgeons, a pediatric surgical nurse practitioner, a knowledge translation researcher, and a clinical nurse, all of whom have experience working with children with rare diseases. We obtained approval for the study from the University of Manitoba Health Research Ethics Board and the Health Sciences Centre Pediatric Research Impact Committee, and all participants provided informed consent.

The study involved 3 stages, which were as follows: surveying HD caregivers and pediatric surgeons to identify priority information needs regarding HD management; prioritization of the information needs, and a literature review to summarize the existing literature. Survey results are reported with guidance from the Checklist for Reporting Results of Internet E-Surveys using internet analytics as previously described [8].

Needs Assessment

Caregiver Survey

This survey was developed to prioritize the information needs of caregivers and those living with HD. The survey asked caregivers to describe the problems most frequently encountered in caring for a child with HD and their satisfaction with HD-related resources that were available at that time (Multimedia Appendix 1). We also asked parents about the format they preferred to receive the study results. The survey was conducted using Fluid Surveys (Ontario, Canada), allowing for secure and anonymous data collection. We launched the survey in November 2013. It was posted for 1 month on the HD Facebook page along with information about the study. Reminders to complete the survey were frequently posted to promote participation by the site administrator. Facebook "likes," reposts, and other sharing mechanisms (eg, to other sites such as reachhd.org [9] and HD-related group pages on Facebook) were also used to increase reach. We used an affiliated HD organization as the survey landing page (reachhd.org) and collected Google Analytics data to track survey metrics [8].

Pediatric Surgeon Survey

A survey was also administered to all members of the Canadian Association of Paediatric Surgeons (CAPS) using their Web-based survey tool. The survey was available for 1 month and reminder emails were sent to promote participation. We collected surgeon demographics (eg, number of years in practice and number of patients they manage with HD) and the top 5 HD-related medical issues they encounter. CAPS members were surveyed a second time 4 months later to determine the resources they used to guide the management of their patients with HD (Multimedia Appendix 2).

One team member (KHM) conducted a content analysis of the results from both caregiver and CAPS surveys. Common themes of information needs were identified and categorized as priority issues for each stakeholder group. A pediatric surgical nurse practitioner (CH) reviewed and verified the prioritized information needs from the caregiver survey, whereas 2 of the pediatric surgeon team members (RK and MM) verified those from the surgeon survey. These issues were then merged to create 1 list of the top 7 most common issues identified by caregivers and pediatric surgeons.

Merging Priorities

A modified Delphi approach was then used to seek consensus among team members on the prioritized list. This process combined direct discussion and 2 rounds of anonymous survey of the research team members [10]. An a priori decision to give primary importance to the caregiver-identified priorities was adhered to. The top 7 most common information needs were then presented back to the Web-based caregiver community and surgeon stakeholder groups for validation.

Literature Review

We then conducted a literature review to identify evidence to address the top 7 prioritized issues. A health science librarian in collaboration with an expert in review methodology (AAS)

developed the search strategy. A second librarian conducted an independent peer review of the search strategy using the Peer Review of Electronic Search Strategies checklist [11]. The search was limited to systematic reviews or clinical practice guidelines published since 2000 and randomized controlled trials, clinical trials, cohorts, or case series from January 2010 to March 2015. Databases searched included Ovid Medline, Ovid EMBASE, CENTRAL, and EBSCOhost CINAHL. Two reviewers (KHM and CH) screened titles and abstracts independently for inclusion based on the predetermined inclusion criteria (Multimedia Appendix 3). For studies that were accepted by both reviewers, full texts were obtained and evaluated independently. Conflicts were resolved through consensus or after discussion with a third reviewer.

Results

Needs Assessment

Caregiver Survey

Of those who consented to participate, 54.9% (89/162) completed the caregiver survey. Moreover, 66% of the surveys were completed within the first week of posting the link. Short-segment HD was the most common diagnosis (36/89, 40%) reported, and most patients had received a diagnosis within the first month of life (78/89, 88%; Table 1).

When asked about satisfaction with current HD-related resources, only 20% (18/89) were satisfied or very satisfied with current resources, 42% (37/89) somewhat satisfied and (34/89) 38% not very or not satisfied.

Canadian Association of Paediatric Surgeons Survey

Of the 218 CAPS members, 23.8% (52/218) responded to the first survey and 46 (46/218, 21.1%) responded to the second survey (Table 1). The sources that Canadian pediatric surgeons report using to guide their clinical HD-related practice are shown in Figure 1.

Merging Priorities

The final top 7 priority information needs were summarized as an infographic (Figure 2) that was shared with the Web-based HD caregiver community through Facebook. Through a survey, community members were asked whether they agreed with the listed priority issues and for their feedback on the infographic format. Of the 96 individuals who provided feedback (90 caregivers, 6 people with HD), 91% (87/96) agreed or strongly agreed with the priority issues.

Literature Review

We identified 8 reviews, 2 randomized trials, 59 retrospective cohort studies, 15 prospective cohort studies, and 3 practice guidelines that provided evidence on the 7 priority needs. The evidence is summarized by topic area in the following sections with emphasis on the highest levels of evidence in sections where a large number of studies were included (surgical complications and long-term outcomes).

Bowel Management

One of the major problems faced by children with HD relates to bowel routines. Even after corrective surgery, many patients with HD still experience bowel management issues. In children who undergo surgery, systematic evaluation and the use of a structured and tailored approach to successfully treat persistent incontinence or soiling is recommended [12-14]. No studies covered the subtheme of toilet training specific to HD. Some authors suggest Botox injections to relax the sphincter muscle and promote defecation [15]. Single-center studies have also shown improvement in bowel function post-Soave with pelvic floor exercises [16] or Malone antegrade enemas [17]. All these interventions aim to achieve social continence.

Infection

Hirschsprung's associated enterocolitis (HAEC) is a serious, potentially life-threatening complication with an estimated incidence ranging from 4.6% to 54% [18]. Risk factors for developing HAEC are unclear. One systematic review that examined HAEC in relation to *Clostridium difficile* infection found 98 reported cases of HAEC related to *Clostridium difficile* infection from 1974 to 2014 [19]. There was insufficient data to analyze the role of other pathogens. One retrospective study found 58% of patients with HAEC had allergy to cow's milk [20]. In terms of prevention, 1 report found no effect of probiotics administration when compared with placebo on HAEC incidence or recurrent HAEC postsurgery [21]. Similarly, a retrospective review found no difference in HAEC incidence or anastomotic stricture rates in children who either had or did not have routine anal dilations prescribed post the pull-through procedure [22].

With respect to surgical techniques, a systematic review reported a low incidence (10.2%) of postoperative HAEC after the transanal 1-stage pull-through procedure technique with HAEC successfully managed conservatively in majority of patients (81.5%) [18]. A single-center retrospective cohort study found that HAEC incidence decreased from 33.9% to 1.9% postsurgery with a transanal rectal mucosectomy and partial internal anal sphincterotomy [23]. We found no studies regarding caregivers' concerns about the association between HD and susceptibility to common flus and colds.

Perianal Irritation

One small (n=4) single-center pilot study suggested that the use of zinc oxide ointment with potato-derived protease inhibitors may reduce the otherwise intractable protease-induced perianal skin irritation in infants with long-segment HD [24].

Nutrition and Growth

One single-center retrospective study found that growth and development in the first year of life were not different between infants with short-segment HD and those with long-segment HD [25].

Gastrointestinal Pain

No studies that met our inclusion criteria were identified for this topic.

Table 1. Demographics of survey respondents.

Demographics	n (%)
HD[a] community survey	
Year of birth of individual living with HD (n=86)	
2012-2014	21 (24)
2009-2011	34 (40)
2006-2008	16 (19)
1995-2005	13 (15)
Before 1995	2 (2)
Type of HD (n=89)	
Ultrashort segment	16 (18)
Short segment	36 (40)
Long segment	16 (18)
Total colonic	13 (15)
Unsure	8 (9)
Age of HD diagnosis (n=89)	
0-1 months	78 (88)
2-12 months	8 (9)
13 months-4 years	2 (2)
>4 years	1 (1)
Pediatric surgeon survey[b] (n=52)	
Length of practice (years)	
0-5	11 (21)
6-15	18 (35)
>15	23 (44)
Cases of HD seen per year	
≤5	30 (58)
>5	22 (42)
Multidisciplinary follow up offered in the clinic	
Yes	12 (23)

[a]HD: Hirschsprung disease.

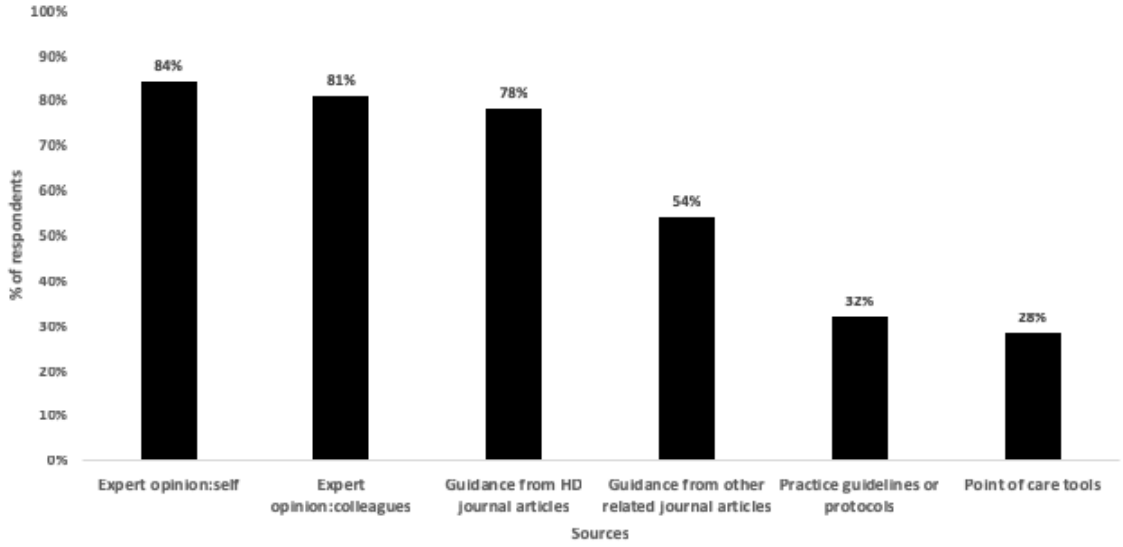
[b]Surgeon demographics were similar for both surveys; demographics from the initial survey only are reported in Table 1.

Diagnostics

Contrast Enema

Contrast enema (CE) has often been used as an adjunct to rectal biopsies for diagnosing HD. Barium or water-soluble contrast is instilled into the rectum to assess the transition zone. We found 10 articles published between 2006 and 2014 related to CE and its utility in HD [26-35]; five primary studies concluded that CE had either a low specificity, was not useful, or had a high false positive and negative rate [26,29-32]. A systematic review reported a sensitivity rate of 70% and specificity rate of 83% [28]. Wong et al [35] found that the addition of a delayed radiograph following CE raised the sensitivity from 69% to 100% but reduced the specificity from 89% to 78%. The literature suggests that CE is not as sensitive or specific as rectal biopsy to diagnose HD. The addition of a delayed film may increase sensitivity but lower specificity. CE may be helpful to raise suspicion of total colonic HD if certain criteria are found [35]. CE should not be used alone as a single method in the diagnosis of HD because it could be misleading and underestimate the extent of HD.

Figure 1. Pediatric surgeon's sources of information related to clinical Hirschsprung disease (HD) practice.



Rectal Biopsies

Classically, hematoxylin and eosin with or without acetylcholinesterase (AChE) are used in rectal biopsies to evaluate ganglion cells and nerve trunk hypertrophy respectively. Morris et al [36] state that calretinin immunohistochemistry may be superior to AChE in the context of total aganglionosis, superficial biopsies, and prematurity. Volpe et al [37] describe that with nerve hypertrophy, calretinin is a reliable marker for the transition zone. Uniformly, all authors concluded that calretinin immunohistochemistry is a reliable modality to diagnose HD and is equivalent if not superior to AChE with a sensitivity of 93.3% and specificity of 100% [38].

Several authors assessed different rectal biopsy techniques. Rectal suction biopsy (reserved for infants) [39], jumbo forceps [40], and full thickness biopsies [41] were all adequate to obtain tissue diagnosis for HD. Hematoxylin and eosin with or without AChE and with the addition of calretinin is effective in making the diagnosis via different techniques of procuring the rectal biopsy.

Surgical Complications & Long-Term Outcomes

Complications

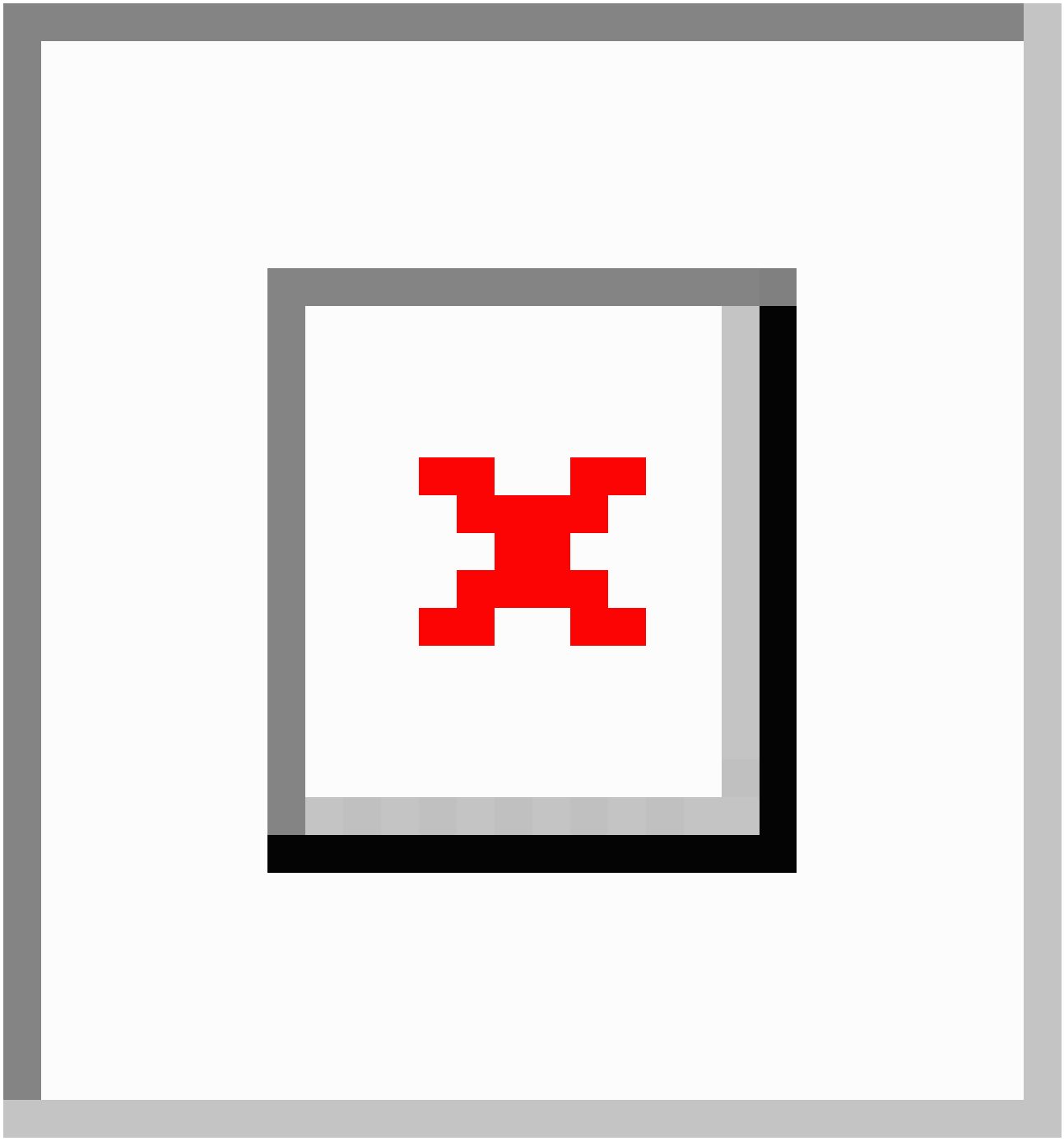
A 2013 systematic review [42] compared the outcomes between 444 transanal endorectal pull-through procedures and 348 conventional transabdominal approaches (including Soave, Duhamel, Swenson, and Rehbein procedures). Transanal endorectal pull-through procedures had shorter operative time and hospital stay, less postoperative soiling or incontinence and constipation, and no difference in postoperative enterocolitis. Similarly, Gosemann et al [43] found an advantage of the transanal over the open approaches in their systematic review.

Yang and Tang published in an abstract [44] the randomization of 54 children to a laparoscopic endorectal pull-through with a long or short cuff and reported that patients with a long cuff had a lower incidence of enterocolitis and better defecation in the first 6 months after surgery. The defecation frequency was similar 12 months after surgery. All authors considered a single stage approach standard of care for the treatment of short-segment HD. Most studies confirm the benefit of a transanal endorectal pull-through approach with or without laparoscopy over an open abdominal approach. Some studies reported that the transanal approach is associated with an increased rate of incontinence after the surgery [45,46]. Incidence rates of postoperative enterocolitis seem comparable among the different surgical approaches.

Long-Term Outcomes

Two prospective cohort studies evaluated the bowel function and quality of life in adults operated for HD during childhood. Ieiri et al [47] found that more than 85% of patients reported satisfactory bowel function (“good or excellent score”). Only 21.4% reported a normal score, and 16.7% and 19% reported incontinence and soiling, respectively. Jarvi et al performed a population-based study that included age- and sex-matched controls and reported that the overall bowel function score was lower in patients with HD, resulting in social problems associated with bowel function [48].

Figure 2. Infographic of top priorities and concerns related to Hirschsprung disease (HD).



Discussion

Research funders and public involvement organizations advocate for public participation to ensure that research is relevant, of high priority, and more easily translated into practice upon completion [1,5,49]. The public's involvement can range from receiving the results of research findings to providing input and guidance to the research team to a partnership role as primary or coinvestigators [3,50,51]. In this study, a parent partnered with our research team, helping to set the agenda for the research and contributing to project decisions. The involvement of a parent on the research team and input from the survey respondents via a social media campaign were critical in highlighting areas in which more resources and research are needed to guide the diagnosis and care of people living with HD.

The James Lind Alliance is an organization that has developed standardized methods of involving the public in setting priorities for research around specific diagnoses [52]. Their priority-setting partnerships are designed to create a top 10 list of uncertainties in a given area to prioritize future research. In 2000, 1 study described a mismatch in funded research when compared with patient priorities [53]. More recently, Crowe et al [54] compared the James Lind Alliance recommended top 10 lists with registered trials and suggested that this mismatch has not yet been remedied. Registered research mainly focuses on drug treatments (37%-86%), whereas priority-setting partnerships mention drugs as a treatment priority less than 20%

of the time [54]. We see a similar trend in this study. Nutrition and growth, perianal irritation, and gastrointestinal (GI) pain are the 2nd, 3rd, and 5th priority areas, respectively, identified by caregivers. Our literature search found only 1 study addressing nutrition and growth, 1 addressing perianal irritation, and no studies addressing GI pain in HD. Conversely, the topics that yielded the largest amount of information were diagnostics and surgical complications, which were the 3rd and 5th priorities of pediatric surgeons but not mentioned by caregivers. Thus, it is important for future research to consider both caregiver and health care provider priorities.

Our study demonstrates that it is feasible to engage with a global caregiver community online, via social media, to prioritize health research topics. We received responses from 89 caregivers; the majority responded within 1 week of posting the survey. Because HD is a rare disease, we would have been unable to recruit this large of a sample size from a single or even multicenter collaboration using more traditional methods within such a short period. However, it is prudent to acknowledge that the use of social media and an exclusively Web-based survey may have precluded the involvement of some individuals and therefore may not be entirely representative.

We undertook a number of measures to ensure that a rigorous literature review was conducted to identify studies addressing the prioritized areas, including a prespecified review methodology, peer-reviewed search strategy, and the inclusion of systematic review experts on our research team to oversee this process. A limitation is that we did not conduct a formal systematic review and as such may not have identified all of the relevant literature. Also, only English language studies were included. Furthermore, on the topic of surgical complications, owing to the large number of studies in this area, only specific study designs (systematic reviews, clinical practice guidelines, and randomized controlled trials) representing the highest level of evidence available were summarized in the evidence synthesis. Finally, within the individual topic areas, the reviewed research revealed limitations that include a large number of retrospective studies (often with small sample sizes) with a paucity of prospective research, data from large cohorts, or randomized trials. The use of nonstandardized outcome measures or clinical practice differences between sites further affected comparisons between studies.

In conclusion, involving caregivers through a parent-initiated social media campaign allowed us to identify and prioritize HD-related information needs of caregivers and surgeons. We have provided a summary of the evidence that is available to address these needs. Notably, the caregiver priority areas of nutrition and growth, perianal irritation, and GI pain have received very little attention in the literature to date. Future research should address these topics, in addition to the priority areas identified by surgeons, in collaboration with caregivers and individuals living with HD.

> *What began as a search for information resulted in establishing a support group. This has since grown into a hub for sharing the most recent medical information on HD. The HD community has partnered with the only HD not-for-profit organization: REACH (www.reachhd.org) and created a medical advisory board. This board consists of pediatric surgeons, researchers, geneticists, dieticians and gastroenterologists with a particular interest in HD. We use social media now not only to collect information but also to prioritize the information and share it so that families can have access to more than just emotional support.* [Parent of a child with HD]

Acknowledgments

Swish Productions Ltd has been responsible for the campaign strategy and public relations relating to the campaign since inception. We would like to thank REACH (Raising Education and Awareness for Children with Hirschsprung's Disease) for their affiliated support toward common goals. We would like to thank Mrs Gabrielle Deraugh for her help with preparing this manuscript for submission. The Thorlakson Chair in Surgical Research, Department of Surgery geographic full time surgeons—University of Manitoba, and Children's Hospital Research Institute of Manitoba, supported this work. RK is the Thorlakson Chair in Surgical Research for the University of Manitoba and received a New Investigator Salary Award from the Canadian Institutes of Health Research, Manitoba Lung Association, and Children's Hospital Research Institute of Manitoba. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Caregiver survey (prefaced with survey consent disclosure).

[PDF File (Adobe PDF File), 29KB - jmir_v20i12e297_app1.pdf]

Multimedia Appendix 2

Pediatric surgeon survey (prefaced with survey consent disclosure).

[PDF File (Adobe PDF File), 40KB - jmir_v20i12e297_app2.pdf]

Multimedia Appendix 3

Inclusion criteria.

[PDF File (Adobe PDF File), 24KB - jmir_v20i12e297_app3.pdf]

References

1. INVOLVE. About INVOLVE URL: http://www.invo.org.uk/about-involve/ [accessed 2018-07-18] [WebCite Cache ID 710aVJIRQ]
2. Patient-Centred Oriented Research Institute (PCORI). 2016. About Us URL: https://www.pcori.org/ [accessed 2018-07-18] [WebCite Cache ID 710aog3IZ]
3. Canadian Institutes of Health Research. 2014. Strategy for Patient-Oriented Research - Patent Engagement Framework URL: http://www.cihr-irsc.gc.ca/e/48413.html [accessed 2018-07-18] [WebCite Cache ID 710arZVaq]
4. Nass P, Levine S, Yancy C. PCORI. Washington DC; 2012. An International Perspective Methods for Involving Patients in Topic Generation for Patient-Centered Comparative Effectiveness Research URL: https://tinyurl.com/ycwqwovu [accessed 2018-12-03] [WebCite Cache ID 74OhWNqM2]
5. PCORI. 2014. PCORI Engagement Rubric URL: https://www.pcori.org/sites/default/files/Engagement-Rubric.pdf [accessed 2018-07-18] [WebCite Cache ID 710atdNH0]
6. Oliver SR, Rees RW, Clarke-Jones L, Milne R, Oakley AR, Gabbay J, et al. A multidimensional conceptual framework for analysing public involvement in health services research. Health Expect 2008 Mar;11(1):72-84 [FREE Full text] [doi: 10.1111/j.1369-7625.2007.00476.x] [Medline: 18275404]
7. Wittmeier K, Holland C, Hobbs-Murison K, Crawford E, Beauchamp C, Milne B, et al. Analysis of a parent-initiated social media campaign for Hirschsprung's disease. J Med Internet Res 2014 Dec 11;16(12):e288 [FREE Full text] [doi: 10.2196/jmir.3200] [Medline: 25499427]
8. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
9. REACH. URL: https://www.reachhd.org/ [WebCite Cache ID 710Zg4qsd]
10. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PLoS One 2011;6(6):e20476-e20476 [FREE Full text] [doi: 10.1371/journal.pone.0020476] [Medline: 21694759]
11. McGowan J, Sampson M, Salzwedel D, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration. 2016. URL: https://www.cadth.ca/press-2015-guideline-explanation-and-elaboration [accessed 2018-12-03] [WebCite Cache ID 74OiQ1snw]
12. Levitt MA, Dickie B, Peña A. The Hirschsprungs patient who is soiling after what was considered a “successful” pull-through. Semin Pediatr Surg 2012 Nov;21(4):344-353. [doi: 10.1053/j.sempedsurg.2012.07.009] [Medline: 22985840]
13. Chumpitazi BP, Nurko S. Defecation disorders in children after surgery for Hirschsprung disease. J Pediatr Gastroenterol Nutr 2011 Jul;53(1):75-79. [doi: 10.1097/MPG.0b013e318212eb53] [Medline: 21694540]
14. Langer JC. Persistent obstructive symptoms after surgery for Hirschsprung's disease: development of a diagnostic and therapeutic algorithm. J Pediatr Surg 2004 Oct;39(10):1458-1462. [Medline: 15486887]
15. Patrus B, Nasr A, Langer JC, Gerstle JT. Intrasphincteric botulinum toxin decreases the rate of hospitalization for postoperative obstructive symptoms in children with Hirschsprung disease. J Pediatr Surg 2011 Jan;46(1):184-187. [doi: 10.1016/j.jpedsurg.2010.09.089] [Medline: 21238663]
16. Sun X, Wang R, Zhang L, Li D, Li Y. Efficacy of pelvic floor muscle training for the treatment of fecal incontinence after Soave procedure for Hirschsprung disease. Eur J Pediatr Surg 2012 Aug;22(4):300-304. [doi: 10.1055/s-0032-1313351] [Medline: 22648188]
17. Peeraully M, Lopes J, Wright A, Davies B, Stewart R, Singh S, et al. Experience of the MACE procedure at a regional pediatric surgical unit: a 15-year retrospective review. Eur J Pediatr Surg 2014 Feb;24(1):113-116. [doi: 10.1055/s-0033-1357502] [Medline: 24443094]
18. Ruttenstock E, Puri P. Systematic review and meta-analysis of enterocolitis after one-stage transanal pull-through procedure for Hirschsprung's disease. Pediatr Surg Int 2010 Nov;26(11):1101-1105. [doi: 10.1007/s00383-010-2695-1] [Medline: 20711596]
19. Mc Laughlin D, Friedmacher F, Puri P. The impact of Clostridium difficile on paediatric surgical practice: a systematic review. Pediatr Surg Int 2014 Aug;30(8):853-859. [doi: 10.1007/s00383-014-3543-5] [Medline: 25008231]
20. Umeda S, Kawahara H, Yoneda A, Tazuke Y, Tani G, Ishii T, et al. Impact of cow's milk allergy on enterocolitis associated with Hirschsprung's disease. Pediatr Surg Int 2013 Nov;29(11):1159-1163. [doi: 10.1007/s00383-013-3379-4] [Medline: 23982385]

21. El-Sawaf M, Siddiqui S, Mahmoud M, Drongowski R, Teitelbaum D. Probiotic prophylaxis after pullthrough for Hirschsprung disease to reduce incidence of enterocolitis: a prospective, randomized, double-blind, placebo-controlled, multicenter trial. J Pediatr Surg 2013 Jan;48(1):111-117. [doi: 10.1016/j.jpedsurg.2012.10.028] [Medline: 23331802]
22. Aworanti O, Hung J, McDowell D, Martin I, Quinn F. Are routine dilatations necessary post pull-through surgery for Hirschsprung disease? Eur J Pediatr Surg 2013 Oct;23(5):383-388. [doi: 10.1055/s-0033-1333635] [Medline: 23444066]
23. Zhang J, Li L, Hou W, Liu S, Diao M, Zhang J, et al. Transanal rectal mucosectomy and partial internal anal sphincterectomy for Hirschsprung's disease. J Pediatr Surg 2014 May;49(5):831-834. [doi: 10.1016/j.jpedsurg.2014.02.042] [Medline: 24851780]
24. Berger S, Rufener J, Klimek P, Zachariou Z, Boillat C. Effects of potato-derived protease inhibitors on perianal dermatitis after colon resection for long-segment Hirschsprung's disease. World J Pediatr 2012 May;8(2):173-176. [doi: 10.1007/s12519-012-0356-2] [Medline: 22573429]
25. More K, Rao S, McMichael J, Minutillo C. Growth and developmental outcomes of infants with hirschsprung disease presenting in the neonatal period: a retrospective study. J Pediatr 2014 Jul;165(1):73-77.e2. [doi: 10.1016/j.jpeds.2014.02.062] [Medline: 24721468]
26. Chen JZ, Jamieson DH, Skarsgard ED. Does pre-biopsy contrast enema delay the diagnosis of long segment Hirschsprung's disease? Eur J Pediatr Surg 2010 Nov;20(6):375-378. [doi: 10.1055/s-0030-1262842] [Medline: 20954107]
27. Das K, Kini U, Babu MK, Mohanty S, D'Cruz AJ. The distal level of normally innervated bowel in long segment colonic Hirschsprung's disease. Pediatr Surg Int 2010 Jun;26(6):593-599. [doi: 10.1007/s00383-010-2603-8] [Medline: 20419377]
28. de Lorijn F, Kremer L, Reitsma J, Benninga M. Diagnostic tests in Hirschsprung disease: a systematic review. J Pediatr Gastroenterol Nutr 2006 May;42(5):496-505. [doi: 10.1097/01.mpg.0000214164.90939.92] [Medline: 16707970]
29. Flanagan S, Dietz C, Hoggard E. Evaluation of the role of fluoroscopic enema diagnosis of Hirschsprung disease, clinical considerations and determining need for radiographic evaluation. Pediatric Radiology 2011;41(1):S349. [doi: 10.1007/s00247-011-2025-3]
30. Maerzheuser S, Bassir C, Rothe K. Hirschsprung disease in the older child: diagnostic strategies. Clin Pediatr (Phila) 2012 Nov;51(11):1087-1090. [doi: 10.1177/0009922812458354] [Medline: 22935219]
31. Muller CO, Mignot C, Belarbi N, Berrebi D, Bonnard A. Does the radiographic transition zone correlate with the level of aganglionosis on the specimen in Hirschsprung's disease? Pediatr Surg Int 2012 Jun;28(6):597-601. [doi: 10.1007/s00383-012-3094-6] [Medline: 22534881]
32. Saxena AM, Sodhi K, Rao K, Vaiphei K, Khandelwal N. Role of contrast enema study in diagnosis of Hirschsprung's disease. Pediatric Radiology 2013;43(2):301. [doi: 10.1007/s00247-013-2638-9]
33. Sheng T, Wang C, Lo W, Lien R, Lai J, Chang P. Total colonic aganglionosis: Reappraisal of contrast enema study. Chinese Journal of Radiology (Taiwan) 2012;37(1):11-19 [FREE Full text]
34. Wang C, Sheng T, Lo W, Lai J. The difficulty in radiographic diagnosis of total colonic aganglionosis. Pediatric Radiology (Suppl) 2011;41:333. [doi: 10.1007/s00247-011-2025-3]
35. Wong A, Tsang D, Lam W. How Useful is Contrast Enema in the Diagnosis of Hirschsprung's Disease? Five-year Experience from a Local Referral Centre. Hong Kong J Radiol 2014 Mar 28;17(1):30-35. [doi: 10.12809/hkjr1413206]
36. Morris MI, Soglio DB, Ouimet A, Aspirot A, Patey N. A study of calretinin in Hirschsprung pathology, particularly in total colonic aganglionosis. J Pediatr Surg 2013 May;48(5):1037-1043. [doi: 10.1016/j.jpedsurg.2013.02.026] [Medline: 23701779]
37. Volpe A, Alaggio R, Midrio P, Iaria L, Gamba P. Calretinin, β-tubulin immunohistochemistry, and submucosal nerve trunks morphology in Hirschsprung disease: possible applications in clinical practice. J Pediatr Gastroenterol Nutr 2013 Dec;57(6):780-787. [doi: 10.1097/MPG.0b013e3182a934c7] [Medline: 23969533]
38. Hiradfar M, Sharifi N, Khajedaluee M, Zabolinejad N, Taraz Jamshidi S. Calretinin Immunohistochemistery: An Aid in the Diagnosis of Hirschsprung's Disease. Iran J Basic Med Sci 2012 Sep;15(5):1053-1059 [FREE Full text] [Medline: 23493854]
39. Hayes CE, Kawatu D, Mangray S, LeLeiko NS. Rectal suction biopsy to exclude the diagnosis of Hirschsprung disease. J Pediatr Gastroenterol Nutr 2012 Sep;55(3):268-271. [doi: 10.1097/MPG.0b013e31824c0acc] [Medline: 22922333]
40. Hirsch BZ, Angelides AG, Goode SP, Garb JL. Rectal biopsies obtained with jumbo biopsy forceps in the evaluation of Hirschsprung disease. J Pediatr Gastroenterol Nutr 2011 Apr;52(4):429-432. [doi: 10.1097/MPG.0b013e3181ecd644] [Medline: 21206383]
41. Vollmer DD, Fair K, Hong YA, Beaudoin CE, Pulczinski J, Ory MG. Apps seeking theories: results of a study on the use of health behavior change theories in cancer survivorship mobile apps. JMIR Mhealth Uhealth 2015;3(1):e31 [FREE Full text] [doi: 10.2196/mhealth.3861] [Medline: 25830810]
42. Chen Y, Nah S, Laksmi N, Ong C, Chua J, Jacobsen A, et al. Transanal endorectal pull-through versus transabdominal approach for Hirschsprung's disease: a systematic review and meta-analysis. J Pediatr Surg 2013 Mar;48(3):642-651. [doi: 10.1016/j.jpedsurg.2012.12.036] [Medline: 23480925]
43. Gosemann J, Friedmacher F, Ure B, Lacher M. Open versus transanal pull-through for Hirschsprung disease: a systematic review of long-term outcome. Eur J Pediatr Surg 2013 Apr;23(2):94-102. [doi: 10.1055/s-0033-1343085] [Medline: 23572464]

44. Yang L, Tang S. A prospective study of laparoscopic transanal endorectal pull-through for subtotal colectomy in Hirschsprung's disease: Anastomosis using long cuff or short cuff? Journal of Laparoendoscopic & Advanced Surgical Techniques 2013;23(12):A20. [doi: 10.1089/lap.2013.9995]
45. Romero P, Kroiss M, Chmelnik M, Königs I, Wessel LM, Holland-Cunz S. Outcome of transanal endorectal vs. transabdominal pull-through in patients with Hirschsprung's disease. Langenbecks Arch Surg 2011 Oct;396(7):1027-1033. [doi: 10.1007/s00423-011-0804-9] [Medline: 21695592]
46. Granström AL, Husberg B, Nordenskjöld A, Svensson P, Wester T. Laparoscopic-assisted pull-through for Hirschsprung's disease, a prospective repeated evaluation of functional outcome. J Pediatr Surg 2013 Dec;48(12):2536-2539. [doi: 10.1016/j.jpedsurg.2013.07.017] [Medline: 24314199]
47. Ieiri S, Nakatsuji T, Akiyoshi J, Higashi M, Hashizume M, Suita S, et al. Long-term outcomes and the quality of life of Hirschsprung disease in adolescents who have reached 18 years or older--a 47-year single-institute experience. J Pediatr Surg 2010 Dec;45(12):2398-2402. [doi: 10.1016/j.jpedsurg.2010.08.040] [Medline: 21129554]
48. Jarvi K, Laitakari E, Koivusalo A, Rintala R, Pakarinen M. Bowel function and gastrointestinal quality of life among adults operated for Hirschsprung disease during childhood: a population-based study. Ann Surg 2010;252(6):977-981.
49. CIHR. Knowledge Synthesis Grant 2015 URL: https://www.researchnet-recherchenet.ca/rnr16/vwOpprtntyDtls.do?prog=2265&view=currentOpps&type=EXACT&resultCount=25&sort=program&all=1&masterList=true [WebCite Cache ID 710bJXSUJ]
50. Arnstein S. A Ladder Of Citizen Participation. Journal of the American Institute of Planners 1969 Jul;35(4):216-224 [FREE Full text] [doi: 10.1080/01944366908977225]
51. Tritter JQ, McCallum A. The snakes and ladders of user involvement: Moving beyond Arnstein. Health Policy 2006 Apr;76(2):156-168. [doi: 10.1016/j.healthpol.2005.05.008] [Medline: 16006004]
52. James Lind Alliance. 2016. About the James Lind Alliance URL: http://www.jla.nihr.ac.uk/about-the-james-lind-alliance/) [accessed 2018-07-18] [WebCite Cache ID 710bMaWmM]
53. Tallon D, Chard J, Dieppe P. Relation between agendas of the research community and the research consumer. Lancet 2000 Jun 10;355(9220):2037-2040. [doi: 10.1016/S0140-6736(00)02351-5] [Medline: 10885355]
54. Crowe S, Fenton M, Hall M, Cowan K, Chalmers I. Patients', clinicians' and the research communities' priorities for treatment research: there is an important mismatch. Res Involv Engagem 2015 Jun 25;1(1):10. [doi: 10.1186/s40900-015-0003-x]

Abbreviations

AChE: acetylcholinesterase
CAPS: Canadian Association of Paediatric Surgeons
CE: contrast enema
GI: gastrointestinal
HAEC: Hirschsprung's associated enterocolitis
HD: Hirschsprung disease

Edited by G Eysenbach; submitted 09.01.18; peer-reviewed by A Goldstein, C Chen; comments to author 24.03.18; revised version received 16.08.18; accepted 10.09.18; published 21.12.18.

Please cite as:
Wittmeier KDM, Hobbs-Murison K, Holland C, Crawford E, Loewen H, Morris M, Lum Min S, Abou-Setta A, Keijzer R
Identifying Information Needs for Hirschsprung Disease Through Caregiver Involvement via Social Media: A Prioritization Study and Literature Review
J Med Internet Res 2018;20(12):e297
URL: http://www.jmir.org/2018/12/e297/
doi:10.2196/jmir.9701
PMID:30578208

Original Paper

The Advantages and Disadvantages of Online and Blended Therapy: Survey Study Amongst Licensed Psychotherapists in Austria

Raphael Schuster[1], MSc; Raffaela Pokorny[2], MSc; Thomas Berger[3]; Naira Topooco[4], PhD; Anton-Rupert Laireiter[1,2], Prof Dr

[1]Outpatient Center for Clinical Psychology, Psychotherapy and Health Psychology, Department of Psychology, University of Salzburg, Salzburg, Austria

[2]Faculty of Psychology, University of Vienna, Vienna, Austria

[3]Department of Clinical Psychology and Psychotherapy, University of Berne, Berne, Switzerland

[4]Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

Corresponding Author:
Raphael Schuster, MSc
Outpatient Center for Clinical Psychology, Psychotherapy and Health Psychology
Department of Psychology
University of Salzburg
Hellbrunnerstraße 34
Salzburg, 5020
Austria
Phone: 43 6644156146
Email: raphael.schuster@stud.sbg.ac.at

Abstract

Background: Web-based and blended (face-to-face plus Web-based) interventions for mental health disorders are gaining significance. However, many licensed psychotherapists still have guarded attitudes toward computer-assisted therapy, hindering dissemination efforts.

Objective: The objective of this study was to provide a therapist-oriented evaluation of Web-based and blended therapies and identify commonalities and differences in attitudes toward both formats. Furthermore, it aimed to test the impact of an information clip on expressed attitudes.

Methods: In total, 95 Austrian psychotherapists were contacted and surveyed via their listed occupational email address. An 8-minute information video was shown to half of the therapists before 19 advantages and 13 disadvantages had to be rated on a 6-point Likert scale.

Results: The sample resembled all assessed properties of Austrian psychotherapists (age, theoretical orientation, and region). Therapists did not hold a uniform overall preference. Instead, perceived advantages of both interventions were rated as neutral (t_{94}=1.89, P=.06; d=0.11), whereas Web-based interventions were associated with more disadvantages and risks (t_{94}=9.86, P<.001; d=0.81). The information clip did not excerpt any detectable effect on therapists' attitudes (r_{95}=−.109, P=.30). The application of modern technologies in the own therapeutic practice and cognitive behavioral orientation were positively related to the given ratings.

Conclusions: This study is the first to directly compare therapists' attitudes toward Web-based and blended therapies. Positive attitudes play a pivotal role in the dissemination of new technologies, but unexperienced therapists seem to lack knowledge on how to benefit from technology-aided treatments. To speed up implementation, these aspects need to be addressed in the development of new interventions. Furthermore, the preference of blended treatments over Web-based interventions seems to relate to avoidance of risks. Although this study is likely to represent therapists' attitudes in countries with less advanced electronic health services, therapists' attitudes in more advanced countries might present differently.

(J Med Internet Res 2018;20(12):e11007) doi:10.2196/11007

KEYWORDS

computer-assisted therapy; eHealth; psychotherapy; attitude of health personnel; attitude to health; mobile phone

Introduction

In recent years, the amount of research on Web-based interventions has increased exponentially, and evidence supporting the efficacy of guided Web-based interventions in treating common mental health disorders has grown substantially [1-3]. As a consequence, knowledge obtained about Web-based interventions is also being transferred into conventional psychotherapy.

Web-based interventions are usually regarded as self-guided or therapist-guided internet or mobile phone-based programs, following one or more predefined treatment paths and entailing a given number of modules or exercises to be completed [4]. Blended interventions, in turn, are integrated combinations of face-to-face therapy together with the above-mentioned Web-based or mobile phone-based programs. In blended treatment, the application of computer-supported elements is intended to optimize the therapeutic process [5,6] or increase treatment efficiency [7] or effectiveness [8,9]. Blended treatment can be provided in individual or group therapy settings [10,11].

Web-based interventions exhibit a variety of advantages, such as good accessibility, flexibility, and cost and time savings [12]. Patients can enter Web-based interventions anonymously wherever and whenever they wish, resulting in low barriers to treatment and the exploitation of new ways of treating mental health disorders [13]. With regard to achievable treatment effects, meta-analyses indicate good effectiveness for Web-based interventions [1-3]. Despite the promising results, internal and external factors seem to hinder the broad dissemination of ready-to-use programs [14]. Important internal challenges are restricted tailoring to patient needs, challenges with managing comorbidity and acute crisis [12], and low patient engagement and high dropout rates [15]. External challenges exist in the form of national legal restrictions (as in force in Austria or Germany) or stakeholders´ cautious attitudes toward Web-based therapy [16-18].

For blended therapy, the internal and external preconditions appear to be different. Given the nature of blended therapy, some of the inherent drawbacks of Web-based treatment seem to be less challenging (eg, handling of crisis or suicide risk). Additionally, external restrictions are either nonexistent (eg, national laws) or appear to be less critical (eg, stakeholder attitudes [18]). On the other hand, blended therapy is typically associated with drawbacks such as reduced scalability owing to the reliance on personal therapist time and concomitantly higher treatment costs. Furthermore, the current evidence base of blended therapy is less comprehensive compared with Web-based interventions [10]. On the patient's side, risks concerning restricted time for personal interaction [11], a potentially weakened patient-therapist bonding, or difficulties in communicating less apparent aspects of disease-related problems are of interest [19].

Regarding therapists' attitudes toward Web-based (and blended) therapy, most findings from previous studies have shown that therapist appraisals of *Web-based interventions* range from cautiously positive to generally positive [20-22]. One study (N=1532) found that therapists were more skeptical regarding Web-based therapy compared with addressed patients (η_p^2=0.38) [20]. Although partly inconsistent, several studies have identified associations between theoretical orientation (eg, psychodynamic vs others) and attitudes toward the use of Web-based interventions [20,22,23]. Furthermore, therapists' personal experience with the use of computer and media was found to positively relate to given appraisals (preference for Web-based treatment: 17.5% vs 6.4%; N=1104) [24]. Additionally, perceived applicability seems to depend on the appraisal of specific treatment features; for example, Web-based interventions were considered better suited to treat mild to moderate disorders [24]. Equally, therapist-level barriers relate to perceived disadvantages of Web-based therapy. Concerns exist with regard to potential negative effects (eg, on the working alliance) or doubtful treatment efficiency [17,25]. Literature on therapists' attitudes of *blended therapy*, however, is less extensive, and some studies have not fully differentiated between Web-based interventions and more blended forms of therapy [17,26]. Therapists frequently have reported benefits such as improvements in patients' self-management skills, improved access to therapy materials and treatment transparency, less traveling time, and possible reductions in so-called therapist drift-offs [5,19]. In a survey on the acceptability of computer-assisted therapy (N=1067) [17], professionals reported they were likely to integrate computer-supported therapies into their practice, but some doubted that the use of technology would actually improve treatment outcomes (low performance expectancy). Attitudes were also related to the general openness to new treatments (beta=.35) and computer literacy (beta=.19), and therapists varied in their ability and willingness to use computer-assisted programs. In a Delphi study (N=21), lack of nonverbal communication and the unsuitability for all patients were identified as disadvantages, and some therapists were concerned about blended therapy being time consuming or hindering to the rapport of less clear disease aspects or the establishment of a therapeutic relation [19]. Literature from neighboring disciplines, such as therapy monitoring or virtual reality, reveals comparable findings [27,28].

Psychotherapists play multiple roles in the dissemination of technology-assisted treatments [29,30] and will be end users of blended therapy; for example, Web-based interventions can be prescribed as an initial, adjunctive, or maintenance program. At the same time, psychotherapists hold important occupational and political positions in mental health systems. Therefore, it is important to improve the understanding of therapists' attitudes toward Web-based and blended therapies.

Important issues of therapists' attitudes toward Web-based and blended therapies refer to different levels of detail. For a global picture, the overall appraisal of both treatment strategies is of interest; for example, do psychotherapists hold a uniform preference for blended therapy over Web-based therapy? At a deeper level, separate rankings and comparative profiles can

depict specific advantages and disadvantages and can lead to a better understanding of each intervention's assigned strengths and weaknesses; for example, which advantages of Web-based therapy do therapists value most? At the highest level of resolution, stakeholders and developers might be interested in specific aspects of both treatments. Here the study provides an item-level analysis of both treatments; for example, do therapists believe that blending can improve current therapy practices? As a last aspect, we were interested in whether a short information video would influence therapists' appraisals.

Methods

Survey Development and Design

To explore the outlined issues, a survey study was conducted. All randomly selected subjects received an email invitation to participate in the study. The corresponding survey contained demographic information as well as items on perceived advantages and disadvantages of Web-based and blended therapies. Questions on both therapy forms were organized in separate blocks, which were presented in randomized order.

Demographic Information

Therapists reported on their educational and professional background, their years in profession (training period excluded), their basic professions (psychology, pedagogics, social work, etc), and their working region (urban vs rural). Additionally, we gathered descriptive information on self-reported computer and internet usage behavior and a ranking of blended therapy applications.

Construction of Survey Questions and Factor Analysis

Because there was no questionnaire designed to contrast the differences in the perceived (dis) advantages of Web-based and blended therapies among psychotherapists, we constructed a survey based on previous literature in the field. In the first step, an extensive literature search was conducted. In the next step, 4 previous studies with high relevance were identified [17,22,31,32] and served as the basis for this survey. In the last step, additional research was regarded during the construction of the items. The item selection was based on different criteria with a scope on 3 main categories (basic characteristics, therapeutic process, and health care perspective). Several items regarded the basic characteristics of Web-based and blended interventions (eg, treatment flexibility, age, or suitability). Further items were related to advantages and disadvantages for the therapeutic process and the therapeutic alliance (eg, repetition of therapy material, complexity of treatment, or nonverbal signals). The last category contained items assessing therapists' attitudes about occupational interests, the psychotherapy supply, and the evidence base of both treatments (eg, treatment quality, data security, or health care provision). Finally, 32 items were selected from a total of 54 candidate items. Selection criteria were redundancy and relevancy of items as well as fit for both intervention types. The selection was carried out consensually by the first, second, and last author (RS, RP, and AL, respectively). A detailed assignment of each item's theoretical background is provided in Multimedia Appendix 1.

Ratings were made on a 6-point Likert scale, and items were divided into "perceived advantages" and "perceived disadvantages" for each intervention strategy (6=I definitely agree, 5=I agree, 4=I somewhat agree and so on). With the exception of the respective intervention name, the items of both scales were identical. Both scales showed high internal consistency (18 items for advantages, Cronbach alpha=.931; 13 items for disadvantages, alpha=.930). Because factor analyses in small samples (100 individuals) can be applied, when the observed communalities were high (λ>0.6) [33], we conducted a maximum likelihood factor analysis (rotation based on the Varimax method) to roughly explore the basic factor structure of our questionnaire. The analysis revealed a single factor with high factor loadings (average λ=0.680). Here perceived advantages were related positively and disadvantages negatively to the identified factor. Detailed results of the factor analysis are listed in Multimedia Appendix 1.

Production of the Video Clip

An 8-minute video clip presented the definitions and the usual content of Web-based and blended interventions as well as their evidence base. There was no particular sequence on advantages and disadvantages of both interventions. The video clip consisted of a sequence of presentation slides depicting graphs, tables, and text-based information on unguided and guided Web-based interventions as well as on blended therapy. A professional rehearsal voice recorded the audio stream. The video did not feature any visible speaker or interview partner (eg, psychologist, professor, or patient).

Procedure

Therapists were contacted via the national register for licensed psychotherapists administered by the Federal Ministry of Health and Women in Austria, renamed and reorganized into the Federal Ministry of Labour, Social Affairs, Health and Consumer Protection (Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz) after the 2017 state elections. The register contains a comprehensive list of all licensed psychotherapists in Austria (N=8643) and is frequently used for research purposes. The entire register was downloaded, and 12.14% (1050/8643) of the addresses were selected at random. Therapists were invited to participate in the survey via email, and the survey was provided via a Web-based survey platform (LimeSurvey). The cover letter was entitled *Survey on Web-based and blended interventions in psychotherapy* and entailed information on the study background, purpose, privacy issues, and detailed contact information. Following best practice guidelines (eg, Tailored Design Method [34]), efforts were made to keep the perceived costs of responding low (eg, easy to complete), to address the relevance (eg, currency of the topic) and the benefits of participating (eg, 3×20 Euro tombola), and to establish trust by ensuring data security and a professional presentation. Additionally, we attempted to provide therapists with basic knowledge about both interventions by screening an 8-minute video clip at the beginning of the survey. Owing to the conflicting priorities of providing some information on the topic but not interacting with personal attitudes, we decided to randomly present this video clip to 50% of the surveyed

psychotherapists. Answering the survey took 23 minutes on average.

Analyses

Statistical analyses were conducted using SPSS Statistics 23 (IBM SPSS Statistics). Responses were based on mandatory field completion; thus, no missing data arose. Results were not normally distributed and nonparametrical statistics therefore would be indicated. For reasons of interpretability, we preferred to present the investigated differences in terms of Cohen *d* [35]. Therefore, obtained data were analyzed parametrically (*t* tests) and nonparametrically (Wilcoxon tests). Because the results corresponded almost perfectly, we decided to present the results based on *t* tests together with effect sizes in Cohen *d*. Differences in demographic variables between surveyed therapists and the population of Austrian therapists were analyzed using chi-square tests. Influences of demographic variables (eg, occupational computer usage or therapeutic orientation) and the impact of the presented video clip were analyzed using point-biserial correlations.

Dependent sample item-level *t* tests were applied to contrast both interventions against each other (19 positive items and 13 negative items). We decided to adjust for type l error inflation by applying Bonferroni correction to each scale separately, resulting in a critical *t* value of t=2.78 for advantages and t=2.65 for disadvantages. Results below the critical threshold are labeled in the corresponding tables. Here the average value of a given item (eg, treatment flexibility) was tested against the total value of the corresponding subscale (eg, advantages of Web-based interventions). According to power analyses (G*Power 3 [36]), the calculated power to detect a given effect size of d=0.3 and d=0.5 was beta=.83 and beta=.99, respectively.

Results

Surveyed Therapists

In response to our nationwide invitation, 95 out of 1050 contacted therapists completed the survey between May 2016 and June 2016, resulting in a response rate of 9.31%. The information clip was presented to 48% (46/95) therapists. For estimating representativeness and potential selection biases, information on therapists' theoretical orientation (eg, cognitive behavioral therapy) and other features are provided in Tables 1 and 2. Among the surveyed psychotherapists, 65% (62/95) were female, which corresponded to the population of psychotherapists in Austria (71.8%). The proportion of behavioral psychotherapists in Austria (11.9%) is traditionally lower than that in other German-speaking countries, such as Germany (35%) [37]. This was reflected in our sample (14/95, 15%). Apart from humanistic therapists, our sample seems to largely reassemble the population of Austrian therapists. Another important feature of our sample is the full range of possible professions a licensed psychotherapist in Austria may originate from. Only 44% (42/95) of the surveyed therapists were psychologists. The remaining 56% (53/95) stemmed from diverse professional areas, such as medicine, social work, etc. Survey results can benefit from this heterogeneity because many different perspectives entered the appraisal of Web-based and blended therapies.

Therapists' Computer and Internet Behavior

The vast majority of our sample used computers regularly for email correspondence and for Web-based search (Table 3). Regular email contact with clients was substantially lower, and only 12%-13% already used computers for videoconferencing or to supply modern media, videos, or book chapters to patients.

Table 1. Demographic characteristics of the sample.

Characteristics	Sample (N=95)	Population of psychotherapists in Austria (N=8643)	Statistics	
			χ^2 value	*P* value
Gender, n (%)				
Female	62 (65.26)	6205 (71.79)	2.1	.14
Male	33 (34.73)	2438 (28.21)	2.1	.14
Age in years, mean (SD)	48.7 (12.2)	N/A[a]	N/A	N/A
Theoretical orientation, n (%)				
Psychodynamic or analytic	28 (29.47)	2230 (25.80)	0.6	.44
Humanistic	26 (27.36)	3232 (37.39)	4.0	.04
Behavioral	14 (14.73)	1029 (11.90)	0.7	.39
Systemic	27 (28.42)	2152 (24.90)	0.6	.44
Region (urban/rural), %	76.8/23.2	70/30	2.1	.15

[a]N/A: not applicable.

Table 2. Professional characteristics of the sample.

Characteristics	Value
Basic profession, n (%)	
Psychology	42 (44)
Counseling	7 (7)
Medicine	5 (5)
Social work	6 (6)
Education	6 (6)
Pedagogics	6 (6)
University professor	2 (2)
Theology or philosophy	3 (3)
Nursing	2 (2)
Economy or management	4 (4)
Other	4 (4)
No specification	8 (8)
Years in profession, mean (SD)	12.4 (11.3)

Table 3. Therapists' occupational computer usage data (N=95).

Computer usage	Yes, n (%)	No, n (%)
General computer use (daily)	93 (98)	2 (2)
General email use (daily)	91 (96)	4 (4)
Conduct Web-based search	83 (87)	12 (13)
General administration tasks	77 (81)	18 (19)
Patient related documentation tasks[a]	41 (43)	54 (57)
Daily patient contact (email)[a]	45 (47)	50 (53)
Application of modern media during therapy[a]	13 (14)	82 (86)
Use of video conferencing[a]	12 (13)	83 (87)

[a]Activities that are relevant to Web-based and blended therapies.

Overall Differences in Perceived Advantages and Disadvantages

The primary aim of this survey was to depict advantages and disadvantages of each intervention strategy at the item level. Still, the overall perception of each method's (dis)advantages helps to reveal general attitudes. With scores of mean values of 3.45 and 3.61, the rating of perceived advantages can best be described as neutral (3="I somewhat disagree;" 4="I somewhat agree"). Although average perceived advantages of blended and Web-based interventions only differed tentatively with a small effect (t_{94}=1.89, P=.06; d=0.11), the appraisal of possible disadvantages differed strongly with a high effect to the detriment of Web-based interventions (t_{94}=9.86, P=.01; d=0.81).

Rankings of Advantages and Disadvantages

Tables 4, 5, 6, and 7 present rankings of the most important advantages and disadvantages separated for each intervention. For each table, the average deviation from the scale mean was calculated. Perceived advantages of Web-based interventions (Table 4) on average scored mean of 3.45 (SD 0.72). With an average of mean of 3.61 (SD 0.58), blended interventions scored slightly above this value (4="I somewhat agree"). Perceived disadvantages of Web-based interventions (Table 6) on average scored mean of 4.24 (SD 0.59). With an average of mean of 3.66 (SD 0.45), blended interventions scored significantly below this value.

Comparison Between Both Interventions

This section analyzes the most salient differences between both interventions. Table 8 presents differences in perceived advantages between Web-based and blended interventions. Besides absolute deviations of both scores, effect sizes of the deviations are also provided as a standardized indicator. Table 9 presents differences in perceived disadvantages between both interventions.

Table 4. Ranking of advantages of Web-based interventions, deviation from average (N=95).

Rank number	Advantage	Score
1	Bridging distances	4.80[a]
2	Discrete	4.36[a]
3	Timewise flexible	4.35[a]
4	Psychoeducation	3.97[b]
5	Repetition of work material	3.97[b]
6	Suitable for young patients	3.92[b]
7	Helping minorities or underserved	3.77[c]
8	Contemporary	3.76[c]
9	Bridging waiting time	3.71[c]
10	Low threshold to care	3.58
11	Web-based disinhibition effect	3.41
12	Suitable for people with age >50	3.28
13	Improve self-management	3.13[c]
14	Delivering evidence-based treatment	2.99[b]
15	Easy to share with family	2.93[b]
16	Improvement of treatment quality	2.47[a]
17	Can support therapist	2.45[a]
18	Independency from therapist	2.40[a]
19	Treatment intensification	2.33[a]
Average	N/A[d]	3.45

[a]*P*<.001 of deviation from average.

[b]*P*<.01 deviation from average.

[c]*P*<.05 deviation from average.

[d]N/A: not applicable.

Table 5. Ranking of advantages of blended interventions, deviation from average (N=95).

Rank number	Advantage	Score
1	Bridging distances	4.47[a]
2	Discrete	4.45[a]
3	Psychoeducation	4.21[b]
4	Contemporary	4.03[b]
5	Bridging waiting time	4.03[b]
6	Helping minorities or underserved	3.97[b]
7	Repetition of work material	3.95[b]
8	Suitable for young patients	3.91[c]
9	Low threshold to care	3.87[c]
10	Timewise flexible	3.75
11	Suitable for people with age >50	3.63
12	Treatment intensification	3.43
13	Improvement of treatment quality	3.38[c]
14	Delivery of evidence-based treatment	3.24[c]
15	Improve self-management	3.22[b]
16	Web-based disinhibition effect	3.04[b]
17	Easy to share with family	2.85[a]
18	Can support therapist	2.72[a]
19	Independency from therapist	2.38[a]
Average	N/A[d]	3.61

[a]*P*<.001 deviation from average.

[b]*P*<.01 deviation from average.

[c]*P*<.05 deviation from average.

[d]N/A: not applicable.

Table 6. Ranking of disadvantages of Web-based interventions, deviation from average (N=95).

Rank number	Disadvantage	Score
1	Lack of nonverbal signals	5.11[a]
2	Missing important disease aspects	4.87[a]
3	Missing problems in therapeutic process	4.83[a]
4	Not applicable for the majority	4.66[b]
5	Data security issues	4.57[c]
6	Avoidance of difficult situation	4.49[c]
7	Risk of therapy discontinuation	4.22
8	Dealing with crisis	4.18
9	Too much technology	4.03[c]
10	Might result in side effects	3.89[c]
11	Transfer into daily life	3.66[c]
12	Technology devaluates therapist's work	3.53[c]
13	More complicated than classical therapy	3.08[c]
Average	N/A[d]	4.24

[a]*P*<.001.
[b]*P*<.01.
[c]*P*<.05.
[d]N/A: not applicable.

Table 7. Ranking of disadvantages of blended interventions, deviation from average (N=95).

Rank number	Disadvantage	Score
1	Data security issues	4.4[a]
2	Lack of nonverbal signals	4.08[b]
3	Not applicable for the majority	4.02[c]
4	Missing problems in therapeutic Process	3.87[c]
5	Missing important disease aspects	3.85[c]
6	More effortful than classical therapy	3.78
7	Avoidance of difficult situation	3.77
8	Might result in side effects	3.77
9	Risk of therapy discontinuation	3.57
10	Transfer into daily life	3.43[c]
11	Too much technology	3.32[c]
12	Technology devaluates therapist's work	3.02[a]
13	Dealing with crisis	2.74[a]
Average	N/A[d]	3.66

[a]*P*<.001 deviation from average.
[b]*P*<.01 deviation from average.
[c]*P*<.05 deviation from average.
[d]N/A: not applicable.

Table 8. Comparison of advantages between Web-based and blended interventions (independent *t* tests; N=95).

Advantages	Blended interventions	Web-based interventions	Mean (SD)	Mean Cohen *d*
Treatment intensification	3.43	2.33	1.11[a] (1.14)	0.97
Improvement of treatment quality	3.38	2.47	0.91[a] (1.17)	0.77
Suitable for people with age >50 years	3.63	3.28	0.35[a] (0.78)	0.45
Bridging waiting time	4.03	3.71	0.32[b] (1.17)	0.27
Low threshold care	3.87	3.58	0.29[c] (1.18)	0.25
Contemporary	4.03	3.76	0.27[b] (0.86)	0.31
Can support the therapist	2.72	2.45	0.27[b] (0.93)	0.29
Delivering evidence-based treatments	3.24	2.99	0.25[b] (1.01)	0.25
Psychoeducation	4.21	3.97	0.24[c] (0.97)	0.25
Helping minorities or underserved	3.97	3.77	0.20 (1.06)	0.19
Improve self-management	3.22	3.13	0.09 (0.77)	0.12
Discrete	4.45	4.36	0.09 (1.27)	0.07
Suitable for young patients	3.91	3.92	−0.01 (1.05)	−0.01
Independency from therapist	2.38	2.40	−0.02 (1.19)	−0.02
Repetition of work material	3.95	3.97	−0.02 (0.85)	−0.02
Easy to share with family	2.85	2.93	−0.08 (1.02)	−0.08
Bridging distances	4.47	4.80	−0.33[b] (1.05)	−0.31
Web-based disinhibition	3.04	3.41	−0.37[b] (1.30)	−0.28
Timewise flexible	3.75	4.35	−0.60[c] (1.51)	−0.40

[a]*P*<.001.

[b]*P*<.01.

[c]*P*<.05.

Additional Findings

Additionally, we investigated the relation between demographic variables as well as the 2 variants of therapists' occupational computer usage (wide and narrow perspective) and therapist attitudes. Age (r_{95}=−.019, *P*=.85), years in profession (r_{95}=−.062, *P*=.55), gender (r_{95}=.039, *P*=.71), rural workplace (r_{95}=−.060, *P*=.57), or presentation of the short video clip (r_{95}=−.109, *P*=.30) did not relate to given appraisals but computer usage did. In the wide perspective of therapists' occupational computer usage (all 4 marked variables from Table 3), a trend toward more favorable attitudes was found (r_{95}=.177, *P*=.09). In the narrow perspective (application of modern media or videoconferencing; the last 2 items presented in Table 3), this relation became more evident (r_{95}=.241, *P*=.02). Finally, we correlated the therapeutic orientation with attitudes, and found a trend toward more positive attitudes among behavioral therapists (r_{95}=.188, *P*=.07). As the last aspect, we were interested in the perceived applicability of blended therapy elements as well as in therapists' interest in potentially applying such elements. Corresponding results are listed in Tables 10 and 11.

Table 9. Comparison of disadvantages between Web-based and blended interventions (independent t tests; N=95).

Disadvantages	Web-based interventions	Blended interventions	Mean (SD)	Mean Cohen *d*
Dealing with crisis	4.18	2.74	1.44[a] (1.17)	1.22
Lack of nonverbal signals	5.11	4.08	1.03[a] (1.16)	0.89
Missing important disease aspects	4.87	3.85	1.02[a] (1.14)	0.89
Missing problems in therapeutic process	4.83	3.87	0.96[a] (1.02)	0.94
Avoidance of difficult situation	4.49	3.77	0.72[a] (1.25)	0.58
Too technological	4.03	3.32	0.71[a] (1.38)	0.51
Risk of therapy discontinuation	4.22	3.57	0.65[a] (1.16)	0.56
Not applicable for the majority	4.66	4.02	0.64[a] (1.31)	0.49
Technology devaluates therapist's work	3.53	3.02	0.51[a] (1.39)	0.37
Transfer into daily life	3.66	3.43	0.23[b] (1.06)	0.22
Data issues	4.57	4.40	0.17 (1.13)	0.15
Might result in side effects	3.89	3.77	0.12 (1.11)	0.11
More effortful than classical therapy	3.08	3.78	−0.70[a] (1.24)	−0.56

[b]$P<.05$.

[a]$P<.001$.

Table 10. Applicability of blended therapy elements (N=95).

Applicability of elements	%
Psychoeducation	96
Record about mood and activities	85
Web-based diary	84
Exercises at home (homework)	84
Videos and multimedia (like YouTube)	78
Mediation and relaxation exercises	74
Diary on smartphone	63
Reflection of therapy elements	59
Introduction into treatment	52
Debriefing of the session	32

Table 11. Interest in blended therapy elements (N=95).

Interest in elements	%
Videos and multimedia (psychoeducation, short videos)	54
Communication (short message service text message, email, feedback about exercises)	45
E-learning (short texts, case example, Web-based exercises)	41
Smartphone or app (diary, behavioral observation, real-time-monitoring)	34
None of the components	26

Discussion

Principal Findings

This study contributes to the understanding of licensed psychotherapists' attitudes toward Web-based and blended therapies. By focusing on different levels of detail (general appraisal, internal and comparative profiles, and item-level analyses), the perceived advantages and disadvantages of both interventions are depicted. Major findings concern the neutral perception of both interventions' advantages as well as the increased perception of disadvantages of Web-based interventions. Additionally, a mismatch between therapists' concepts about both interventions and the corresponding empiric evidence can be identified at the item level of analysis. Finally, the effect of an 8-minute information video was found to be negligible.

Therapists' overall perception of advantages in Web-based and blended therapies can be described as neutral because average ratings ranged around the midpoint of the survey's scale. Although this finding does not suggest negative attitudes toward both interventions, it seems to be more in line with studies suggesting that psychotherapists are reserved and cautious in their views [20,28,29]. Even though therapists might be expected to have more positive attitudes toward technology-aided, face-to-face therapy (blended format), there was no overall preference nor was there significant difference between both intervention formats' advantages. Blended therapy pursues the frequently stated goal of unifying the advantages of traditional face-to-face and computer-supported treatments [10], and our results suggest that this relates primarily to risk-related aspects. According to the Diffusion Of Innovations theory [38] and the Unified Theory of Acceptance and Use of Technology [39], perceived advantages as well as compatibility with personal beliefs and preferences play a pivotal role in the successful dissemination of new technologies. Accordingly, the lack of perceived advantages or benefits can result in reduced interest and consequently, in the possible obstruction of dissemination efforts. In this context, several studies [17,24,40] have stressed the relevance of electronic health (eHealth) knowledge and experience as facilitators of more positive attitudes. We found a tendency toward more positive attitudes among therapists who had already used some computer or media support in their practice.

However, when comparing perceived disadvantages, the results were a bit different. Although the attributed disadvantages of blended therapy again can be described as neutral, the surveyed professionals showed particularly more negative attitudes toward the presented risks associated with Web-based interventions. This finding is in accordance with results from previous studies [20,41,42]. Some authors have ascribed the low acceptance of Web-based interventions to professionals' concerns that their work may be replaced by such technologies [12,43]. Although surveyed therapists did not per se agree with the statement that technology would devaluate a therapist's work (a minor disadvantage in Table 6), more negative attitudes toward Web-based interventions emerged when the approach was compared with blended treatment (Table 9). Still, the most salient disadvantages of Web-based therapy concerned therapeutic process aspects, such as the lack of nonverbal signals, missing important disease aspects, or dealing with crisis. Given that the relevance of these aspects differs between guided and unguided forms of Web-based therapy, further differentiation between both forms would have been advisable. In this regard, both guided and unguided forms of Web-based therapy are represented equally in this study. In the synopsis, both formats failed to elicit positive responses among psychotherapists. Additionally, risks and disadvantages seem to be particulary relevant to Web-based interventions, resulting in a more negative perception of this format. This result is in line with previous studies that reported on stakeholders' and therapists' overall preferences of blended therapy over Web-based interventions that are completely delivered via the internet [18,44] and suggests that perceived risks could play a pivotal role in this regard.

At the item level, a mismatch between empirical evidence and therapists' personal beliefs was found. Recognizing such differences can help improve training and consumer information and thus improve the dissemination of internet-based interventions; for example, the empirical base of Web-based interventions in delivering evidence-based treatments was not acknowledged by surveyed professionals (Table 4 Rank 14). The same applies to the improvement in patients' self-management abilities in blended therapy approaches—a benefit suggested in previous literature [6,19]. Finally, the increased salience of potential risks of Web-based interventions is currently not supported by evidence [45-47]. In this context, previous studies have successfully promoted positive attitudes toward eHealth in general and patient populations by providing text or video-based information [48,49]. At the same time, comparable studies yielded less successful results [50,51]. In this context, the mode of presentation (text vs video-based) and the use of persuasive methods (eg, expert evaluations or testimonials) [52] could influence the impact of the presented material. Whether such a strategy could change therapists' attitudes toward Web-based or blended treatments for now remains an open question. In this study, the randomized presentation of a short information clip did not effect therapists' attitudes. Ultimately, more profound implementation strategies appear most promising [53]. Among others, such strategies should focus on teaching, therapist trainings, incentives, and reimbursement policies.

In the light of the above-mentioned innovation theories (Diffusion Of Innovations theory and Unified Theory of Acceptance and Use of Technology [38,39]), a further strategy to improve uptake emphasizes on therapist-oriented co-design. On one hand, practitioners agreed that blended interventions constitute a contemporary and flexible approach. Simultaneously, a strikingly high number of therapists doubted that blended therapy would support them in their daily work (Table 4, rank 18 of 19). Thus, the criterion of *performance expectancy*— which was an identified key factor for (patient-based) acceptance and use in previous Web-based therapy studies [21,40]—remains unsatisfied from the therapists' perspective on blended therapy. Furthermore, therapist-based *effort expectancy* for blended therapy is very high (Table 9, last

item) because therapists do expect more workload from using blended formats. Although therapists frequently participate in the development of new interventions [54], developers should particularly emphasize how therapist-based performance and effort expectancies can be addressed in blended therapy.

Therapist attitudes were related to their personal experience in using modern technologies but not to work experience (years in profession) or other demographic variables. The relevance of personal experience has frequently been stressed in previous qualitative and quantitative studies [17,22-24,44]. Concerning the role of therapeutic orientation, our results revealed only a statistical tendency toward more positive attitudes among cognitive behavioral therapists. Thus, although our results contradict findings from several studies showing more negative attitudes among psychodynamic and humanistic therapists [20,23], they support studies identifying more liberal attitudes among behavioral therapists [22]. When interpreting these results, the small sample size, which further spreads over several different therapeutic orientations, needs to be taken into account.

Regarding the study's validity, certain factors support representativeness, whereas others restrict generalizability. Essential features of the sample resemble available population characteristics (therapeutic orientation, gender, or regionality), suggesting that the attitudes of the respective sample represent those of Austrian therapists. However, recent literature indicates critical regional differences in the knowledge about and acceptability of internet-assisted and blended interventions. Stakeholders in countries with more advanced eHealth services tend to have more positive attitudes [18] and as previously mentioned, personal experience with technology- and media-supported therapy elements relates to more positive evaluations of both treatment formats [24]. Consequently, this study seems to primarily represent therapist attitudes in surroundings with less advanced eHealth services, whereas therapists in advanced eHealth environments might hold more positive attitudes.

Limitations

This study has several limitations. First, considering that the study was carried out online and therapists were only contacted via email, selection bias may have been introduced. To counteract this tendency, it would have been advisable to use an additional paper-pencil version of the questionnaire [33]. Second, the low response rate increased the risk of introducing response bias. To estimate this risk, available population data on essential sample characteristics are provided, and corresponding deviations from the population range from around 3.1 to 10.4 percentage points. Although the sample can be considered representative for psychodynamic, behavioral, and systemic therapists (deviation=3.1%-3.2%), therapists with a humanistic orientation were underrepresented (deviation=10.4%). Third, the sample size in this study was rather small. Consequently, the study lacks sufficient power to detect small effects or subgroup effects reliably. Therefore, findings on the influence of therapeutic orientation or the relevance of personal experience in therapists' appraisals should be interpreted with caution. Fourth, many previous studies have employed standardized questionnaires [31,32]. Owing to the specific aim of this study and the lack of a corresponding pretested questionnaire, we have not been able to implement any validated survey. As a result, the translation of the survey is prone to language errors, and assumptions about its factor structure are unconfirmed. However, the reported exploratory factor analysis does indicate a single factor structure in which factor loadings of advantages and disadvantages load according to expectations. Additionally, the full translation of each survey question is provided in Multimedia Appendix 1. As the last aspect, we assessed therapists' daily personal computer usage, but we did not assess computer literacy by means of a standardized questionnaire. Applying computer literacy questionnaires might have led to additional findings.

Conclusion

This study is the first to investigate therapist attitudes toward blended therapy and to directly compare therapist appraisals of Web-based and blended intervention formats. Therapists' general attitudes can be described as neutral to cautious, and therapists' preferences of blended therapy over Web-based interventions seem to be risk-driven. According to two mentioned innovation theories, positive beliefs and preferences play a pivotal role in the successful dissemination of new technologies. As one crucial aspect, therapists seem to lack knowledge on how to benefit from technology-aided treatments. This aspect should be regarded in the development of new interventions. However, contrary to personal experience with technology- and media-supported therapy, an unspecific information video did not influence therapists' appraisals. In this context, the study provides a starting point for improved therapist education (eg, fostering knowledge on potential benefits or addressing frequent mismatches between empirical evidence and therapists' concepts). Although this study is likely to represent therapist opinions in countries with less advanced eHealth services, the small sample size restricts its sensitivity to detect small or subgroup effects.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Full translation of the questionnaire and factor loadings.

[PDF File (Adobe PDF File), 63KB - jmir_v20i12e11007_app1.pdf]

References

1. Andersson G. Internet-Delivered Psychological Treatments. Annu Rev Clin Psychol 2016;12:157-179. [doi: 10.1146/annurev-clinpsy-021815-093006] [Medline: 26652054]
2. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the Internet: a systematic review of applications, clinical efficacy and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2012 Dec;12(6):745-764. [doi: 10.1586/erp.12.67] [Medline: 23252357]
3. Păsărelu CR, Andersson G, Bergman NL, Dobrean A. Internet-delivered transdiagnostic and tailored cognitive behavioral therapy for anxiety and depression: a systematic review and meta-analysis of randomized controlled trials. Cogn Behav Ther 2017 Jan;46(1):1-28. [doi: 10.1080/16506073.2016.1231219] [Medline: 27712544]
4. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
5. Månsson KNT, Skagius RE, Gervind E, Dahlin M, Andersson G. Development and initial evaluation of an Internet-based support system for face-to-face cognitive behavior therapy: a proof of concept study. J Med Internet Res 2013 Dec 10;15(12):e280 [FREE Full text] [doi: 10.2196/jmir.3031] [Medline: 24326065]
6. Schuster R, Sigl S, Berger T, Laireiter AR. Patients' Experiences of Web- and Mobile-Assisted Group Therapy for Depression and Implications of the Group Setting: Qualitative Follow-Up Study. JMIR Ment Health 2018 Jul 11;5(3):e49 [FREE Full text] [doi: 10.2196/mental.9613] [Medline: 29997106]
7. Kleiboer A, Smit J, Bosmans J, Ruwaard J, Andersson G, Topooco N, et al. European COMPARative Effectiveness research on blended Depression treatment versus treatment-as-usual (E-COMPARED): study protocol for a randomized controlled, non-inferiority trial in eight European countries. Trials 2016 Dec 03;17(1):387 [FREE Full text] [doi: 10.1186/s13063-016-1511-1] [Medline: 27488181]
8. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: Results of a pragmatic randomized controlled trial. J Affect Disord 2018 Dec;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]
9. Zwerenz R, Becker J, Knickenberg RJ, Siepmann M, Hagen K, Beutel ME. Online Self-Help as an Add-On to Inpatient Psychotherapy: Efficacy of a New Blended Treatment Approach. Psychother Psychosom 2017 Nov;86(6):341-350. [doi: 10.1159/000481177] [Medline: 29131090]
10. Erbe D, Eichert H, Riper H, Ebert DD. Blending Face-to-Face and Internet-Based Interventions for the Treatment of Mental Disorders in Adults: Systematic Review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
11. Schuster R, Fichtenbauer I, Sparr VM, Berger T, Laireiter A. Feasibility of a blended group treatment (bGT) for major depression: uncontrolled interventional study in a university setting. BMJ Open 2018 Dec 12;8(3):e018412 [FREE Full text] [doi: 10.1136/bmjopen-2017-018412] [Medline: 29530905]
12. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
13. Wykes T, Haro JM, Belli SR, Obradors-Tarragó C, Arango C, Ayuso-Mateos JL, ROAMER consortium. Mental health research priorities for Europe. Lancet Psychiatry 2015 Nov;2(11):1036-1042. [doi: 10.1016/S2215-0366(15)00332-6] [Medline: 26404415]
14. Vis C, Kleiboer A, Prior R, Bønes E, Cavallo M, Clark SA, et al. Implementing and up-scaling evidence-based eMental health in Europe: The study protocol for the MasterMind project. Internet Interventions 2015 Nov;2(4):399-409. [doi: 10.1016/j.invent.2015.10.002]
15. 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]
16. Batterham PJ, Calear AL. Preferences for Internet-Based Mental Health Interventions in an Adult Online Sample: Findings From an Online Community Survey. JMIR Ment Health 2017 Jun 30;4(2):e26 [FREE Full text] [doi: 10.2196/mental.7722] [Medline: 28666976]
17. Becker EM, Jensen-Doss A. Computer-assisted therapies: examination of therapist-level barriers to their use. Behav Ther 2013 Dec;44(4):614-624. [doi: 10.1016/j.beth.2013.05.002] [Medline: 24094787]
18. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, et al. Attitudes towards digital treatment for depression: A European stakeholder survey. Internet Interventions 2017 Jun;8:1-9. [doi: 10.1016/j.invent.2017.01.001]
19. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer ET, van Gemert-Pijnen LJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014;14(1):355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
20. Schröder J, Berger T, Meyer B, Lutz W, Hautzinger M, Späth C, et al. Attitudes Towards Internet Interventions Among Psychotherapists and Individuals with Mild to Moderate Depression Symptoms. Cogn Ther Res 2017 Apr 22;41(5):745-756. [doi: 10.1007/s10608-017-9850-0]

21. Stallard P, Richardson T, Velleman S. Clinicians' attitudes towards the use of computerized cognitive behaviour therapy (cCBT) with children and adolescents. Behav Cogn Psychother 2010 Oct;38(5):545-560. [doi: 10.1017/S1352465810000421] [Medline: 20615273]
22. Vigerland S, Ljótsson B, Bergdahl Gustafsson F, Hagert S, Thulin U, Andersson G, et al. Attitudes towards the use of computerized cognitive behavior therapy (cCBT) with children and adolescents: A survey among Swedish mental health professionals. Internet Interventions 2014 Jul;1(3):111-117 [FREE Full text] [doi: 10.1016/j.invent.2014.06.002]
23. Wangberg SC, Gammon D, Spitznogle K. In the eyes of the beholder: exploring psychologists' attitudes towards and use of e-therapy in Norway. Cyberpsychol Behav 2007 Jun;10(3):418-423. [doi: 10.1089/cpb.2006.9937] [Medline: 17594266]
24. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. Australas Psychiatry 2011 Jun;19(3):259-264. [doi: 10.3109/10398562.2011.562295] [Medline: 21682626]
25. Cook JM, Biyanova T, Coyne JC. Barriers to adoption of new treatments: an internet study of practicing community psychotherapists. Adm Policy Ment Health 2009 Mar;36(2):83-90 [FREE Full text] [doi: 10.1007/s10488-008-0198-3] [Medline: 19104928]
26. Overholser JC. Technology-Assisted Psychotherapy (TAP): Adapting Computerized Treatments into Traditional Psychotherapy for Depression. J Contemp Psychother 2013 Jul 4;43(4):235-242. [doi: 10.1007/s10879-013-9241-0]
27. Glegg SMN, Holsti L, Velikonja D, Ansley B, Brum C, Sartor D. Factors influencing therapists' adoption of virtual reality for brain injury rehabilitation. Cyberpsychol Behav Soc Netw 2013 May;16(5):385-401. [doi: 10.1089/cyber.2013.1506] [Medline: 23713844]
28. Kaiser T, Schmutzhart L, Laireiter AR. Attitudes of Austrian Psychotherapists Towards Process and Outcome Monitoring. Adm Policy Ment Health 2018 Sep;45(5):765-779 [FREE Full text] [doi: 10.1007/s10488-018-0862-1] [Medline: 29520535]
29. Carper MM, McHugh RK, Barlow DH. The dissemination of computer-based psychological treatment: a preliminary analysis of patient and clinician perceptions. Adm Policy Ment Health 2013 Mar;40(2):87-95. [doi: 10.1007/s10488-011-0377-5] [Medline: 22001968]
30. Lovejoy TI, Demireva PD, Grayson JL, McNamara JR. Advancing the practice of online psychotherapy: An application of Rogers' diffusion of innovations theory. Psychotherapy (Chic) 2009 Mar;46(1):112-124. [doi: 10.1037/a0015153] [Medline: 22122574]
31. Moritz S, Schröder J, Meyer B, Hauschildt M. The more it is needed, the less it is wanted: attitudes toward face-to-face intervention among depressed patients undergoing online treatment. Depress Anxiety 2013 Feb;30(2):157-167. [doi: 10.1002/da.21988] [Medline: 22930656]
32. Schröder J, Sautier L, Kriston L, Berger T, Meyer B, Späth C, et al. Development of a questionnaire measuring Attitudes towards Psychological Online Interventions-the APOI. J Affect Disord 2015 Nov 15;187:136-141. [doi: 10.1016/j.jad.2015.08.044] [Medline: 26331687]
33. Henson RK, Roberts JK. Use of Exploratory Factor Analysis in Published Research. Educational and Psychological Measurement 2016 Jul 02;66(3):393-416. [doi: 10.1177/0013164405282485]
34. Dillman D, Smyth J, Christian L. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th ed. Hoboken, NJ: John Wiley & Sons; 2014:15-17.
35. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ: Erlbaum; 1988:1.
36. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146]
37. Schäfer S. Aktuelle Statistiken der KBV aus dem Blickwinkel der Psychotherapeuten. 2010. URL: http://www.sabineschaefer.com/fileadmin/user_upload/Artikel_S.Schaefer/PA-01_10__Aktuelle_Stat._KBV_Schaefer_ms.pdf
38. Rogers E. Diffusion of Innovations. 5th ed. New York: Free Press; 2003.
39. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
40. Hennemann S, Beutel ME, Zwerenz R. Drivers and Barriers to Acceptance of Web-Based Aftercare of Patients in Inpatient Routine Care: A Cross-Sectional Survey. J Med Internet Res 2016 Dec 23;18(12):e337 [FREE Full text] [doi: 10.2196/jmir.6003] [Medline: 28011445]
41. Eichenberg C, Grabmayer G. Green N: Acceptance of serious games in psychotherapy: An inquiry into the stance of therapists and patients. Telemedicine and e-Health 2016;22(11):A. [doi: 10.1089/tmj.2016.0001]
42. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
43. Caspar F. Technological developments and applications in clinical psychology and psychotherapy: introduction. J Clin Psychol 2004 Mar;60(3):221-238. [doi: 10.1002/jclp.10260] [Medline: 14981788]
44. Kivi M, Eriksson MC, Hange D, Petersson E, Björkelund C, Johansson B. Experiences and attitudes of primary care therapists in the implementation and use of internet-based treatment in Swedish primary care settings. Internet Interventions 2015 Sep;2(3):248-256. [doi: 10.1016/j.invent.2015.06.001]
45. Boettcher J, Rozental A, Andersson G, Carlbring P. Side effects in Internet-based interventions for Social Anxiety Disorder. Internet Interventions 2014 Mar;1(1):3-11. [doi: 10.1016/j.invent.2014.02.002]

46. Rozental A, Boettcher J, Andersson G, Schmidt B, Carlbring P. Negative effects of internet interventions: a qualitative content analysis of patients' experiences with treatments delivered online. Cogn Behav Ther 2015 Feb;44(3):223-236. [doi: 10.1080/16506073.2015.1008033] [Medline: 25705924]
47. Ebert DD, Donkin L, Andersson G, Andrews G, Berger T, Carlbring P, et al. Does Internet-based guided-self-help for depression cause harm? An individual participant data meta-analysis on deterioration rates and its moderators in randomized controlled trials. Psychol Med 2016 Oct;46(13):2679-2693. [doi: 10.1017/S0033291716001562] [Medline: 27649340]
48. Casey LM, Joy A, Clough BA. The impact of information on attitudes toward e-mental health services. Cyberpsychol Behav Soc Netw 2013 Aug;16(8):593-598. [doi: 10.1089/cyber.2012.0515] [Medline: 23679567]
49. Ebert DD, Berking M, Cuijpers P, Lehr D, Pörtner M, Baumeister H. Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms. A randomized controlled trial. J Affect Disord 2015 May 1;176:9-17. [doi: 10.1016/j.jad.2015.01.056] [Medline: 25682378]
50. Baumeister H, Seifferth H, Lin J, Nowoczin L, Lüking M, Ebert D. Impact of an Acceptance Facilitating Intervention on Patients' Acceptance of Internet-based Pain Interventions. The Clinical Journal of Pain 2015;31(6):528-535. [doi: 10.1097/ajp.0000000000000118]
51. Baumeister H, Nowoczin L, Lin J, Seifferth H, Seufert J, Laubner K, et al. Impact of an acceptance facilitating intervention on diabetes patients' acceptance of Internet-based interventions for depression: a randomized controlled trial. Diabetes Res Clin Pract 2014 Jul;105(1):30-39. [doi: 10.1016/j.diabres.2014.04.031] [Medline: 24862240]
52. Apolinário-Hagen J, Fritsche L, Bierhals C, Salewski C. Improving attitudes toward e-mental health services in the general population via psychoeducational information material: A randomized controlled trial. Internet Interv 2018 Jun;12:141-149 [FREE Full text] [doi: 10.1016/j.invent.2017.12.002] [Medline: 30135778]
53. Metz MJ, Veerbeek MA, Franx GC, van der Feltz-Cornelis CM, de BE, Beekman ATF. A National Quality Improvement Collaborative for the clinical use of outcome measurement in specialised mental healthcare: results from a parallel group design and a nested cluster randomised controlled trial. BJPsych Open 2017 May;3(3):106-112 [FREE Full text] [doi: 10.1192/bjpo.bp.116.004366] [Medline: 28507769]
54. Kooistra LC, Ruwaard J, Wiersma JE, van Oppen P, van der Vaart R, van Gemert-Pijnen JEWC, et al. Development and initial evaluation of blended cognitive behavioural treatment for major depression in routine specialized mental health care. Internet Interv 2016 May;4:61-71 [FREE Full text] [doi: 10.1016/j.invent.2016.01.003] [Medline: 30135791]

Abbreviations

eHealth: electronic health

Edited by G Eysenbach; submitted 09.05.18; peer-reviewed by EM Rathner, J Apolinário-Hagen, J Lin; comments to author 07.06.18; revised version received 27.06.18; accepted 28.06.18; published 18.12.18.

Please cite as:
Schuster R, Pokorny R, Berger T, Topooco N, Laireiter AR
The Advantages and Disadvantages of Online and Blended Therapy: Survey Study Amongst Licensed Psychotherapists in Austria
J Med Internet Res 2018;20(12):e11007
URL: https://www.jmir.org/2018/12/e11007/
doi:10.2196/11007
PMID:30563817

Original Paper

Exploring the Utility of Community-Generated Social Media Content for Detecting Depression: An Analytical Study on Instagram

Benjamin J Ricard[1,2], MS; Lisa A Marsch[2,3], PhD; Benjamin Crosier[1,2], PhD; Saeed Hassanpour[1,2,4,5], PhD

[1]Department of Biomedical Data Science, Dartmouth College, Lebanon, NH, United States

[2]Center for Technology and Behavioral Health, Dartmouth College, Hanover, NH, United States

[3]Department of Psychiatry, Dartmouth College, Hanover, NH, United States

[4]Department of Epidemiology, Dartmouth College, Hanover, NH, United States

[5]Department of Computer Science, Dartmouth College, Hanover, NH, United States

Corresponding Author:
Saeed Hassanpour, PhD
Department of Biomedical Data Science
Dartmouth College
One Medical Center Drive, HB 7261
Lebanon, NH, 03756
United States
Phone: 1 603 650 1983
Fax: 1 603 650 1966
Email: Saeed.Hassanpour@dartmouth.edu

Abstract

Background: The content produced by individuals on various social media platforms has been successfully used to identify mental illness, including depression. However, most of the previous work in this area has focused on *user-generated* content, that is, content created by the individual, such as an individual's posts and pictures. In this study, we explored the predictive capability of *community-generated* content, that is, the data generated by a community of friends or followers, rather than by a sole individual, to identify depression among social media users.

Objective: The objective of this research was to evaluate the utility of community-generated content on social media, such as comments on an individual's posts, to predict depression as defined by the clinically validated Patient Health Questionnaire-8 (PHQ-8) assessment questionnaire. We hypothesized that the results of this research may provide new insights into next generation of population-level mental illness risk assessment and intervention delivery.

Methods: We created a Web-based survey on a crowdsourcing platform through which participants granted access to their Instagram profiles as well as provided their responses to PHQ-8 as a reference standard for depression status. After data quality assurance and postprocessing, the study analyzed the data of 749 participants. To build our predictive model, linguistic features were extracted from Instagram post captions and comments, including multiple sentiment scores, emoji sentiment analysis results, and meta-variables such as the number of likes and average comment length. In this study, 10.4% (78/749) of the data were held out as a test set. The remaining 89.6% (671/749) of the data were used to train an elastic-net regularized linear regression model to predict PHQ-8 scores. We compared different versions of this model (ie, a model trained on only user-generated data, a model trained on only community-generated data, and a model trained on the combination of both types of data) on a test set to explore the utility of community-generated data in our predictive analysis.

Results: The 2 models, the first trained on only community-generated data (area under curve [AUC]=0.71) and the second trained on a combination of user-generated and community-generated data (AUC=0.72), had statistically significant performances for predicting depression based on the Mann-Whitney *U* test (P=.03 and P=.02, respectively). The model trained on only user-generated data (AUC=0.63; P=.11) did not achieve statistically significant results. The coefficients of the models revealed that our combined data classifier effectively amalgamated both user-generated and community-generated data and that the 2 feature sets were complementary and contained nonoverlapping information in our predictive analysis.

Conclusions: The results presented in this study indicate that leveraging community-generated data from social media, in addition to user-generated data, can be informative for predicting depression among social media users.

(J Med Internet Res 2018;20(12):e11817) doi:10.2196/11817

KEYWORDS

machine learning; depression; social media; mental health

Introduction

Background

Major depressive disorder (MDD) is estimated to be the second leading cause of disability burden worldwide and can contribute to a variety of health complications, particularly contributing to increased rates of suicide and ischemic heart disease [1]. However, many cases of MDD remain untreated due to the difficulty in identifying the disease: nonpsychiatric physicians diagnose depression in less than half of their patients with MDD, even with 5 years of follow-up care [2]. The United States Preventative Services Task Force recommends depression screening in the general adult population, particularly in pregnant and postpartum women [3]. Screening in hospital settings and medical practices may aid providers in identifying depression; however, screening methods need to be efficiently and feasibly implemented in medical care settings [4,5]. In addition to the promise of providing information about patients at risk of MDD, these methods may also be applicable to other mental illnesses, such as drug addiction [6,7]. Furthermore, social media's potential causal effects on depression may be controlled for by exploring this data source's predictive capability [8].

Social media content may be useful in expanding efforts to identify mental disorders at a population level and in facilitating the delivery of interventions to otherwise undiagnosed social media users. Designing a mental health screening methodology using social media data offers the potential to reach a broad population, including lower-income and minority individuals who may be undiagnosed and untreated for MDD, as teenagers and adults use social media in comparable levels across these socioeconomic and demographic groups [8-12]. As of May 2018, 35% of US adults use Instagram, an online social media platform for users to share pictures and video, including 64% of individuals aged 18 to 29 years [9]. Instagram posts consist of a photo or video and, optionally, a caption provided by the user and comment and likes from other users. Instagram photos have been shown to contain information relevant to depression status [10].

Objectives

Recent evidence shows the significant predictive power of social media to identify MDD, particularly in users of Twitter and Facebook [13-16]. Reece and Danforth created a model trained on signals indicative of depression in Instagram posts, such as the number of comments and the color of images, to predict depression as measured by the Center for Epidemiologic Studies Depression Scale [13]. Word-based score approaches, where each word corresponds to a specific score, have been shown to contain information regarding depression. Specifically, Affective Norms for English Words (ANEW) and Language assessment by Mechanical Turk (LabMT) were used to predict depression among Twitter users [14]. Previous work has also established the ability of Facebook status updates to predict postpartum depression as measured using the Patient Health Questionnaire (PHQ) [15]. Other studies indicate that depression in users of Twitter and Facebook could be identified through word sentiment analysis, particularly with regard to daily variation of sentiment in users, as well as through the use of metadata [16,17].

The majority of research conducted to predict MDD in social media users has focused on *user-generated* content, that is, the content created by the users themselves. This includes Twitter "Tweets," Facebook status updates, and images/videos created by the user and subsequently shared with their peers. The other information available for a given user is *community-generated* content, such as a post's "likes"/comments, friends' "wall" posts, and followers, all of which are not generated by individual users themselves but contain information on a given user and friend pair's bidirectional engagement on a social media platform. In this study, we hypothesized that word-based community-generated content contains information that can be utilized for MDD screening. We also aimed to directly test if user-generated and community-generated content contain complementary information indicative of an individual's MDD status.

Methods

Recruitment

The Clickworker crowdsourcing platform was used to recruit study participants. This platform is similar to Amazon's Mechanical Turk; however, its policy on sharing social media content on the platform was more suitable for this study. Participants' time completing surveys was compensated via monetary payment. Following consent, participants were asked to respond to survey questions, including the PHQ-8 questionnaire responses, and provide access to their Instagram profiles. Instagram profiles consist of a series of posts (an example Instagram post is shown in Multimedia Appendix 1). These posts can have captions, which are written by the user, and comments, which are mostly written by the user's followers. The Instagram application programming interface was used to automatically mine relevant features, with "/users/self," "/users/self/media/recent," and "/media/{media-id}/comments" as the end points. This data collection, the study's methodology, and the use of data in our study were approved by the Dartmouth Institutional Review Board. The research presented in this paper was conducted with participants' informed consent and complies with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects.

Patient Health Questionnaire-8 Mental Health Questionnaire

To quantify MDD in our study, PHQ-8 was completed by Instagram users. This 8-question inventory surveys the incidence

of MDD symptoms over the prior 2 weeks and was created according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition criteria for depression over the prior 2 weeks. The PHQ-8 is identical in content and scoring to the PHQ-9; however, it does not include the last question of the PHQ-9 regarding suicidal/self-injury thoughts, as previous studies have shown this question does not provide significant additional information regarding MDD risk [18]. For each symptom, the respondent is asked to identify whether they felt each symptom, for example, "Feeling down, depressed, or hopeless," and to select how often they felt each symptom: not at all (value of "0"), several days ("1"), more than half the days ("2"), and nearly every day ("3") over the past 2 weeks. The full PHQ-8 questionnaire is included in Multimedia Appendix 2. These values were added together to create a composite score between "0" and "24" for each user. A score above "10" usually signifies MDD; however, the scores between "10" and "14" have been called a "gray zone" in which some individuals are false positives for MDD, whereas a score above "15" is strongly indicative of MDD [19,20]. For the purposes of this study, we defined MDD using a PHQ-8 score cutoff of "15," following the example of a previous study that used the same cutoff to analyze postpartum depression in Facebook users [21].

Sentiment Analysis

Replicating the variable selection protocol from a previous study [14], information was extracted from the texts of post captions and comments using 3 unigram frequency–based approaches: ANEW, LabMT, and an emoji sentiment score [22-24]. These 3 methods map a unigram of a word or emoji to a word score. ANEW consists of 3 scores per word, 1 relating to valence or happiness, 1 to arousal or excitement, and 1 relating to dominance or being influenced [22]. Similarly, LabMT is a measure of happiness, mapping words to a score for valence [23]. Finally, we used an emoji sentiment scale, which maps emojis, Unicode-based emoticons, to a happiness score [24]. After calculating the mean unigram score for the caption or comment section for each post, the average and SD for each unigram score were calculated as a feature.

Data Postprocessing and Feature Extraction

In the postprocessing, individual profiles were filtered for quality and sufficient content. Following the guidelines of the CLPsych 2015 shared task, which asked participants to create methods for predicting depression in Twitter users, we restricted data from our original cohort of 2040 to individuals with at least 25 Instagram posts (removed n=755) and 75% English content (removed n=51), as measured by Google's Compact Language Detector [25]. We further filtered the data by including only individuals with complete data in our dataset (removed n=482). The final included individuals have corresponding values for all variables, except for emoji scores, where a neutral value of "0" was imputed if the variable was missing. Finally, due to the small sample size, we only included male and female responders (removed n=3), for a total sample size of 749. The characteristics of the individuals in our cohort, including their extracted features, and the text-based scores are shown in Table 1. Compared with the initial cohort of 2040 individuals, there is no significant difference of the final cohort in gender proportion measured using a binomial test and in PHQ-8 scores as measured using a *t* test. However, there is a significant difference ($P<.001$) of ages that resulted from restricting our dataset to only active Instagram users, who are generally younger than the general population [9,10].

Model Development

Figure 1 shows an overview of the machine learning methodology in this study. For all individuals, text-based features, including ANEW, LabMT, and emoji sentiment, were calculated from unigrams within texts of comments and captions to generate community-generated and user-generated features. The mean and SD of the text-based scores for the most recent *k* posts were utilized as features in our model training, with *k* as a hyperparameter tuned through cross-validation. We considered the summed PHQ-8 score as our target output and our extracted features as variables in a linear regression model, using an elastic-net regularization penalty to prevent overfitting.

To generate a test set to independently evaluate the performance of the model, 10% of the original 749 data points were randomly selected and excluded before training. For each *k* between 5 and 30, the training data were split into a 90/10 percentage training and validation set for 20 separate iterations. For each iteration, a linear regression model with an elastic-net regularization was fit to the sums of the PHQ-8 scores on the training data using the *glmnet* R package, whereas the results were evaluated on the internal validation data [26]. To find the optimal number of recent posts to use (k) and the regularization parameter (λ), the average validation area under curve (AUC) was calculated. This cross-validation found $k=20$ as the optimal value. We also used the median of the optimal λ for 20 iterations as a regularization parameter. In total, we trained 3 separate models: (1) based on only community-generated data, (2) based on only user-generated data, and (3) based on the combination of all variables from both sources. The discriminatory power of the generated models was compared on the held-out test set, using a binary indicator variable of depression as an input (ie, PHQ-8 ≥15). The AUC for all 3 models was calculated using the *ROCR*, *pROC*, and *verification* R packages [27,28].

Table 1. Our cohort characteristics and their associated features. The last column specifies which models, if any, contain the variable.

Characteristic	Statistics	Model inclusion (user/community)
Subjects (n)	749	Both
Text-based scores, mean (SD)		
Emoji sentiment, captions	0.39 (0.25)	User-based
Emoji sentiment, comments	0.47 (0.17)	Community-based
ANEW[a] valence, captions	6.55 (0.4)	User-based
SD ANEW valence, captions	1.05 (0.36)	User-based
ANEW domination, captions	5.66 (0.25)	User-based
SD ANEW domination, captions	0.66 (0.23)	User-based
ANEW arousal, captions	5.36 (0.25)	User-based
SD ANEW arousal, captions	0.65 (0.2)	User-based
LabMT[b] score, captions	5.81 (0.23)	User-based
SD LabMT score, captions	0.57 (0.21)	User-based
ANEW valence, comments	6.83 (0.55)	Community-based
SD ANEW valence, comments	0.99 (0.5)	Community-based
ANEW domination, comments	5.77 (0.32)	Community-based
SD ANEW domination, comments	0.63 (0.3)	Community-based
ANEW arousal, comments	5.51 (0.3)	Community-based
SD ANEW arousal, comments	0.59 (0.23)	Community-based
LabMT score, comments	0.62 (0.29)	Community-based
SD LabMT score, comments	5.91 (0.34)	Community-based
Metadata, mean (SD)		
Number of posts	333.55 (476.59)	Both
Number of likes	27.25 (55.46)	Both
Number of comments per post	1.63 (1.8)	Both
Number of comments, total	245.25 (616.41)	Both
Fraction of posts with no captions	0.03 (0.07)	User-based
Fraction of posts with no comments	0.48 (0.24)	Community-based
Caption length by word	12.39 (10.07)	User-based
Comment length by word	10.09 (13.21)	Community-based
Demographics		
Age (years), mean (SD)	26.7 (7.29)	Neither
Female, n (%)	515 (68.8)	Both
Male, n (%)	234 (31.2)	Both
Asian, n (%)	51 (6.8)	Neither
Black, n (%)	143 (19.1)	Neither
Hispanic/Latino, n (%)	91 (12.1)	Neither
Native American/Alaskan Native, n (%)	10 (1.3)	Neither
Native Hawaiian/Pacific Islander, n (%)	2 (0.2)	Neither
Other, n (%)	27 (3.6)	Neither
White, n (%)	425 (56.7)	Neither
Depression		

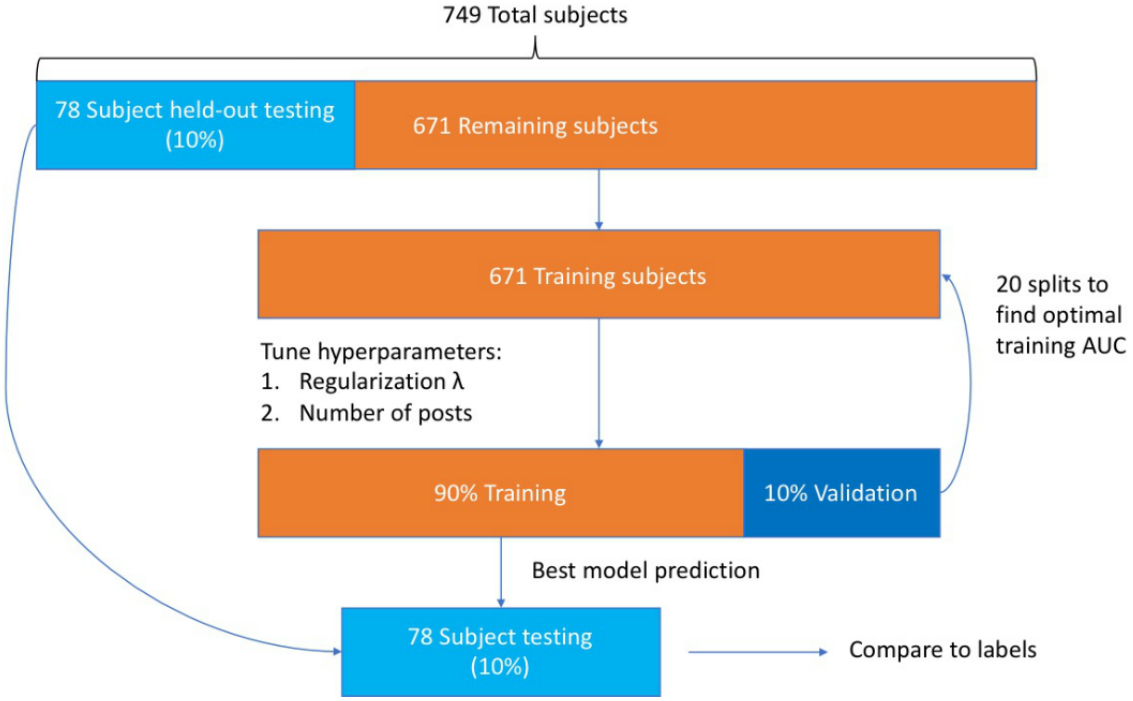
Characteristic	Statistics	Model inclusion (user/community)
PHQ-8[c] score, mean (SD)	6.62 (5.22)	Neither
PHQ-8 ≥15, n (%)	69 (9.2)	Neither

[a]ANEW: Affective Norms for English Words.

[b]LabMT: Language assessment by Mechanical Turk.

[c]PHQ-8: Patient Health Questionnaire-8.

Figure 1. Overview of our machine learning methodology. From the original 749 participating individuals, 78 (ie, 10% of the dataset) were randomly selected and held out for testing. The remaining 671 cases were used for training and parameter-tuning through cross-validation. AUC: area under curve.



Results

Outcome

The evaluation of the trained models on our held-out test set of 78 (10.4% (78/749) of the total dataset; in which 8 of the 78 had a PHQ-8 score at or above 15) indicated that community-generated data had significant predictive capacity for determining moderately severe to severe depression according to the PHQ-8 assessment (AUC=0.71; *P*=.03), whereas user-generated data were not significantly predictive (AUC=0.63; *P*=.11). When all features were combined to train a single, combined model, the model performed slightly better than the community-generated model (AUC=0.73; *P*=.02) alone, but this improvement was not statistically significant. Figure 2 shows the receiver operating characteristic curves of these 3 models on our independent test set.

Our sensitivity analysis showed that at different values of *k*, the cutoff for the most recent posts, the model based on the combination of community-generated and user-generated data still outperforms the other 2 models. To understand the composition of the model, we utilized the linear regression weights of minimum-maximum normalized variables to identify the indicative features in each model (see Figure 3 and Multimedia Appendix 3).

Outcome

Importantly, the model combining the 2 different feature sets did not simply use the community-generated or user-generated data alone. The highest corresponding weights in this combined model were features extracted both from user-generated and community-generated data; the SD of ANEW arousal caption scores, and the SD of ANEW dominance comment scores, respectively, as opposed to only using information from either dataset individually. Furthermore, the other influential variables also consisted of a combination of user-generated variables (SD of ANEW arousal caption scores, percentage of posts without captions, and SD of LabMT caption scores) and community-generated variables (SD of ANEW dominance comment scores, percentage of posts without comments, and ANEW valence comment scores).

Figure 2. Classification receiver operating characteristic curves for the predictive capability of user-generated data, community-generated data, and the combination of both to predict major depressive disorder in 78 social media users. The models that included community-generated data were significantly better than random classification, as measured with a Mann-Whitney *U* test (*P*=.03 and *P*=.02 for community-generated and combined, respectively), whereas the model trained on only user-generated data was not (*P*=.11). AUC: area under curve.

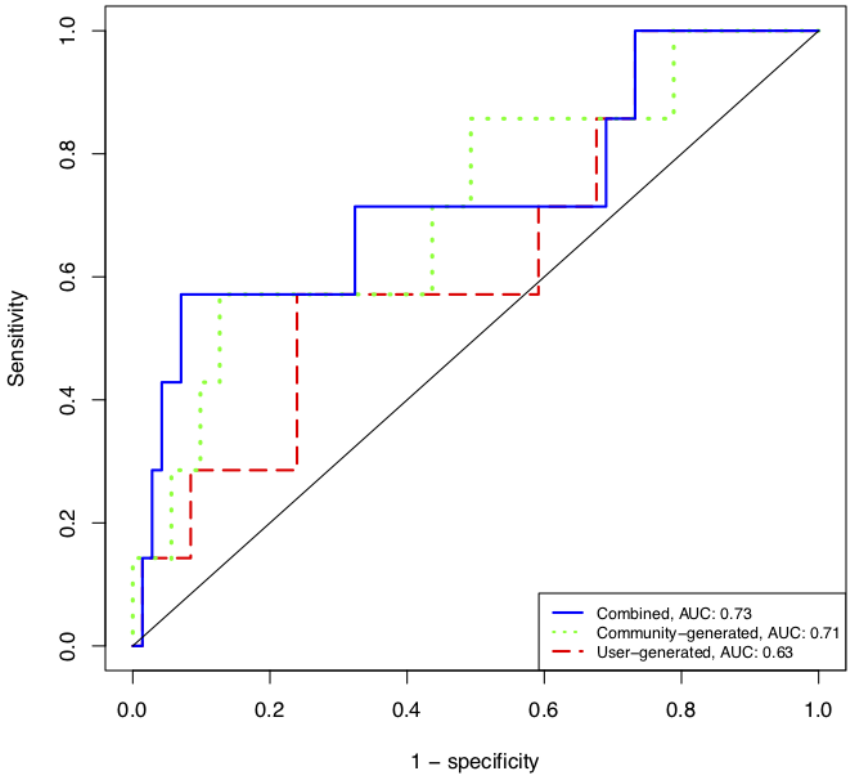
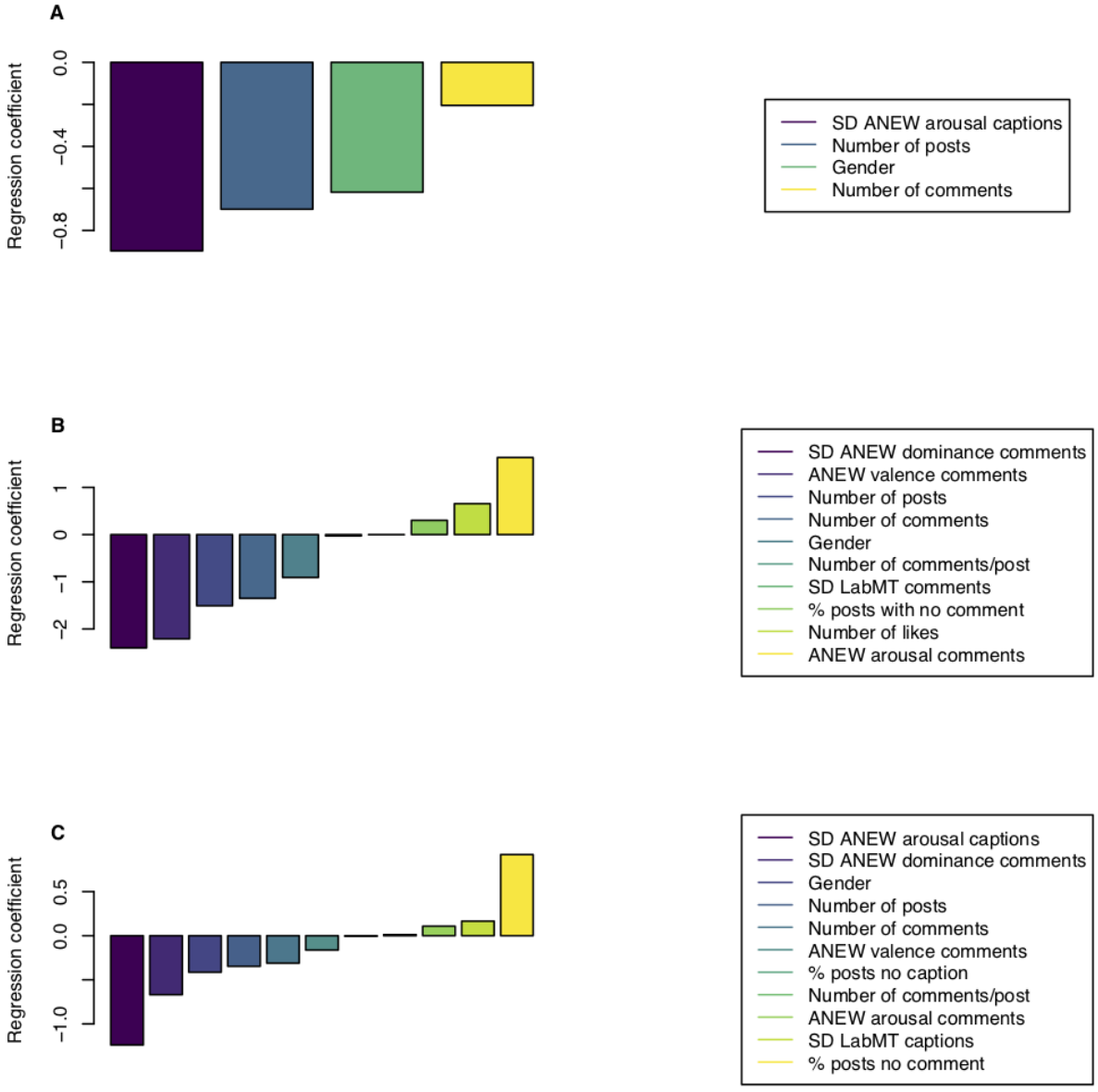


Figure 3. Minimum-maximum normalized linear regression coefficients for the model based on (A) user-generated data, (B) community-generated data, and (C) both. "Gender" variable indicates if the individual is male. These weights indicate the relative importance of each feature in the corresponding model. ANEW: Affective Norms for English Words; LabMT: Language assessment by Mechanical Turk.



To identify the influence on the number of recent posts on our models, we performed a sensitivity analysis by adjusting different numbers of recent posts *k* for each model. This analysis revealed that decreasing the number of incorporated posts resulted in user-generated data becoming more informative than comments in terms of predicting MDD. Conversely, comments became more informative than captions with the inclusion of more posts. However, the combination of both consistently outperformed either user-generated or community-generated data alone, indicating that community-generated data contain vital information on mental health status that is not captured within user-generated data alone.

Discussion

Principal Findings

This study explored the utility of community-generated social media data for identifying depression among social media users. The results indicate that community-generated social media content contains information indicative of a social media user's depression and that a model trained on the combination of user-generated and community-generated social media data outperforms models using either data source alone. Further analysis of the resulting models reveals that the indicative features from community-generated and user-generated data for this task are largely complementary and nonoverlapping.

Using Community-Generated Data Improves Detection of Depression

To the best of our knowledge, the study presented in this paper shows for the first time that information extracted from community-generated content on social media, specifically "post comments," can be used to identify mental illness in individuals with similar capacity as user-generated data. Although previous work has incorporated community-generated metadata, such as number of comments, much of the previous

research has largely been focused on understanding a user's mental well-being through information that the user posts on social media, such as Twitter "Tweets," Facebook status updates, or Instagram images [13-15,21]. Other studies also have suggested that community-generated data are correlated with user-generated data, as alcohol-related posts have more positive community-generated data [29]. Our results presented in this paper add to the body of evidence that inclusion of community-generated data may benefit analysis of social media users.

In this study, we showed that data generated from the interaction of other users with an individual carry information about a clinically validated depression assessment (PHQ-8). The model trained on community-generated data classifies individuals with a PHQ-8 score ≥15 significantly better than random classification, whereas the model trained with user-generated data did not perform as well. The model using the combination of both community-generated and user-generated datasets outperformed both; however, the improvement of this model upon the community-generated model is not statistically significant. These results indicate that future research may benefit from incorporating community-generated data, in addition to user-generated data, to understand and predict mental health in social media users.

To determine the potential for our models to be used for clinical purposes, model performance characteristics were calculated for optimal threshold values. Our model's results for detecting depression are comparable to unaided physician performance [30], demonstrating the potential for community-generated content to be implemented into screening populations for MDD. These results are not meant to be directly comparable, due to differences in population and methodology between our study and the meta-analysis performed by Mitchell et al [30]. However, the results suggest that the use of community-generated data can be beneficial for mental health screening and can be improved to the point of clinical relevance.

Community-Generated Data Contain Unique Information Not Captured in User-Generated Data

A concern regarding the combined model is the distribution of its indicative features among user-generated and community-generated variables and whether community-generated data and user-generated data provide similar information or a high degree of correlation, which would limit the utility of including community-generated data in future research. To analyze the variable distribution and their overlap among different models, the minimum-maximum normalized variable weights in each of the models were examined. This normalization allows a direct comparison of the feature coefficients within each model by rescaling all variables between "0" and "1."

These model weights indicate that for the model using both user-generated and community-generated data, the extracted features are considered informative, indicating that user-generated and community-generated data contain unique, complementary information and are nondegenerate. Furthermore, in all 3 models, there were variables that were given more weight than gender, a variable consistently shown to be attributed to different rates of depression, with women having a higher predisposition to depression [31-33]. This indicates the utility of social media–based features, both community-generated and user-generated, as an informative source for detecting depression, in addition to previously explored demographic information.

Surprisingly, the user-generated model did not perform as well as the community-generated model. A potential explanation concerns the lack of time data. The model presented was optimized to prioritize comment data over user-generated data, particularly in choosing the number *k* of recent posts to use. Sensitivity analysis indicated that captions generally performed better with fewer recent posts used and comments performed better with more recent posts used. A user may have a more variable mood through timeline given for a series of posts, and potentially, the community may provide a more stable signal over a longer period. However, the combination of features consistently outperformed either when used alone.

Comparison With Previous Work

Prediction of MDD based on social media data is well established with strong results [13-17,21,34]. However, the existing literature has largely focused on using user-generated data for this purpose, with minimal amount of community-generated data analysis. This study demonstrates that community-generated content contains information complementary to user-generated data, which can be used to predict MDD in a given user. In particular, these results suggest that community-generated text (eg, "comments") may be useful for predicting MDD, as opposed to only network/graph type features (eg, "followers") currently used in research [17].

In previous work on predicting MDD based on user-generated data, a random forest model trained on user-generated Twitter data showed promising results for predicting depression [14]. Our study had a significantly larger dataset, with 749 total individuals compared with 204, from a different social media platform (Instagram). The features incorporated in our models were partially inspired by the variables in this study, which included ANEW and LabMT scores, as well as word counts. In future work, we plan to improve the presented models through incorporating data-driven feature extraction, instead of a priori feature selection.

Table 2. Optimal cutoffs using the highest observed *F* score for user-generated, community-generated, combined, and bag-of-words models and comparison with physician rates.

Method	F_1 score	Sensitivity	Specificity
Physician (meta-analysis [30])	0.62	0.50	0.81
Baseline feature set (BOW[a])	0.58	0.50	0.69
User-generated	0.66	0.57	0.77
Community-generated	0.69	0.57	0.87
Community- and user-generated	0.70	0.57	0.92

[a]BOW: bag-of-words.

Potential for Clinical Use

A prior meta-analysis of 118 studies indicated that physicians, without the use of scales or other diagnostic tools, had an average sensitivity of 50% and a specificity of 81% for detecting depression [30]. At respective optimal cutoffs, our models had a sensitivity of 57% and specificities of 0.76, 0.86, and 0.93 for user-generated data, community-generated data, and the combination of both, respectively (Table 2). This analysis indicates the potential of community-generated data alone to diagnose moderately severe to severe depression at levels comparable to physician diagnosis. The model's performance further improves with the addition of user-generated data. It is important to note that the population and methodology used for the meta-analysis are fundamentally different from the research presented here, and interpretations between the 2 studies should be performed with caution.

To evaluate the use of a baseline feature set, a bag-of-words (BOW) model was used. Each post's caption and comment were tokenized, and English stop words were removed using the Natural Language Tool Kit library [35]. The processed captions and comments of a user's most recent 20 posts were aggregated and converted to a feature vector according to the BOW model. A regularized linear regression was trained based on these feature vectors according to the same procedure applied to the previously generated models, and its performance was compared with the previously presented models in this paper (Table 2).

The results indicate that this baseline model does not perform as well as the other presented models. This low performance can be due to the smaller sample size in this study and the simplicity and sparsity of the features in the BOW model. The features in this baseline model only rely on the frequency of words and do not capture explicit information about the word semantics and sentiment. In addition, many captions and comments are short, with an average of 12.39 and 10.09 words, respectively, in our dataset. Of note, the number of features (ie, the number of unique words) in the BOW model was 49,497 for 671 training samples, which contributed to the feature sparsity in the baseline model.

This method may also potentially be used as a cost-effective metric for the evaluation of interventions. Similar to approaches analyzing the effectiveness of other behavioral or pharmacological interventions, this method could be used as a low-cost means of patient monitoring. This is especially valuable among youth and adolescent populations, who tend to display less compliance with ecological momentary assessment reporting [36,37].

Limitations and Future Work

Due to our deanonymization protocol, time stamps, in addition to other identifiers, were removed from posts in our dataset. Therefore, we only had access to the chronological order of the posts rather than their exact time stamps. Other studies have used time series and chronologically dependent variables to understand depression in social media users [14,16]. Such analysis was not possible in this project. The PHQ-8 questionnaire represents a timeline of the previous 2 weeks; however, one of the shortcomings of utilizing the most recent *k* posts is that these *k* posts may not fully represent the posts in the last 2-week period. Future studies should incorporate time data to potentially improve outcomes. Another current limitation is that the comment section may contain some user-generated information, specifically comments generated by the user themselves. These user-generated comments could not be recognized and removed in our current dataset due to our deanonymization protocol of removing user identification. However, a significant portion of the comments is not generated by the user themselves, and most information comes from extrinsically defined sources. Finally, we acknowledge the relatively small sample size of MDD-positive individuals in our testing set (8 of 78); however, the statistical hypothesis test determining the presence of non-null significant difference in ranks between MDD-positive and -negative individuals considers sample size intrinsically in *P* values generated at the 95% confidence level. Leveraging a larger dataset with data-driven feature selection in future work can improve the training of models.

Conclusions

Social media content has been utilized previously to identify depression; however, much research to date has focused mostly on the information that individuals generate as opposed to content generated by other users, such as comments or "likes." The results presented in this paper indicate that data generated from persons who interact with posts made by other social media users contain information about the mental health of those users, specifically depression status. Furthermore, this study found that community-generated data are complementary and nonoverlapping, with respect to the content generated by the user themselves.

Acknowledgments

The authors would like to thank Lamar Moss, Timothy DeLise, Andrej Ficnar, Bruno Korbar, Cara Van Uden, Sophie Montgomery, and Madison Weatherly for their help in this project. This work was supported by a pilot grant from Dartmouth College's Office of the Provost and National Institute on Drug Abuse #P30 DA029926.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example Instagram post. Each post consists of an image or a video, with an optional caption generated by the user. Friends or followers of the user can "Like" or comment on the photo.

[PNG File, 1MB - jmir_v20i12e11817_app1.png]

Multimedia Appendix 2

Patient Health Questionnaire-8. For each question, responders are asked for the number of days they have been affected by each symptom. The numeric responses are summed for a response. In this study, a score at or above 15 is considered positive for major depressive disorder.

[PDF File (Adobe PDF File), 22KB - jmir_v20i12e11817_app2.pdf]

Multimedia Appendix 3

Coefficients of models based on user-generated, community-generated, and combined data.

[PDF File (Adobe PDF File), 28KB - jmir_v20i12e11817_app3.pdf]

References

1. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med 2013 Nov;10(11):e1001547 [FREE Full text] [doi: 10.1371/journal.pmed.1001547] [Medline: 24223526]
2. Jackson JL, Passamonti M, Kroenke K. Outcome and impact of mental disorders in primary care at 5 years. Psychosom Med 2007 Apr;69(3):270-276 [FREE Full text] [doi: 10.1097/PSY.0b013e3180314b59] [Medline: 17401055]
3. Siu AL, Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, Ebell M, 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] [Medline: 26813211]
4. IsHak WW, Collison K, Danovitch I, Shek L, Kharazi P, Kim T, et al. Screening for depression in hospitalized medical patients. J Hosp Med 2017 Feb;12(2):118-125. [doi: 10.12788/jhm.2693] [Medline: 28182810]
5. Kitchell M, Barnes R, Veith R, Okimoto J, Raskind M. Screening for depression in hospitalized geriatric medical patients. J Am Geriatr Soc 1982 Mar;30(3):174-177. [Medline: 7061789]
6. Kim SJ, Marsch LA, Hancock JT, Das AK. Scaling up research on drug abuse and addiction through social media big data. J Med Internet Res 2017 Oct 31;19(10):e353 [FREE Full text] [doi: 10.2196/jmir.6426] [Medline: 29089287]
7. 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;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
8. Escobar-Viera CG, Whitfield DL, Wessel CB, Shensa A, Sidani JE, Brown AL, et al. For better or for worse? A systematic review of the evidence on social media use and depression among lesbian, gay, and bisexual minorities. JMIR Ment Health 2018 Jul 23;5(3):e10496 [FREE Full text] [doi: 10.2196/10496] [Medline: 30037786]
9. Anderson M, Jiang J. Pew Research Center. Teens, Social Media & Technology 2018 URL: http://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/ [accessed 2018-06-14] [WebCite Cache ID 736XOoHlP]
10. Smith A, Anderson M. Pew Research Center. 2018. Social Media Use in 2018 URL: http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/ [accessed 2018-06-14] [WebCite Cache ID 71NaotcuY]
11. Williams DR, González HM, Neighbors H, Nesse R, Abelson JM, Sweetman J, et al. Prevalence and distribution of major depressive disorder in African Americans, Caribbean blacks, and non-Hispanic whites: results from the National Survey of American Life. Arch Gen Psychiatry 2007 Mar;64(3):305-315. [doi: 10.1001/archpsyc.64.3.305] [Medline: 17339519]
12. Wells AA, Palinkas LA, Shon E, Ell K. Low-income cancer patients in depression treatment: dropouts and completers. J Behav Health Serv Res 2013 Oct;40(4):427-441. [doi: 10.1007/s11414-013-9354-y] [Medline: 23868016]

13. Reece AG, Danforth CM. Instagram photos reveal predictive markers of depression. EPJ Data Sci 2017 Aug 8;6(1):15. [doi: 10.1140/epjds/s13688-017-0110-z]
14. Reece AG, Reagan AJ, Lix KL, Dodds PS, Danforth CM, Langer EJ. Forecasting the onset and course of mental illness with Twitter data. Sci Rep 2017 Oct 11;7(1):13006 [FREE Full text] [doi: 10.1038/s41598-017-12961-9] [Medline: 29021528]
15. Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. : ACM; 2013 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2013; Paris, France p. 3267-3276. [doi: 10.1145/2470654.2466447]
16. Seabrook EM, Kern ML, Fulcher BD, Rickard NS. Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. J Med Internet Res 2018 May 08;20(5):e168 [FREE Full text] [doi: 10.2196/jmir.9267] [Medline: 29739736]
17. Shen G, Jia J, Nie L, Feng F, Zhang C, Hu T, et al. Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution. 2017 Presented at: 26th International Joint Conference on Artificial Intelligence; 2017; Melborne, Australia p. 3838-3844. [doi: 10.24963/ijcai.2017/536]
18. Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen Hosp Psychiatry 2010 Jul;32(4):345-359. [doi: 10.1016/j.genhosppsych.2010.03.006] [Medline: 20633738]
19. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals 2002 Sep 01;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
20. 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]
21. Choudhury M, Counts S, Horvitz E, Hoff A. Characterizing and predicting postpartum depression from shared Facebook data. : ACM; 2014 Presented at: 17th ACM conference on Computer supported cooperative work & social computing; February 14-19. 2014; Baltimore, MD p. 626-638 URL: http://erichorvitz.com/FB-cscw2014.pdf [doi: 10.1145/2531602.2531675]
22. Bradly MM, Lang PJ. The Center for Research in Psychophysiology, University of Florida. 1999. Affective Norms for English Words (ANEW): Instruction manual and affective ratings. Technical Report C-3 URL: https://www.uvm.edu/pdodds/teaching/courses/2009-08UVM-300/docs/others/everything/bradley1999a.pdf [WebCite Cache ID 71Nb3UFXF]
23. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. Temporal patterns of happiness and information in a global social network: hedonometrics and Twitter. PLoS One 2011;6(12):e26752 [FREE Full text] [doi: 10.1371/journal.pone.0026752] [Medline: 22163266]
24. Kralj NP, Smailović J, Sluban B, Mozetič I. Sentiment of emojis. PLoS One 2015 Dec;10(12):e0144296 [FREE Full text] [doi: 10.1371/journal.pone.0144296] [Medline: 26641093]
25. Resnik P, Armstrong W, Claudino L, Nguyen T. The University of Maryland CLPsych 2015 Shared Task System. : Association for Computation Linguistics; 2015 Presented at: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, CO p. 2015. [doi: 10.3115/v1/W15-1207]
26. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. J Stat Softw 2011 Mar;39(5):1-13 [FREE Full text] [doi: 10.18637/jss.v039.i05] [Medline: 27065756]
27. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005 Oct 15;21(20):3940-3941. [doi: 10.1093/bioinformatics/bti623] [Medline: 16096348]
28. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011 Mar 17;12:77 [FREE Full text] [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
29. Hendriks H, Van den Putte B, Gebhardt WA, Moreno MA. Social drinking on social media: content analysis of the social aspects of alcohol-related posts on Facebook and Instagram. J Med Internet Res 2018 Jun 22;20(6):e226 [FREE Full text] [doi: 10.2196/jmir.9355] [Medline: 29934290]
30. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet 2009 Aug;374(9690):609-619. [doi: 10.1016/S0140-6736(09)60879-5]
31. Faravelli C, Alessandra SM, Castellini G, Lo SC. Gender differences in depression and anxiety: the role of age. Psychiatry Res 2013 Dec 30;210(3):1301-1303. [doi: 10.1016/j.psychres.2013.09.027] [Medline: 24135551]
32. Goodwin RD, Gotlib IH. Gender differences in depression: the role of personality factors. Psychiatry Res 2004 Apr 30;126(2):135-142. [doi: 10.1016/j.psychres.2003.12.024] [Medline: 15123392]
33. Weissman MM, Klerman GL. Sex differences and the epidemiology of depression. Arch Gen Psychiatry 1977 Jan;34(1):98-111. [Medline: 319772]
34. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC. Detecting depression and mental illness on social media: an integrative review. Curr Opin Behav Sci 2017 Dec;18:43-49. [doi: 10.1016/j.cobeha.2017.07.005]

35. Loper E, Bird S. NLTK: The Natural Language Toolkit. USA: Association for Computational Linguistics; 2002 Presented at: Proc ACL-02 Workshop Eff Tools Methodol Teach Nat Lang Process Comput Linguist; 2006; Stroudsburg, PA p. 63-70. [doi: 10.3115/1225403.1225421]
36. Kalf RR, Makady A, Ten HR, Meijboom K, Goettsch WG, IMI-GetReal Workpackage 1. Use of social media in the assessment of relative effectiveness: explorative review with examples from Oncology. JMIR Cancer 2018 Jun 08;4(1):e11 [FREE Full text] [doi: 10.2196/cancer.7952] [Medline: 29884607]
37. 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 Apr 26;19(4):e132. [doi: 10.2196/jmir.6641] [Medline: 28446418]

Abbreviations

ANEW: Affective Norms for English Words
AUC: area under curve
BOW: bag-of-words
LabMT: Language assessment by Mechanical Turk
MDD: major depressive disorder
PHQ-8: Patient Health Questionnaire-8

Edited by G Eysenbach; submitted 06.08.18; peer-reviewed by H Hendriks, M Conway, H Karim; comments to author 30.08.18; revised version received 12.10.18; accepted 24.10.18; published 06.12.18.

Please cite as:
Ricard BJ, Marsch LA, Crosier B, Hassanpour S
Exploring the Utility of Community-Generated Social Media Content for Detecting Depression: An Analytical Study on Instagram
J Med Internet Res 2018;20(12):e11817
URL: http://www.jmir.org/2018/12/e11817/
doi:10.2196/11817
PMID:30522991

Original Paper

eHealth Engagement as a Response to Negative Healthcare Experiences: Cross-Sectional Survey Analysis

Nicole Senft[1*], PhD; Jordan Everson[2*], PhD

[1]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States

[2]Department of Health Policy, Vanderbilt University School of Medicine, Nashville, TN, United States

[*]all authors contributed equally

Corresponding Author:
Nicole Senft, PhD
Department of Medicine
Vanderbilt University Medical Center
2525 West End Avenue
Nashville, TN,
United States
Phone: 1 6153226135
Email: nicole.senft@vumc.org

Abstract

Background: eHealth provides individuals with new means of accessing health information and communicating with providers through online channels. Prior evidence suggests that patients use eHealth to find information online when they receive care that is low in patient centeredness. However, it is unclear how other problems with the healthcare-delivery system motivate the use of eHealth, how these problems relate to different kinds of eHealth activities, and which populations are most likely to use eHealth when they receive low-quality care.

Objective: We aimed to determine how two types of negative care experiences—low patient centeredness and care coordination problems—motivate the use of different eHealth activities, and whether more highly educated individuals, who may find these tools easier to use, are more likely to use eHealth following negative experiences than less highly educated individuals.

Methods: Using nationally representative data from the 2017 Health Information National Trends Survey, we used factor analysis to group 25 different eHealth activities into categories based on the correlation between respondents' reports of their usage. Subsequently, we used multivariate negative binomial generalized linear model regressions to determine whether negative healthcare experiences predicted greater use of these resulting categories. Finally, we stratified our sample based on education level to determine whether the associations between healthcare experiences and eHealth use differed across groups.

Results: The study included 2612 individuals. Factor analysis classified the eHealth activities into two categories: provider-facing (eg, facilitating communication with providers) and independent (eg, patient-driven information seeking and communication with non-providers). Negative care experiences were not associated with provider-facing eHealth activity in the overall population (care coordination: *P*=.16; patient centeredness: *P*=.57) or among more highly educated respondents (care coordination: *P*=.73; patient centeredness: *P*=.32), but respondents with lower education levels who experienced problems with care coordination used provider-facing eHealth more often (IRR=1.40, *P*=.07). Individuals engaged in more independent eHealth activities if they experienced problems with either care coordination (IRR=1.15 *P*=.01) or patient-centered communication (IRR=1.16, *P*=.01). Although care coordination problems predicted independent eHealth activity across education levels (higher education: IRR=1.13 *P*=.01; lower education: IRR=1.19, *P*=.07), the relationship between low perceived patient centeredness and independent activity was limited to individuals with lower education levels (IRR=1.25, *P*=.02).

Conclusions: Individuals use a greater number of eHealth activities, especially activities that are independent of healthcare providers, when they experience problems with their healthcare. People with lower levels of education seem particularly inclined to use eHealth when they have negative healthcare experiences. To maximize the potential for eHealth to meet the needs of all patients, especially those who are traditionally underserved by the healthcare system, additional work should be performed to ensure that eHealth resources are accessible and usable to all members of the population.

(J Med Internet Res 2018;20(12):e11034) doi:10.2196/11034

KEYWORDS

care coordination; eHealth; health disparities; patient-centered care

Introduction

eHealth is defined as "health services and information delivered or enhanced through the Internet and related technologies" [1] and provides patients with a set of tools to engage in their health and healthcare. Along with the increase in provider use of electronic health records and associated tools over the past decade, the variety of eHealth tools available to patients has also increased [2-4]. eHealth resources may allow patients to more actively engage in their health and address problems unaddressed by their providers [5,6]. For example, patients can seek health information online that was not provided or was poorly provided by their care provider; in addition, they can use secure messaging to ensure their test results are seen by specific providers when they perceive poor coordination among their care team.

Existing evidence indicates that patients who experience deficits in the patient-provider relationship are more likely to seek health information online than those who do not experience such deficits [7-10]. For example, Li and colleagues [7] found that 40% of patients sought information online because they believed their doctor had provided them with inaccurate or incomplete information, or that the doctor's care was not as good as it should have been. In particular, patients who rated their physician as having low patient-centered communication (ie, communication that is respectful of and responsive to individual patients' needs and preferences [11]) were more likely to seek information online following their appointments, suggesting that online health information may help patients meet informational needs that are not adequately met within the patient-provider relationship. However, the focus in existing research on the association between patients' perceptions of providers and health information seeking does not clarify how dissatisfaction with care might relate to other kinds of eHealth tools. In particular, tools that provide a means to communicate with healthcare providers (eg, secure messaging) may meet different needs from tools that provide access to information and support that is relatively independent of the healthcare system (eg, health information seeking). In addition, evidence on the impact of other negative healthcare experiences beyond low patient centeredness on eHealth use is limited. No studies have thus far examined the association between care coordination and the use of eHealth. Care coordination problems reflect a system-level failure to organize patient-care activities across multiple people or organizations. Patients may perceive this problem in a different way than they perceive problems in the patient-provider relationship and may use different kinds of eHealth resources in efforts to facilitate coordination of their care.

Patients' use of the Internet for health-related reasons varies according to the individual's needs [12,13]. It is therefore likely that different kinds of problematic healthcare experiences are associated with the use of different eHealth tools. Understanding which eHealth resources can be categorized together based on their use and how healthcare experiences predict different kinds of eHealth use could allow practitioners to help their patients derive value from available technologies. Demographic differences in eHealth use indicate that groups that have traditionally been able to effectively navigate the healthcare system (ie, wealthier or more highly educated individuals) are best positioned to use eHealth resources available to them [14-18]. Consequently, these groups may be most likely to use eHealth in response to a problem in their care, thereby limiting the protective effect of broad access to eHealth tools.

This study aimed to determine the relationships between patients' experiences of problems with the healthcare system and the use of varied eHealth tools. We first classified the eHealth activities assessed in the 2017 Health Information National Trends Survey (HINTS) into related groups. We then used these factors to determine how two types of negative care experiences (ie, low patient centeredness and care coordination problems) motivated the use of different kinds of eHealth activities and whether education level affected the associations between eHealth use and negative experiences.

Methods

Data

We used data from the first wave of the 2017 HINTS, which is a cross-sectional, nationally representative survey of American adults. HINTS is designed to analyze how people use health information, with a focus on information technology and healthy behaviors. We selected the HINTS data because they contain unique information on people's interactions with the healthcare system and their use of eHealth tools.

Population

Our study included individuals aged ≥18 years in the civilian non-institutionalized population of the United States. Respondents were excluded from analyses if they were missing >25% of data in the measures of eHealth activity (n=143), patient centeredness (n=500), or care coordination (n=30). Thus, our final sample included data from 2612 respondents.

Dependent Variable: eHealth Activity

There are many available eHealth activities, and studies frequently select only one or a few activities for analysis. The HINTS survey includes 25 items related to eHealth activity across 4 instruments. Instead of limiting the activities in our analyses, we categorized these tools into conceptually similar groups. We used exploratory factor analysis to identify the number of underlying constructs onto which eHealth items loaded. These analyses, described in detail in the Analysis section below, resulted in the construction of two dependent variables: eHealth activities used to communicate with healthcare providers (11 provider-facing activities) and eHealth activities performed independent of the provider (10 independent activities).

Independent Variable: Negative Healthcare Experiences

Two sets of survey items in HINTS are related to negative healthcare experiences. *Patient centeredness* of care was measured using a 7-item scale based on the core functions of patient-centered communication identified by Epstein and Street [19] and widely used in past research [19,20]. Respondents were asked 7 questions about how often (Always, Often, Sometimes, Never) providers (1) "Give you the chance to ask all the health-related questions you had," (2) "Give the attention you needed to your feelings and emotions," (3) "Involve you in decisions about your healthcare as much as you wanted," (4) "Make sure you understood the things you needed to do to take care of your health," (5) "Explain things in a way you could understand," (6) "Spend enough time with you," and (7) "Help you deal with feelings of uncertainty about your health or health care." To create a summary of responses, we calculated the mean of each respondents' answers to all 7 questions. As the responses were highly skewed, we operationalized this variable as tertiles rather than as a continuous measure. The tertiles represented relatively low, medium, and high patient centeredness. If the patient reported low or medium patient centeredness, they were considered to have *negative experiences of patient centeredness*. This conceptual categorization of perceived patient centeredness into very-positive versus less-positive perceptions is consistent with the methods of previous studies that used this measure and other measures of patient-centered communication [7,20,21].

The second set of survey items focused on *problems in care coordination*. Four survey items on this concept were included in HINTS. Respondents were asked whether at some point in the last 12 months, they (1) "Had to bring an X-ray, MRI, or other type of test result with you to the appointment," (2) "Had to wait for test results longer than you thought reasonable," (3) "Had to redo a test or procedure because the earlier test results were not available," and (4) "Had to provide your medical history again because your chart could not be found." We excluded the survey item about bringing a test result to an appointment because we believed it lacked face validity; unlike the other 3 items, this item was not considered problematic. Therefore, we were concerned that any patient who underwent imaging might answer this question positively. In support of this reasoning, during initial data cleaning, we empirically observed that this "problem" was reported far more often (572/2612, 22% unweighted, 19% weighted) vs an average of 225/2612 (8.6% unweighted, 9.4% weighted) respondents for the other 3 problems) and that correlations between this "problem" and the other 3 problems were low (0.17 on average). Because each item was relatively rare in the initial data analysis, we operationalized this variable dichotomously. If the patient experienced at least 1 of the 3 problems, they were considered to have *negative experiences of care coordination.*

Stratifying Variables

In the HINTS data, education is measured on a 5-point scale: (1) Less than high school, (2) High school graduation, (3) Some college, (4) Bachelor's degree, and (5) Postbaccalaureate degree. We stratified the sample into higher and lower education levels to determine whether these groups engaged in eHealth differently when they had negative healthcare experiences. The median level of education was some college experience, and more respondents reported a college or higher education level than a high school or lower education level (30.2% vs 36.1%). Therefore, we categorized participants with education level lower than a bachelor's degree as having a lower education level and those with a bachelor's or postbaccalaureate degree as having a higher education level. A subsequent sensitivity analysis grouped participants with at least some college experience along with participants with a higher education level.

Control Variables

We included several variables in our multivariate analysis to account for factors that may introduce bias in the relationships between negative healthcare experiences and the use of eHealth. We included 4 demographic variables (race, gender, age, and income), two variables related to use of the internet (whether they ever use the internet and whether they accessed the internet from home), and patients' self-reported general health, each of which may be associated with both the extent to which individuals experience problems with their healthcare and their use of eHealth resources.

Analysis

Factor Analysis

To test the first research question, that is, how eHealth activities can be categorized on the basis of their usage, we used exploratory factor analysis with oblique promax rotation. We chose this rotation because we did not want to constrain the data based on an assumption of orthogonality. We retained factors with an eigenvalue >1, which is the typical cutoff to retain factors for analysis. Items were excluded if, following rotation, they did not load onto any factor at levels >0.40 or if they loaded onto multiple factors at levels >0.40.

Regression Analysis

We created two multivariate generalized linear model regressions for our second research question about the relationship between negative healthcare experiences and eHealth activities. In one model, we estimated how negative healthcare experiences (medium or low perceptions of patient centeredness and experience of at least one coordination problem) were associated with the use of provider-facing eHealth activities. In the second model, we analyzed the associations between these two negative healthcare experiences and independent eHealth activities. We included covariates related to demographics, internet use, and general health in each model. We used negative binomial regressions because the outcomes were counts of eHealth activities and were overdispersed. We used survey weights to ensure that our estimates were representative of the US population.

We divided our sample into two groups according to higher and lower education levels to address our third research question about whether the relationships between negative healthcare experiences and eHealth activity differed across educational levels. Subsequently, we recreated the two negative binomial regression models described above in this section to estimate

the relationships between negative healthcare experiences and provider-facing and independent eHealth activities within each education group. Finally, we plotted the predicted level of each eHealth activity based on negative healthcare experiences to facilitate comparison of the magnitude of effects. All statistical analyses were conducted in Stata 16 MP (Stata Corporation, College Station, TX).

Results

Summary Statistics

Our final study sample included 2612 individuals (Table 1), of which 63.5% (survey-weighted) were non-Hispanic white, 9.2% were Hispanic, and 13.1% were non-Hispanic African-American. In addition, 62.6% of respondents (weighted) did not receive a bachelor's or higher degree. The mean age of the study population was 49 years, and the modal health rating was "Very good."

Factor Analysis

Only two factors had eigenvalues >1. Following oblique promax rotation, 21 of the 25 eHealth activity items loaded clearly onto one of the two factors, which together accounted for 93% of the variance in reported eHealth use. The first factor included 11 eHealth activities used to communicate with healthcare providers (provider-facing activities). The second factor included 10 eHealth activities independent of the provider (independent activities; Table 2). The remaining 4 items were excluded from analyses because they failed to load onto either factor using the factor-loading cutoff of .40. Use of provider-facing activities and independent activities were positively correlated with each other (r [2,612]=.48). A mean comparison using t-test showed that respondents had used fewer provider-facing activities (mean 1.94, SD 2.70) than independent activities (mean 3.75, SD 2.71; t [2611]=33.42, P<.001, 95% CI 1.70-1.92).

Regression Analysis

Full Sample

Overall, neither problems in care coordination nor perceived patient centeredness predicted the number of provider-facing activities used (Table 3). In contrast, participants who experienced problems with care coordination used an average of 0.50 (14.9%) more independent eHealth activities than those who did not experience such problems (beta=1.15, P=.01). Compared to participants who perceived high levels of patient centeredness, those who perceived moderate levels of patient centeredness used an average of 0.44 (14.0%) more independent activities (beta=1.14, P=.02) and those who perceived low levels used an average of 0.50 (15.9%) more independent activities (beta=1.16, P=.01).

Education-Stratified Groups

Provider-facing eHealth activity was not predicted by problems in care coordination or perceived patient centeredness in the model restricted to more highly educated adults (Figure 1). Among individuals with education below college level, those who experienced problems with care coordination used an average of 0.40 (40.4%) more provider-facing eHealth activities than those who did not experience such problems (beta=1.40, P=.07). However, this finding should be interpreted with caution, as it was not significant in our sensitivity analysis (ie, when participants with some college education were categorized as having higher education levels, Multimedia Appendix 1). The perceived lack of patient centeredness remained nonsignificantly associated with provider-facing eHealth use among adults with lower levels of education.

In the stratified model restricted to more highly educated individuals, problems with care coordination were associated with the use of 0.63 (13.1%) more independent eHealth activities (beta=1.13, P=.009), whereas perceived patient centeredness was not associated with the use of these activities. Among individuals with education below college level, those who experienced problems with care coordination used an average of 0.51 (18.8%) more provider-facing eHealth activities than those who did not experience such problems, showing a marginally significant increase (beta=1.19, P=.07). This finding should also be interpreted with caution, as it was nonsignificant in our sensitivity analysis (Multimedia Appendix 1). Compared to participants who perceived high levels of patient centeredness, those who perceived moderate levels of patient centeredness used an average of 0.55 (22.4%) more independent activities (beta=1.22, P=.02) and those who perceived low levels used an average of 0.62 (25.4%) more independent activities (beta=1.25, P=.02).

Table 1. Survey-weighted summary characteristics of the 2017 Health Information National Trends Survey respondents in current analyses.

Variable	n (%)	95% CI
Education		
Less than high school	148 (7.5)	5.5-9.5
High school graduate	467 (22.7)	20.3-25.1
Some college	752 (32.4)	30.0-34.7
Bachelor's degree	697 (22.2)	20.7-23.6
Postbaccalaureate degree	491 (13.9)	12.6-15.3
Race/ethnicity		
Non-Hispanic white	1568 (63.5)	61.3-65.8
Hispanic	333 (9.2)	8.1-10.3
Non-Hispanic African-American	296 (13.1)	11.6-14.5
Non-Hispanic Asian	95 (4.4)	3.7-5.1
Other	98 (2.8)	2.4-3.2
Gender		
Male	996 (45.6)	43.8-47.4
Female	1580 (53.5)	51.7-55.3
Age		
18-34 years	280 (21.3)	18.0-24.5
35-49 years	514 (26.7)	23.3-30.2
50-64 years	877 (30.0)	28.0-32.1
65-74 years	563 (11.5)	10.8-12.1
≥75 years	292 (7.8)	7.2-8.5
Income (USD)		
$0-$34,999	805 (28.6)	25.7-31.4
$35,000-$100,000	1125 (44.5)	41.0-48.1
≥$100,000	660 (25.8)	23.1-28.4
General health		
Poor	64 (2.4)	1.4-3.4
Fair	410 (14.5)	12.3-16.7
Good	903 (34.1)	30.5-37.6
Very good	944 (37.3)	33.9-40.8
Excellent	266 (10.8)	8.3-13.4
Use Internet		
No	499 (16.4)	14.5-18.2
Yes	2113 (83.6)	81.8-85.5
Use Internet at home		
Not applicable	100 (4.4)	3.0-5.9
Never	482 (16.4)	14.4-18.3
Sometimes	627 (25.2)	22.6-27.7
Daily	1260 (49.8)	46.7-52.9

Table 2. Factor analysis results.

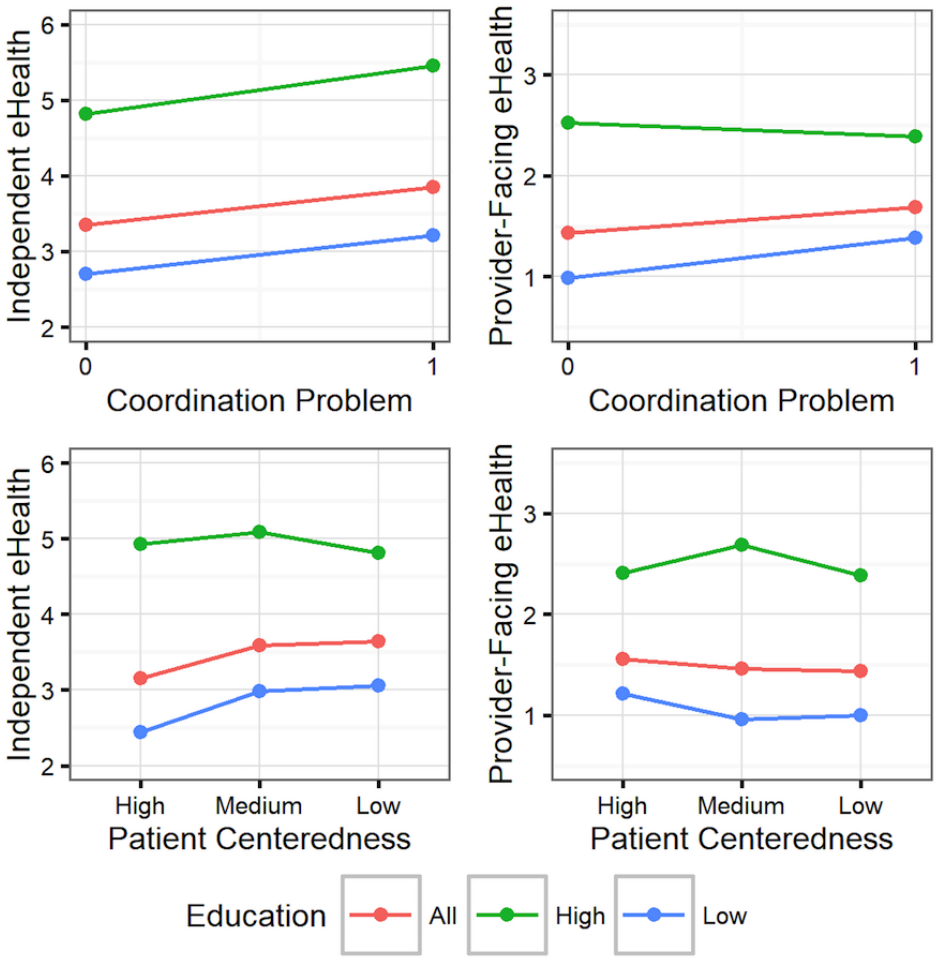
Item	Factor loadings	
	Provider-Facing eHealth	Independent eHealth
In the past 12 months, have you used a computer, smartphone, or any other electronic means to do any of the following?		
Look for health or medical information for yourself	–0.04	.62
Look for health or medical information for someone else	–0.06	.58
Buy medicine or vitamins online	.08	.28
Look for a healthcare provider	–0.04	.55
Use e-mail or the Internet to communicate with a doctor or a doctor's office	.44	.31
Make appointments with a health care provider	.29	.31
Track healthcare charges and costs	.20	.43
Fill out forms or paperwork related to your health care	.16	.44
Look up test results	.56	.23
In the past 12 months, have you used your online medical record to		
Make appointments with a healthcare provider	.74	–0.02
Fill out forms or paperwork related to your healthcare	.62	.07
Request refill of medications	.72	–0.07
Request correction of inaccurate information	.39	–0.03
Add health information to share with your healthcare provider, such as health concerns, symptoms, and side-effects	.60	–0.06
Download your health information to your computer or mobile device such as a cell phone or tablet	.48	–0.01
Help you make a decision about how to treat an illness or condition	.57	–0.03
Securely message a health care provider and staff (eg, e-mail)	.80	–0.04
Monitor your health	.66	–0.05
Look up test results	.76	.03
In the last 12 months, have you used the Internet for any of the following reasons?		
To share health information on social networking sites such as Facebook or Twitter	–0.07	.41
To participate in an online forum or support group for people with a similar health or medical issue	–0.05	.35
To watch a health-related video on YouTube	–0.06	.52
Has your tablet or smartphone		
Helped you track progress on a health-related goal such as quitting smoking, losing weight, or increasing physical activity	.05	.46
Helped you make a decision about how to treat an illness or condition	–0.07	.58
Helped you in discussions with your healthcare provider	.09	.51

Table 3. Weighted negative binomial regression predicting provider-facing and independent eHealth activities.

Variable	Provider-facing eHealth		Independent eHealth	
	Incidence rate ratios (SE)	*P* value	Incidence rate ratios (SE)	*P* value
Care coordination problem	1.18 (0.14)	.16	1.15 (0.06)	.01
Patient centeredness (reference: High)				
Medium	0.94 (0.10)	.58	1.14 (0.06)	.01
Low	0.92 (0.13)	.57	1.16 (0.07)	.01
Race/ethnicity (reference: Non-Hispanic white)				
Hispanic	0.98 (0.16)	.90	1.03 (0.08)	.67
Non-Hispanic African-American	0.84 (0.14)	.29	1.01 (0.05)	.91
Non-Hispanic Asian	1.14 (0.27)	.58	1.23 (0.09)	.01
Other	1.01 (0.23)	.95	1.15 (0.13)	.22
Missing	1.18 (0.31)	.53	1.24 (0.27)	.33
Gender (reference: Male)				
Female	1.22 (0.15)	.12	1.16 (0.06)	.004
Missing	0.92 (0.38)	.85	1.31 (0.29)	.22
Age (reference: 18-34 years)				
35-49 years	1.32 (0.22)	.09	0.95 (0.07)	.47
50-64 years	1.17 (0.19)	.35	0.77 (0.06)	.001
65-74 years	1.06 (0.18)	.74	0.65 (0.06)	<.001
≥75 years	0.98 (0.22)	.91	0.43 (0.05)	<.001
Missing	1.01 (0.43)	.98	0.46 (0.14)	.01
Education (reference: Less than high school graduate)				
High school graduate	1.31 (0.39)	.36	1.12 (0.23)	.57
Some college	1.72 (0.55)	.09	1.31 (0.27)	.20
Bachelor’s degree	2.28 (0.73)	.01	1.44 (0.29)	.08
Postbaccalaureate Degree	2.14 (0.71)	.03	1.38 (0.28)	.12
Missing	0.85 (0.41)	.74	1.01 (0.41)	.98
Income (USD; reference: $0-$34,9999)				
$35000-$99,9999	1.39 (0.31)	.14	1.14 (0.06)	.02
≥$100,000	1.59 (0.37)	.05	1.21 (0.08)	.009
Missing	0.93 (1.01)	.94	0.96 (0.54)	.94
General health (reference: Poor)				
Fair	0.98 (0.57)	.97	1.29 (0.31)	.28
Good	0.78 (0.46)	.68	1.30 (0.33)	.31
Very good	0.92 (0.54)	.89	1.34 (0.33)	.23
Excellent	0.81 (0.47)	.72	1.37 (0.34)	.21
Missing	0.36 (0.19)	.06	1.35 (0.44)	.37
Use Internet	2.70 (0.87)	<.001	2.41 (0.39)	<.001
Use Internet at home (reference: Not Applicable)				
Never	2.15 (1.01)	.11	1.27 (0.22)	.17
Sometimes	2.29 (0.84)	.03	1.48 (0.17)	.001
Daily	2.59 (0.92)	.01	1.55 (0.17)	<.001
Missing	1.44 (0.72)	.47	1.38 (0.24)	.07

Variable	Provider-facing eHealth		Independent eHealth	
	Incidence rate ratios (SE)	*P* value	Incidence rate ratios (SE)	*P* value
Constant	0.11 (0.10)	.11	0.71 (0.22)	.15

Figure 1. Associations between healthcare experiences and eHealth use, stratified by education level.



Discussion

We examined the relationships between negative experiences with the healthcare system and the use of eHealth, and whether these relationships differed across individuals' education levels in a nationally representative sample of adults. We found that eHealth activities were clearly divided into two categories: provider-facing activities (facilitating access to providers and communication with providers) and independent activities (patient-driven information seeking and communication with non-providers). Negative care experiences were not associated with provider-facing eHealth activity in the overall population or among more highly educated respondents; however, respondents with a lower education level were more likely to use these activities if they experienced problems with care coordination. These results were different for independent eHealth activity: Overall, individuals were more likely to engage in these activities if they experienced problems with either care coordination or patient-centered communication. Although care coordination problems predicted independent eHealth activity similarly across education levels, the relationship between low perceived patient centeredness and independent activity seemed limited to individuals with lower levels of education. The cross-sectional nature of these data precludes us from determining whether eHealth use results from these negative care experiences; however, our findings suggest that people may use eHealth to address deficiencies in healthcare, and this potential protective effect is more pronounced in groups that have traditionally struggled to navigate the healthcare system (individuals with lower levels of education).

The two underlying categories we identified using factor analysis resonate with existing literature on eHealth, which tend to focus on provider-facing eHealth tools or independent health information seeking, but rarely on both [6]. Existing individual analyses focused on one or a few eHealth activities, making the comparison of results across studies difficult (eg, [14] vs [16]). The structure we identified provides a framework for determining how the use of one kind of activity might affect the use of other activities. In addition, our approach facilitates measurement of multiple kinds of eHealth activities concurrently and limits potentially arbitrary selection of eHealth activities for analysis. Finally, scale construction facilitates investigation

of the intensity of eHealth use. As internet access and advanced electronic health records become increasingly widespread, binary indicators of eHealth use may become less meaningful and measures of intensity may become more important [22].

Mean levels of eHealth activity indicate that these resources remain underused, with people using, on average, less than half of the eHealth resources available to them. Provider-facing activities are especially underused, indicating the need for researchers and healthcare professionals to identify and remedy barriers to their adoption and use. Although eHealth activities are overall underused, the current results suggest greater use among individuals with negative healthcare experiences than in those without such experiences. Similar to previous studies, we found that perceived low patient centeredness predicted increased independent activity [7,10]. Extending this prior work, we observed that care-coordination problems were associated with greater independent eHealth use. In contrast to these effects for independent use, neither low patient centeredness nor care coordination problems contributed to provider-facing use in this combined, nationally representative sample. This may indicate that, when individuals experience problems with care, eHealth activities that act as *alternatives* to the traditional healthcare system may seem more useful than tools that improve interactions with the system.

Contrary to our expectations, people with lower levels of education may be more likely to seek alternatives or supplements when care problems occur as compared to individuals with higher levels of education. This suggests a potential protective effect for a disadvantaged group (individuals with lower levels of education), as they seek alternatives or supplements when care problems occur. In particular, our findings suggest that individuals with lower education levels may react more to problems with patient-centered communication than individuals with higher education levels. Despite this relationship, individuals with lower education levels used eHealth resources at lower rates than those with higher education levels. This finding is consistent with a persistent digital divide in eHealth use associated with other health disparities [14,17]. These findings indicate that eHealth could help address differences in the quality of care received by different socioeconomic groups, but new strategies are needed to increase its adoption and use in vulnerable populations if these resources are to meet their potential of reducing health disparities [18]. Future work should focus on ensuring equitable access to eHealth resources as well as the creation and dissemination of culturally appropriate eHealth tools.

Our study has several limitations. First, the construction of the HINTS survey may have contributed to sorting of individual items in the factor analysis, as responses to nearby items on instruments are likely correlated by construction. Although some survey instruments loaded fully onto one factor, others contributed items to both or neither category, and the survey construction alone did not fully explain the pattern of our results. Future work should aim to replicate these results in other surveys. Second, our analyses were cross-sectional in nature. We observed associations between negative healthcare experiences and eHealth use and hypothesized that patients use eHealth in response to these care experiences, but our data cannot support this causal inference. As such, our results are subject to potential bias or reverse causality. One possible source of bias is that people with more complex health problems may be more likely to use provider-facing tools and experience coordination problems. To reduce the potential for bias, we included a set of patient demographics, internet access, and health status variables to control for the observed differences in respondents. Finally, although we discuss eHealth as a promising resource, we were unable to test whether its use improves health in individuals with negative care experiences. Measuring the impact of eHealth on outcomes and developing strategies to maximize the potential benefits of eHealth remain important areas of study but are beyond the scope of this research.

A strength of the current work is that the data were sourced from a nationally representative sample of Americans. However, we should ascertain whether these results can be replicated in other cultural contexts. People in "Western, Educated, Industrialized, Rich, and Democratic (WEIRD)" societies like the United States are frequent outliers in behavioral research [23], with a strong focus on independence and autonomy. However, Americans with lower levels of education, similar to the worldwide population, tend to value interpersonal connection more strongly [23,24]. This may partially explain the increased responsiveness to deficits in the patient-provider relationship among individuals with lower education levels. It is possible that the overall patterns of eHealth motivation more closely resemble those of more highly educated Americans in cultural contexts that value independence and those of less highly educated Americans in cultural contexts that value interpersonal connection. Compared to populations of other nations, Americans face shorter wait times to visit their providers. The use of eHealth tools in response to negative care experiences may be more prevalent in nations where followup visits to address these experiences are limited [25]. Therefore, it is possible that the trends observed in this study may be more pronounced in other settings.

Our findings indicate that individuals use eHealth activities, especially those that are independent of healthcare providers, when they experience problems with their healthcare. In particular, individuals with lower levels of education seem to use eHealth in response to negative healthcare experiences. Nonetheless, eHealth use remains low overall, and eHealth is an underused means of improving health outcomes. To maximize the potential for eHealth to meet the needs of all patients, especially those who are traditionally underserved by the healthcare system, additional work should ensure that eHealth resources are accessible to and usable by all members of the population.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sensitivity testing for stratified regression analyses, with "Some college" categorized as higher education.

[XLSX File (Microsoft Excel File), 14KB - jmir_v20i12e11034_app1.xlsx]

References

1. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
2. Heisey-Grove D, Patel V. ONC Data Brief, no. 28. Washington, DC: Office of the National Coordinator for Health Information Technology; 2015 Sep. Any, certified, and basic: Quantifying physician EHR adoption through 2014 URL: https://www.healthit.gov/sites/default/files/briefs/oncdatabrief28_certified_vs_basic.pdf [WebCite Cache ID 72ABn0EDY]
3. Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
4. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228.e1. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]
5. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
6. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for Patient Engagement: A Systematic Review. Front Psychol 2015;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
7. Li N, Orrange S, Kravitz RL, Bell RA. Reasons for and predictors of patients' online health information seeking following a medical appointment. Fam Pract 2014 Oct;31(5):550-556. [doi: 10.1093/fampra/cmu034] [Medline: 24963151]
8. 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] [Medline: 20390974]
9. Bell RA, Hu X, Orrange SE, Kravitz RL. Lingering questions and doubts: online information-seeking of support forum members following their medical visits. Patient Educ Couns 2011 Dec;85(3):525-528. [doi: 10.1016/j.pec.2011.01.015] [Medline: 21315538]
10. Hou J, Shim M. The role of provider-patient communication and trust in online sources in Internet use for health-related activities. J Health Commun 2010;15 Suppl 3:186-199. [doi: 10.1080/10810730.2010.522691] [Medline: 21154093]
11. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press; 2001.
12. Lee SY, Hawkins R. Why do patients seek an alternative channel? The effects of unmet needs on patients' health-related Internet use. J Health Commun 2010 Mar;15(2):152-166. [doi: 10.1080/10810730903528033] [Medline: 20390984]
13. Gidwani R, Zulman D. Association Between Acute Medical Exacerbations and Consuming or Producing Web-Based Health Information: Analysis From Pew Survey Data. J Med Internet Res 2015 Jun 23;17(6):e145 [FREE Full text] [doi: 10.2196/jmir.3801] [Medline: 26104000]
14. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
15. Nguyen A, Mosadeghi S, Almario CV. Persistent digital divide in access to and use of the Internet as a resource for health information: Results from a California population-based study. Int J Med Inform 2017 Jul;103:49-54. [doi: 10.1016/j.ijmedinf.2017.04.008] [Medline: 28551001]
16. 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] [Medline: 26944212]
17. Amo L. Education-Based Gaps in eHealth: A Weighted Logistic Regression Approach. J Med Internet Res 2016 Oct 12;18(10):e267 [FREE Full text] [doi: 10.2196/jmir.5188] [Medline: 27733329]
18. Latulippe K, Hamel C, Giroux D. Social Health Inequalities and eHealth: A Literature Review With Qualitative Synthesis of Theoretical and Empirical Studies. J Med Internet Res 2017 Apr 27;19(4):e136 [FREE Full text] [doi: 10.2196/jmir.6731] [Medline: 28450271]
19. Epstein RM, Street RL. Patient-centered communication in cancer care: Promoting healing and reducing suffering. Bethesda, MD: National Cancer Institute, US Department of Health and Human Services, National Institutes of Health; 2007. URL: https://healthcaredelivery.cancer.gov/pcc/pcc_monograph.pdf

20. Blanch-Hartigan D, Chawla N, Moser RP, Finney RLJ, Hesse BW, Arora NK. Trends in cancer survivors' experience of patient-centered communication: results from the Health Information National Trends Survey (HINTS). J Cancer Surviv 2016 Dec;10(6):1067-1077. [doi: 10.1007/s11764-016-0550-7] [Medline: 27193357]
21. Arora NK, Reeve BB, Hays RD, Clauser SB, Oakley-Girvan I. Assessment of quality of cancer-related follow-up care from the cancer survivor's perspective. J Clin Oncol 2011 Apr 01;29(10):1280-1289 [FREE Full text] [doi: 10.1200/JCO.2010.32.1554] [Medline: 21357781]
22. Lupiáñez-Villanueva F, Anastasiadou D, Codagnone C, Nuño-Solinís R, Garcia-Zapirain SMB. Electronic Health Use in the European Union and the Effect of Multimorbidity: Cross-Sectional Survey. J Med Internet Res 2018 May 03;20(5):e165 [FREE Full text] [doi: 10.2196/jmir.7299] [Medline: 29724702]
23. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? Behav Brain Sci 2010 Jun 15;33(2-3):61-83. [doi: 10.1017/s0140525x0999152x]
24. Cohen AB. Many forms of culture. Am Psychol 2009 Apr;64(3):194-204. [doi: 10.1037/a0015308] [Medline: 19348520]
25. Davis K, Stremlikis K, Squires D, Schoen C. Mirror, Mirror on the Wall, 2014 Update: How the Performance of the U.S. Health Care System Compares Internationally. New York: The Commonwealth Fund; 2014 Jun. URL: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2014_jun_1755_davis_mirror_mirror_2014.pdf [WebCite Cache ID 72ACtCw1L]

Abbreviations

HINTS: Health Information National Trends Survey
WEIRD: Western, Educated, Industrialized, Rich, and Democratic

Edited by G Eysenbach; submitted 14.05.18; peer-reviewed by D Alhuwail, J Alpert; comments to author 09.08.18; accepted 23.09.18; published 05.12.18.

Please cite as:
Senft N, Everson J
eHealth Engagement as a Response to Negative Healthcare Experiences: Cross-Sectional Survey Analysis
J Med Internet Res 2018;20(12):e11034
URL: https://www.jmir.org/2018/12/e11034/
doi:10.2196/11034
PMID:30518513

Review

Remote Patient Monitoring and Telemedicine in Neonatal and Pediatric Settings: Scoping Literature Review

Farzan Sasangohar[1,2,3], BA, BCS, MASc, SM, PhD; Elise Davis[2], BA, MPH; Bita A Kash[2,3], MBA, FACHE, PhD; Sohail R Shah[4,5], MSHA, MD

[1]Industrial and Systems Engineering, Department of Environmental and Occupational Health, Texas A&M University, College Station, TX, United States

[2]National Science Foundation Center for Health Organization Transformation, Department of Health Policy and Management, Texas A&M University, College Station, TX, United States

[3]Houston Methodist Hospital, Center for Outcomes Research, Houston, TX, United States

[4]Division of Pediatric Surgery, Texas Children's Hospital, Houston, TX, United States

[5]Department of Surgery, Baylor College of Medicine, Houston, TX, United States

Corresponding Author:
Farzan Sasangohar, BA, BCS, MASc, SM, PhD
Industrial and Systems Engineering
Department of Environmental and Occupational Health
Texas A&M University
101 Bizzell Street
College Station, TX, 77843
United States
Phone: 1 979 458 2337
Email: sasangohar@tamu.edu

Abstract

Background: Telemedicine and telehealth solutions are emerging rapidly in health care and have the potential to decrease costs for insurers, providers, and patients in various settings. Pediatric populations that require specialty care are disadvantaged socially or economically or have chronic health conditions that will greatly benefit from results of studies utilizing telemedicine technologies. This paper examines the emerging trends in pediatric populations as part of a systematic literature review and provides a scoping review of the type, extent, and quantity of research available.

Objective: This paper aims to examine the role of remote patient monitoring (RPM) and telemedicine in neonatal and pediatric settings. Findings can be used to identify strengths, weaknesses, and gaps in the field. The identification of gaps will allow for interventions or research to improve health care quality and costs.

Methods: A systematic literature review is being conducted to gather an adequate amount of relevant research for telehealth in pediatric populations. The fields of RPM and telemedicine are not yet very well established by the health care services sector, and definitions vary across health care systems; thus, the terms are not always defined similarly throughout the literature. Three databases were scoped for information for this specific review, and 56 papers were included for review.

Results: Three major telemedicine trends emerged from the review of 45 relevant papers—RPM, teleconsultation, and monitoring patients within the hospital, but without contact—thus, decreasing the likelihood of infection or other adverse health effects.

Conclusions: While the current telemedicine approaches show promise, limited studied conditions and small sample sizes affect generalizability, therefore, warranting further research. The information presented can inform health care providers of the most widely implemented, studied, and effective forms of telemedicine for patients and their families and the telemedicine initiatives that are most cost efficient for health systems. While the focus of this review is to summarize some telehealth applications in pediatrics, we have also presented research studies that can inform providers about the importance of data sharing of remote monitoring data between hospitals. Further reports will be developed to inform health systems as the systematic literature review continues.

(J Med Internet Res 2018;20(12):e295) doi:10.2196/jmir.9403

KEYWORDS

neonatal; pediatric; remote patient monitoring; telehealth; telemedicine

Introduction

The United States Department of Health and Human Services defines *telehealth* as the "use of electronic information and telecommunication technologies to support and promote long-distance clinical health care, patient and professional health-related education, public health and health administration" [1]. The Agency for Healthcare Research and Quality classifies telehealth into 3 distinct categories: (1) real-time video telehealth between the patient and a health care professional; (2) store and forward telehealth, such as the sharing of medical images or data between providers; and (3) home monitoring telehealth, involving the use of telehealth to remotely monitor patients and their health status, also known as remote patient monitoring (RPM) [2]. While telehealth is used for both clinical and nonclinical applications, the term *telemedicine* is used more exclusively for clinical applications or to diagnose and treat patients [3]. Various telemedicine technologies are emerging in health care very rapidly, and some of them can potentially be cost and time saving for patients and providers as well as offer improved quality of care. Historically, telemedicine techniques and technologies have been utilized by health systems within acute care settings and patient homes most commonly to improve access to care and monitor those with the greatest need. Technologies vary in terms of cost, patient adherence and utility, effectiveness, implementation success, desired health outcomes, and impact on capacity. Pediatric patients who often lack access to specialty pediatric care are socioeconomically disadvantaged or have chronic medical needs that may especially benefit from telemedicine. There is a need to identify and describe those telemedicine devices and techniques aimed at pediatric populations that are most promising in lowering costs of care, improving patient and family experience, decreasing time spent traveling, and increasing care capacity in hospitals and clinics. In this research, we aimed to shed some light on some noteworthy telemedicine technologies successfully used for pediatric patient segments.

A systematic literature review is being conducted to examine the technologies that are currently used in health systems to effectively provide telemedicine coverage for pediatric patients from remote locations. In this paper, we present the results of the scoping review that provides our preliminary findings on the type, extent, and quantity of research available in the literature. While the overall study takes a comprehensive approach in terms of pediatric patient populations studied by disease category, complexity, and patient segment, this paper aims to highlight some emerging RPM and telemedicine trends in the neonatal and pediatric literature. Results from this research can provide an overview of available evidence to inform practitioners, including hospitals and clinics, as well as health technology developers and care providers about the current state of and opportunities in RPM and telemedicine.

We first discuss the steps taken and update on the progress of the comprehensive systematic review. Additionally, some key findings of innovations and emerging technologies in RPM and telemedicine capabilities for pediatric patients are presented. Incremental updates of this review are intended to reduce unintended consequences and costs that come with failing to utilize telemedicine capabilities within and between health systems in various settings.

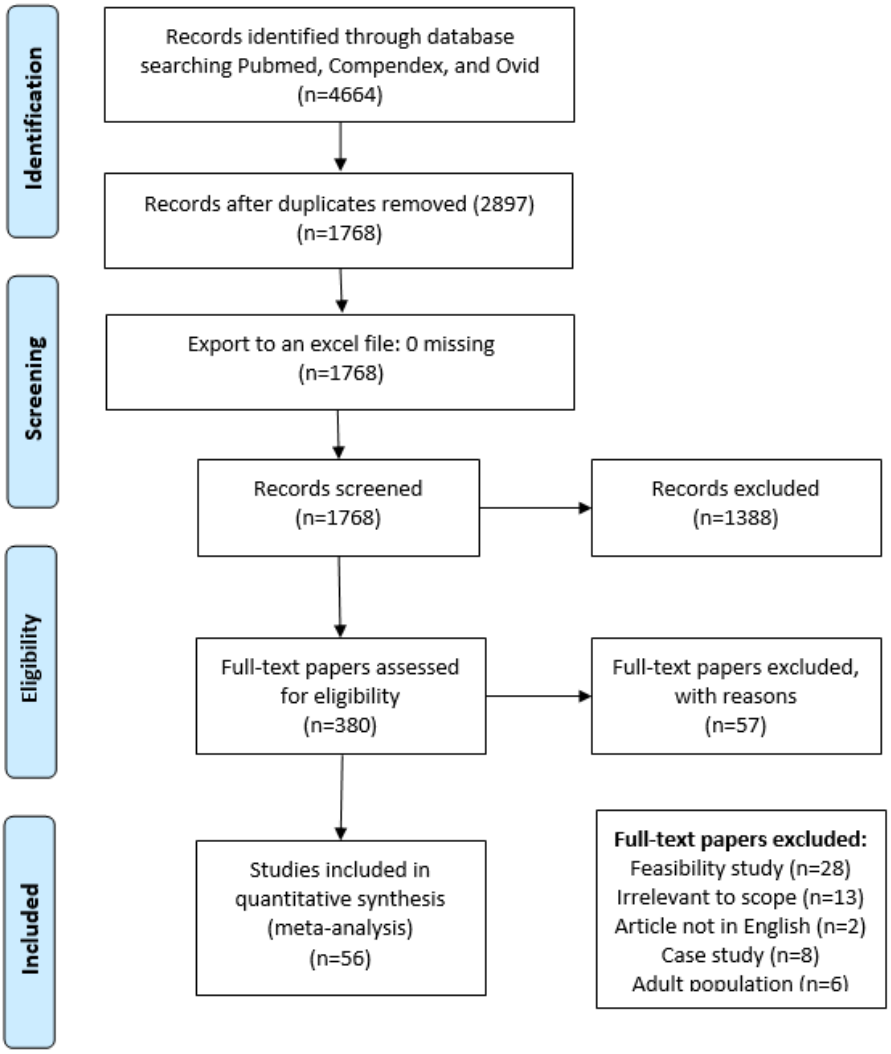
Methods

A systematic literature review is being conducted to gather an adequate amount of relevant research for telehealth in pediatric populations. The fields of RPM and telemedicine are not yet very well established by the health care services sector, and definitions vary across health care systems; thus, the terms are not always defined similarly throughout the literature. A preliminary search helped us to identify which terms provided the most literature on RPM and telemedicine and also helped us identify which databases to use.

A combination of search terms allowed us to obtain 4664 papers, which are relevant to pediatric RPM and telemedicine. All searches included either "child" or "pediatric" and at least one word comparable to "tele-monitoring," "telehealth," "telemedicine," or "remote monitoring." Some other important search terms were "population health" and "population management."

We began this search with a scope of the literature relevant to RPM and telemedicine in pediatric populations in PubMed, Compendex, and Ovid. Our search was restricted to peer-reviewed original studies published after January 1, 2008, and papers were collected between July 24 and September 2, 2016. After deleting duplicates, 1768 papers were included for an abstract review and screening. After applying the exclusion criteria, 380 papers were included for full-text review, of which 56 were selected to be included in this review. This review was conducted according to the guidelines of the Preferred Reporting Items for Systematic Review and Meta-Analyses, which visualizes the process of inclusion and exclusion of papers (Figure 1) [4]. Textbox 1 shows a brief explanation of our inclusion and exclusion criteria, respectively. A thematic analysis was then used to identify common patterns across the studies. One coder reviewed the papers and coded the RPM and telemedicine technologies used or evaluated. This paper summarizes our thematic synthesis.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram.



Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Biometric monitoring (n=59)
- Economic benefit (n=13)
- Patient or provider satisfaction (n=55)
- Teleconsultation (n=36)
- Telediagnosis (n=120)
- Telemanagement (n=43)
- Telepresence (n=14)
- Telesupport (n=40)

Exclusion criteria:

- Adult population (n=224)
- Case study (n=32)
- Duplicate (n=64)
- Irrelevant (n=328)
- No specific findings (n=28)
- No original research (n=13)
- Provider-initiated contact (n=46)
- Tele-education (n=86)
- Telementoring (n=13)
- Report (n=185)
- Subjective (n=9)
- Telephone-based intervention (n=55)
- Review (n=286)

Results

Summary

RPM for pediatric patients can be utilized effectively in many different settings for a variety of diseases and with a variety of emerging technologies. In some cases, pediatric patients are monitored in the hospital by a physician who is in a remote location. In other cases, hospitals are using technologies to monitor patients in the hospital, but without contact, thus, decreasing the likelihood of infection or other adverse health effects. Another exciting aspect of RPM is that of monitoring patients at their home via continuous monitors or via self-uploading of patient data from a monitoring device at the home. The majority of studies we have reviewed demonstrate significant positive results, such as improved health outcomes and cost savings to patients and providers, regarding patients who are vulnerable in terms of cardiac health or diabetes. The following sections summarize the emerging themes identified in our scoping review, which examine the role of RPM technologies and provide support for their efficacy.

Use of Continuous Glucose Monitoring

The management of type 1 diabetes in children can be challenging. Several research teams have examined the role of RPM in the management of type 1 diabetes for children, which especially helps to alert families and health professionals of hyper- and hypoglycemic critical concerns [5-10]. A key concern for these research teams was nocturnal hyper- and hypoglycemia, so glycemic levels were closely monitored throughout the day as well as during the nighttime hours.

Pena et al [9] used a glucose monitor that required patients and patient families to send glycemic information (mean blood glucose, glycated hemoglobin, and indexes of glucose volume) via email at five specific times suggested by providers throughout the day for monitoring. If any critical concerns arose, the families were contacted by the Diabetes Unit of the treating hospital via short message service text message or email. This form of RPM, which requires patients and their families to transfer data via email 5 times per day, led to adherence issues, yet it resulted in a significant decrease in glycated hemoglobin levels and overall improved metabolic control [9]. Additionally, Pena et al's system was well accepted by parents. Unfortunately, this system was not sustainable as metabolic control returned to baseline after the study discontinued. This calls for a model for glucose monitoring that is easier for patients and families to adhere to and that emphasizes the importance of the patient-integrated care model.

The remainder of studies highlighted real-time RPM utilizing a continuous glucose monitoring (CGM) system, which simply required patients to wear a monitoring device. Three systems used CGM in association with an insulin pump so that alarms were triggered when glycemic values were critical, but insulin pumps were used to treat the critical values while the alarm was working to alert both parents and remote clinician teams of the concern [6-8]. One system did not use the insulin pump with CGM but did use alarms to alert caregivers and clinician teams in a remote setting of any critical values, thus, allowing children to be treated with appropriate levels of insulin as needed by parents or caregivers [5]. In all cases, if parents or caregivers did not respond, the remote clinician teams were available 24/7 to attempt further contact to alert caregivers of the critical values in children with type 1 diabetes.

The real-time CGM systems were able to shorten the length of hypoglycemic events in children, thus, preventing any adverse health outcomes associated with hypoglycemic events [5]. Patients and family members felt comfortable using these systems, felt that they were easy to use and understand, expressed that they would recommend the system to other families, and felt a sense of comfort knowing there was a clinician team available for backup throughout the day and at night [8]. Overall, CGM systems improved diabetes management success, and there were no safety issues identified throughout any of the studies mentioned [5-10].

Home Monitoring of Cardiovascular Implantable Devices

Cardiovascular implantable devices are increasingly being used in the pediatric population as a method of long-term RPM [11]. A variety of research studies have examined the role of RPM with implantable devices in decreasing the incidence of adverse cardiac events [11-14]. In these studies, patients with newly implanted cardiac devices were followed either prospectively or retrospectively via RPM and compared with patients with the same devices who were monitored traditionally. All 4 studies highlighted here used automated data sent from patients to a cardiac or pacemaker care center. At the cardiac care center, data were analyzed by a cardiac physician or care nurse and contact to patients and families was initiated via the internet, telephone, or short message service text messaging depending on the results, typically in the form of an electrocardiograph (ECG). In two cases, patients were also able to report symptoms and record specific suspected cardiac events to be sent to the cardiac care team [13,14].

Researchers found several benefits from remotely monitoring pediatric patients with implantable cardiac devices. Leoni et al [14] prevented 72 clinic visits, or an average of 2 hours and 35 minutes of transportation time, for patients by monitoring symptoms remotely and communicating effectively with patients and families. In addition, 87% of patients and families rated the remote monitoring to be “very easy to perform” in the study. Leshem-Rubinow et al [13] achieved a median time between data transmission and viewing ECG data of 7 minutes; interpretation of the ECG was accomplished by trained cardiac staff within 5 hours, and the diagnosis of cardiac events averaged at 16 hours after the data transmission. Malloy et al [11] found that RPM decreased the average number of days that patients went without physician contact, potentially decreasing adverse events. For patients on a 6-month follow-up regimen, there was a temporal gain of 134 days of physician contact, and for patients on a 3-month follow-up regimen, there was a gain of 44 days. Patients in the study by Zartner et al [12] experienced 33 pacemaker shocks that successfully terminated ventricular tachycardia, improving the overall safety and well-being of patients outside a clinical setting. All researchers found that their systems were acceptable and easy to use and had a low number of false alarms from their devices. False alarms can easily be improved with continued use of devices, and they do no harm to patients or their families [11,12].

Mobile Robotic Telemedicine in the Neonatal Intensive Care Unit

One of the benefits of telemedicine is that it allows access to specialists and subspecialists in settings where it may not be feasible or possible. Robotic telepresence (RTP) is a form of telemedicine that allows face-to-face contact between a specialist and a patient in a hospital [15,16]. An increase in preterm deliveries and survival rates with advances in neonatal medicine have resulted in a need for neonatal intensive care units (NICUs) to staff more neonatal specialists during more hours of the day [17]. A solution to these increased pressures on NICUs is the model of RTP to monitor patients in the NICU from remote locations. RTP machines are linked to the NICU and the remote location via the internet and have synchronous bidirectional audio and visual communication capabilities with zoom and a digital camera for image capture. In addition, the video screen is able to move as per the requirement of the physician or neonatal care specialist while caring for patients. A digital stethoscope, otoscope, and pulse oximeter allow the physician to check vital signs, listen to heart and bowel sounds, and better evaluate the patient while in a remote location [15,16]. By working together with onsite nurses, offsite neonatal providers can maneuver the RTP machine on their own from a distant location, and motion sensors keep the machine from bumping into any incubators or medical equipment. Visual and audio capabilities allow remote physicians to communicate with NICU nurses and families of patients.

Garingo et al [15] studied the ability of onsite and offsite neonatologists to physically examine patients in the NICU and found that local and remote physicians had good or excellent agreement for most assessments of patients. Rincon et al [16] showed that NICU nurses felt that physicians were easily accessible via RTP and that they were adequately involved and supportive of both nurses and NICU patients and their families. In addition, nurses felt they had sufficient time to ask questions and had the resources to care for patients with the simple use of RTP. A novel benefit to RTP is that neonatologists are able to monitor NICU patients during the nighttime hours, when fewer nurses are available. Overall, RTP enhanced communication and improved access for NICU patients; furthermore, cost savings are implied with remote physician capabilities.

Telehealth Capabilities for Remote Consultation and Diagnosis

In addition to the capabilities summarized previously, telemedicine can be used for consultations and diagnosis of health concerns from remote settings. Patients and providers can save on travel time and costs, and patients who are unable to travel will benefit from specialty physician consultation via videoconferencing. In emergent cases, physicians are able to provide timely feedback to families and patients who would otherwise have to incur a great deal of costs on ground or air ambulance [18]. When using telehealth capabilities instead of telephone or email for a consult, physicians are also able to provide more accurate diagnoses and, thus, more appropriate treatment for patients [19]. In addition, physicians are able to consult with pediatric patients via a Web camera and a high-quality television screen. This allows open communication between patients, physicians, and patient families and caregivers. The same quality of care is capable of being provided in these video consults according to previous research [18,19]. Rowell et al [18] found that 40% of pediatric patients receiving orthopedic consultations via videoconferencing were discharged after one telehealth consult and 58% of patients did not require a further in-person appointment.

Dharmar et al [19] studied the effectiveness of physicians in prescribing appropriate medications and doses to pediatric patients in critical care via telemedicine. Physicians made significantly fewer medication errors in patients who received a consult via videoconferencing compared with those who received a telephone consult or did not receive a consult at all. This was an important finding as physicians were dealing with critically ill and seriously injured pediatric patients in the emergency department.

Another study examined the role of mobile telemedicine units in low-income, inner-city neighborhoods of Rochester, New York [20]. McIntosh et al [20] used health workers with minimal training to visit acute care patients along with videoconferencing capabilities to a primary care facility. By visiting patients in their homes, health workers with video access to primary care facilities saved 30% of families a trip to the emergency department and 17% of families a trip to the urgent care clinic. Close to 90% of caregivers were highly satisfied with the service and found it to be very convenient. Furthermore, McIntosh et al [20] suggested that the creation of a sustainable plan for this service with payment models included would be highly beneficial to low-income areas in the United States.

Telemedicine Technologies Without the Use of Remote Patient Monitoring

Some technologies discovered from the literature review are relevant to telemedicine, yet they fail to utilize the aspects of RPM. Below we discuss two such systems: closed-loop systems and noncontact heart rate monitoring.

Closed-Loop Systems

Both Ly et al [7] and Tauschmann et al [10] used a closed-loop monitoring system, which does not require remote monitoring or supervision by clinicians. Closed-loop insulin delivery systems use a CGM device along with an automated insulin delivery device. Patients have to calibrate their devices approximately 4 times per day with a finger prick. Overall, these closed-loop systems lower mean glucose levels and reduce the amount of time spent above target glucose levels without altering daily insulin amount. In both studies, patients and families had access to clinicians or nurses 24/7 in the case of emergencies or difficulties with the system [7,10]. In this case, clinicians are left out of the loop, yet data can easily be shared remotely and monitored in case of any emergencies.

Noncontact Heart Rate Monitoring of Infants in the Neonatal Intensive Care Unit

Similar to closed-loop systems, several health systems are using telemedicine in hospitals, which do not require continuous physician monitoring. This creates an opportunity for sharing data with remote locations.

Several previous researchers have documented the development of robust methods for automated computation of heart rate of infants in the NICU [21,22]. Heart rate is a critical vital sign to continually assess for infants in the NICU, but current techniques involve wearing adhesive gel patches or chest straps, which can easily cause skin irritation. For NICU patients who are especially susceptible to infection, a noncontact heart rate monitor would improve overall health and decrease stress among patients and their parents [21,22]. Aarts et al and Bal used photoplethysmography (PPG), which is inexpensive and simple to use, but typically is used as a contact device using adhesive sensors [21,22]. A recent advancement is the use of camera-based PPG, a noncontact method of remotely recording PPG signals from patients using a camera and ambient light [21,22].

Aarts et al studied patients in the NICU in California and the Netherlands through noncontact PPG with an objective of exploring potential challenges of the noncontact PPG technique [21]. A total of 19 infants were examined using noncontact PPG, which provided a good measure of heart rate for >90% of the time. The study team was able to monitor heart rate by setting up a camera approximately 1 m away from infants; the camera monitored infants either through plexiglass or with open incubators. Researchers ensured that the light within the NICU was appropriate for monitoring with the camera and there was never a need for infants to be touched, removed from incubators, or repositioned throughout the study. Recordings were taken from an undressed portion of the skin (head, arm, or thorax). The recordings from the camera were saved and transferred to a computer, where the heart rate was then obtained using pulse oximetry sensors or ECG sensors [21].

There are two major limitations to noncontact PPG as identified by Aarts et al [21]. First, to be feasible, noncontact PPG must record at a random anatomical location on the skin, and noncardiovascular events may negatively affect how PPG signals are recorded. Thus, repositioning of a limb or redistribution of venous blood could affect how heart rate is identified. Additionally, the study team was unable to obtain an appropriate signal for heart rate monitoring if the infant was squirming. To ensure infant stability, the team monitored PPG signals during kangaroo mother care, and despite the slight rocking of the infant, accurate PPG signals were recorded. It is important to

remember, though, that not all infants are able to engage in kangaroo mother care, and the squirming of infants remains a limitation to the noncontact PPG technique.

Bal conducted a similar study of webcam-based PPG for heart rate and oxygen saturation of healthy infants and NICU patients in Turkey [22]. Bal avoided the issue of nonstationary infants using wavelet transform, a technique that has the ability to detect rapid changes in frequency. Instead of strictly using fluorescent lighting, Bal also used sunlight for the proper detection of PPG signals and placed subjects just 50 cm from the camera. Again, recordings were sent to a computer for further analysis using ECG. Overall, Bal was able to conclude that PPG signals were accurate in both sunlight and fluorescent light and that this method monitors heart rate and oxygen saturation accurately and safely without patient contact [22].

Contactless heart rate monitoring is important in the NICU because it can help avoid infection, thus, decreasing health care cost and stress on families. Additionally, this technique is simple, inexpensive, and effective with the appropriate parameters in place such as light and distance to the camera. Neither research mentioned was disruptive of hospital or clinician flow. The avoidance of touching and repositioning infants allows patients proper rest and development within the incubator.

Discussion

Principal Findings

Our scoping review showed that research on telemedicine applications for pediatric populations is limited, and of the existing research, many studies are severely limited by small sample sizes and convenience samples of participants. In addition, much of the research on telemedicine technologies for pediatrics relies on the satisfaction of parents and caregivers of children with varying diseases. Further research can be strengthened with the education of parents about the importance of enrolling their children in studies that utilize telemedicine services to improve adherence to care management plans and sustainability of the care model. While the benefit for a limited set of diseases is apparent, the effects of telemedicine on patient care and clinical outcomes need to be examined further for a wider range of conditions. By filling this gap in research, health care providers will find opportunities for greater utilization of telemedicine in their health systems.

Based on our findings, there are a wide variety of ways in which telehealth can be used effectively in a health system. Our brief report covers a limited scope of the types of services and devices that are being effectively used for RPM and telemedicine in pediatrics. These include CGM of pediatrics with type 1 diabetes, home monitoring of cardiovascular implantable devices, remote robotic telemedicine in the NICU, and remote consultation and diagnosis. We also presented closed-loop insulin delivery without remote monitoring and noncontact heart rate monitoring of infants in the NICU. The results of our systematic literature review may shed more light on potential research areas or adoption decisions by summarizing some of the more innovative and emerging telehealth capabilities being used throughout pediatric and neonatal health systems. However, this scoping review may help health care providers to remain current with the large plethora of emerging technologies and trends.

Our findings presented in this paper are also limited to studies in developed countries. One application has been reviewed in Malawi, Africa, a developing country. While this study was not necessarily relevant to the scope of this brief review, it may be important for future research and telehealth applications. In developing countries, access to a quality internet connection is rare, yet in larger cities, it is becoming more widely utilized by health systems and hospitals. Effective telemedicine consultations require high-quality equipment with appropriate internet connection and strong service coordination [18,23]. In Malawi, Africa, there are a total of 4 pathologists throughout the country, serving a population of 14 million [23]. The Queen Elizabeth Central Hospital in Malawi connected with a highly qualified hospital in Newcastle, United Kingdom, to obtain a speedy and efficient diagnosis of pediatric oncology cases. If the hospital in Malawi had waited for local diagnosis, they could spend anywhere from 3 weeks to 4 months waiting, whereas remote telepathologists were able to send diagnostic information within 24 hours; this is critical time for patients with oncological concerns, especially in resource-poor settings [23].

Some other aspects of RPM and telemedicine that have not been addressed in this report, but will be addressed in future reports, are cost savings to patients, families, and hospitals; the role of telesupport and telepresence between clinicians and providers; telediagnosis of a variety of diseases and medical conditions; and the importance of telemedicine in improving patient, family, and provider satisfaction.

Conclusions

Despite the limited applications of telemedicine in pediatric and neonatal settings, current technologies show promise in several domains. Small sample size continues to be the main limitation of telemedicine studies in pediatrics. Continued research in telemedicine and RPM applications to a wider range of conditions will further emphasize the need for emerging trends in pediatric health systems. The information presented can inform health care providers of the most widely accepted forms of telemedicine for patients and their families and of the telemedicine that is most cost efficient for health systems. While the focus of this report is on RPM, we have presented some research studies that can inform providers about the importance of data sharing of remote monitoring data between hospitals. Continued reports of findings from this scoping literature review will educate key informants about the importance of telemedicine for pediatric populations and their families.

Acknowledgments

We would like to acknowledge Texas Children's Hospital and Rakesh Mehta for the inspiration on a topic for this literature review. This research was funded by a National Science Foundation Industry–University Cooperative Research Centers Program agreement between National Science Foundation Center for Health Organization Transformation and Texas Children's Hospital. One of the coauthors affiliated with Texas Children's Hospital, SRS, has contributed to the design, analysis, and writing of the manuscript.

Conflicts of Interest

None declared.

References

1. Health Resources and Services Administration. Telehealth Programs. 2015. URL: https://www.hrsa.gov/ruralhealth/telehealth/ [accessed 2017-01-03] [WebCite Cache ID 73TNZrIAy]
2. Agency for Healthcare Research and Quality. Health Information Technology. 2017. URL: https://healthit.ahrq.gov/key-topics/telehealth [accessed 2017-01-07] [WebCite Cache ID 73TNb3GCi]
3. World Health Organization. Health and sustainable development, telehealth. 2017. URL: http://www.who.int/sustainable-development/health-sector/strategies/telehealth/en/ [accessed 2017-01-03] [WebCite Cache ID 73TNbYiTX]
4. Moher D, Liberati A, Tetzlaff J, Altman D, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 2009 Jul 21:e1000097. [doi: 10.1371/journal.pmed.1000097]
5. DeSalvo DJ, Keith-Hynes P, Peyser T, Place J, Caswell K, Wilson DM, et al. Remote glucose monitoring in cAMP setting reduces the risk of prolonged nocturnal hypoglycemia. Diabetes Technol Ther 2014 Jan;16(1):1-7. [doi: 10.1089/dia.2013.0139] [Medline: 24168317]
6. Kaiserman K, Buckingham BA, Prakasam G, Gunville F, Slover RH, Wang Y, et al. Acceptability and utility of the mySentry remote glucose monitoring system. J Diabetes Sci Technol 2013 Mar 01;7(2):356-361 [FREE Full text] [doi: 10.1177/193229681300700211] [Medline: 23566993]
7. Ly TT, Keenan DB, Roy A, Han J, Grosman B, Cantwell M, et al. Automated Overnight Closed-Loop Control Using a Proportional-Integral-Derivative Algorithm with Insulin Feedback in Children and Adolescents with Type 1 Diabetes at Diabetes Camp. Diabetes Technol Ther 2016 Dec;18(6):377-384. [doi: 10.1089/dia.2015.0431] [Medline: 27183197]
8. Oron T, Farfel A, Muller I, Miller S, Atlas E, Nimri R, et al. A remote monitoring system for artificial pancreas support is safe, reliable, and user friendly. Diabetes Technol Ther 2014 Nov;16(11):699-705. [doi: 10.1089/dia.2014.0090] [Medline: 25211216]
9. Peña NV, Torres M, Cardona JAC, Iniesta R. Impact of telemedicine assessment on glycemic variability in children with type 1 diabetes mellitus. Diabetes Technol Ther 2013 Feb;15(2):136-142. [doi: 10.1089/dia.2012.0243] [Medline: 23289433]
10. Tauschmann M, Allen J, Wilinska M, Thabit H, Stewart Z, Cheng P, et al. Day-and-Night Hybrid Closed-Loop Insulin Delivery in Adolescents With Type 1 Diabetes: A Free-Living, Randomized Clinical Trial. Diabetes Care 2016 Dec;39(7):1168-1174 [FREE Full text] [doi: 10.2337/dc15-2078] [Medline: 26740634]
11. Malloy LE, Gingerich J, Olson MD, Atkins DL. Remote monitoring of cardiovascular implantable devices in the pediatric population improves detection of adverse events. Pediatr Cardiol 2014 Feb;35(2):301-306 [FREE Full text] [doi: 10.1007/s00246-013-0774-5] [Medline: 23949666]
12. Zartner PA, Toussaint-Goetz N, Photiadis J, Wiebe W, Schneider MB. Telemonitoring with implantable electronic devices in young patients with congenital heart diseases. Europace 2012 Jul;14(7):1030-1037. [doi: 10.1093/europace/eur434] [Medline: 22308087]
13. Leshem-Rubinow E, Berger M, Shacham J, Birati EY, Malov N, Tamari M, et al. New real-time loop recorder diagnosis of symptomatic arrhythmia via telemedicine. Clin Cardiol 2011 Jul;34(7):420-425 [FREE Full text] [doi: 10.1002/clc.20906] [Medline: 21618252]
14. Leoni L, Padalino M, Biffanti R, Ferretto S, Vettor G, Corrado D, et al. Pacemaker remote monitoring in the pediatric population: is it a real solution? Pacing Clin Electrophysiol 2015 May;38(5):565-571. [doi: 10.1111/pace.12600] [Medline: 25645302]
15. Garingo A, Friedlich P, Tesoriero L, Patil S, Jackson P, Seri I. The use of mobile robotic telemedicine technology in the neonatal intensive care unit. J Perinatol 2012 Jan;32(1):55-63. [doi: 10.1038/jp.2011.72] [Medline: 21617643]
16. Rincon F, Vibbert M, Childs V, Fry R, Caliguri D, Urtecho J, et al. Implementation of a model of robotic tele-presence (RTP) in the neuro-ICU: effect on critical care nursing team satisfaction. Neurocrit Care 2012 Aug;17(1):97-101. [doi: 10.1007/s12028-012-9712-2] [Medline: 22547040]
17. Martin JA, Kochanek KD, Strobino DM, Guyer B, MacDorman MF. Annual summary of vital statistics--2003. Pediatrics 2005 Mar;115(3):619-634. [doi: 10.1542/peds.2004-2695] [Medline: 15741364]
18. Rowell P, Pincus P, White M, Smith A. Telehealth in paediatric orthopaedic surgery in Queensland: a 10-year review. ANZ J Surg 2014 Dec;84(12):955-959. [doi: 10.1111/ans.12753] [Medline: 25040240]

19. Dharmar M, Kuppermann N, Romano PS, Yang NH, Nesbitt TS, Phan J, et al. Telemedicine consultations and medication errors in rural emergency departments. Pediatrics 2013 Dec;132(6):1090-1097 [FREE Full text] [doi: 10.1542/peds.2013-1374] [Medline: 24276844]
20. McIntosh S, Cirillo D, Wood N, Dozier AM, Alarie C, McConnochie KM. Patient evaluation of an acute care pediatric telemedicine service in urban neighborhoods. Telemed J E Health 2014 Dec;20(12):1121-1126 [FREE Full text] [doi: 10.1089/tmj.2014.0032] [Medline: 25290233]
21. Aarts LAM, Jeanne V, Cleary JP, Lieber C, Nelson JS, Bambang Oetomo S, et al. Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit - a pilot study. Early Hum Dev 2013 Dec;89(12):943-948. [doi: 10.1016/j.earlhumdev.2013.09.016] [Medline: 24135159]
22. Bal U. Non-contact estimation of heart rate and oxygen saturation using ambient light. Biomed Opt Express 2015 Jan 01;6(1):86-97 [FREE Full text] [doi: 10.1364/BOE.6.000086] [Medline: 25657877]
23. Carey P, Fudzulani R, Scholfield D, Chagaluka G, Tomoka T, Liombe G, et al. Remote and rapid pathological diagnosis in a resource challenged unit. J Clin Pathol 2014 Jun;67(6):540-543. [doi: 10.1136/jclinpath-2013-202099] [Medline: 24561316]

Abbreviations

CGM: continuous glucose monitoring
ECG: electrocardiograph
NICU: neonatal intensive care unit
PPG: photoplethysmography
RPM: remote patient monitoring
RTP: robotic telepresence

Edited by K Eddens; submitted 13.11.17; peer-reviewed by E Pullen; comments to author 21.03.18; revised version received 17.04.18; accepted 17.09.18; published 20.12.18.

Please cite as:
Sasangohar F, Davis E, Kash BA, Shah SR
Remote Patient Monitoring and Telemedicine in Neonatal and Pediatric Settings: Scoping Literature Review
J Med Internet Res 2018;20(12):e295
URL: http://www.jmir.org/2018/12/e295/
doi:10.2196/jmir.9403
PMID:30573451

Original Paper

Questionnaire Breakoff and Item Nonresponse in Web-Based Questionnaires: Multilevel Analysis of Person-Level and Item Design Factors in a Birth Cohort

Cauane Blumenberg[1*], MSc; Daniela Zugna[2*], PhD; Maja Popovic[2*], MSc; Costanza Pizzi[2*], PhD; Aluisio J D Barros[1*], PhD; Lorenzo Richiardi[2*], PhD

[1]Center for Epidemiological Research, Post-Graduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil

[2]Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Turin, Italy

[*]all authors contributed equally

Corresponding Author:
Cauane Blumenberg, MSc
Center for Epidemiological Research
Post-Graduate Program in Epidemiology
Federal University of Pelotas
1160 Marechal Deodoro Street, 3rd Floor
Pelotas, 96020-220
Brazil
Phone: 55 53 32841300
Fax: 55 53 32841300
Email: cauane.epi@gmail.com

Abstract

Background: Web-based questionnaires are increasingly used in epidemiologic studies, as traditional methods are facing a decrease in response rates and an increase in costs. However, few studies have investigated factors related to the level of completion of internet-based epidemiologic questionnaires.

Objective: Our objective was to identify person-level characteristics and item design factors associated with breakoff (not finishing the questionnaire) and item nonresponse in a Web-based questionnaire.

Methods: This study was a cross-sectional analysis of the baseline questionnaire, applied from 2005 to 2016, of the Italian NINFEA (Nascita e Infanzia: gli Effetti dell'Ambiente) birth cohort. The baseline questionnaire was administered to enrolled women, who could register at any time during pregnancy. We used logistic regression to analyze the influence of person-level factors on questionnaire breakoff, and a logistic multilevel model (first level: items of the questionnaire; second level: sections of the questionnaire; third level: study participants) to analyze the influence of person-level and item design factors on item nonresponse. Since the number of applicable items depended on the respondent's characteristics and breakoff, we used inverse probability weighting to deal with missing by design.

Results: Of 5970 women, 519 (8.69%) did not finish the questionnaire. Older age (adjusted odds ratio 1.40, 95% CI 1.05-1.88), lower educational level (adjusted odds ratio [OR] 1.53, 95% CI 1.23-1.90), and earlier stage of pregnancy (adjusted OR 3.01, 95% CI 2.31-3.92) were positively associated with questionnaire breakoff. Of the 1,062,519 applicable items displayed for the participants, 22,831 were not responded to (overall prevalence of item nonresponse 2.15%). Item nonresponse was positively associated with older age (adjusted OR 1.25, 95% CI 1.14-1.38), being in the first trimester of pregnancy (adjusted OR 1.18, 95% CI 1.06-1.31), and lower educational level (adjusted OR 1.23, 95% CI 1.14-1.33). Dropdown menu items (adjusted OR 1.77, 95% CI 1.56-2.00) and items organized in grids (adjusted OR 1.69, 95% CI 1.49-1.91) were positively associated with item nonresponse.

Conclusions: It is important to use targeted strategies to keep participants motivated to respond. Item nonresponse in internet-based questionnaires is affected by person-level and item design factors. Some item types should be limited to reduce item nonresponse.

(J Med Internet Res 2018;20(12):e11046) doi:10.2196/11046

KEYWORDS

epidemiology; internet; surveys and questionnaires; epidemiologic research design; data collection

Introduction

Background

Novel data collection methods are increasingly used in epidemiologic studies [1,2], as traditional methods, including mail questionnaires, face-to-face interviews, and telephone interviews, are facing a decrease in response rates [3] and an increase in costs [4]. Given the limitations of traditional methods and the growing internet penetration, the number of Web-based e-epidemiologic studies is increasing worldwide [1].

Compared with traditional methods, Web-based epidemiologic questionnaires have clear advantages, such as higher data quality (if filtering questions and consistency checks are used) and lower costs [1]. However, they may also have weaknesses that should be explored empirically [5]. In particular, the validity of epidemiologic studies may be jeopardized by lower response rates [6], questionnaire breakoff (not finishing the questionnaire), and item nonresponse [7], which can depend on participants' characteristics and item design factors [8,9]. For instance, in a study investigating homosexual rights, the item nonresponse rates were higher among heterosexual individuals than among homosexual individuals [10]. In this case, the item nonresponse rates varied according to individual characteristics that were relevant to the objectives of the study, and this could bias the results [11].

Objective

Although item nonresponse may have a great impact on study validity, few studies have investigated factors related to the level of completion of internet-based epidemiologic questionnaires [12,13]. Thus, in the context of the internet-based NINFEA (Nascita e Infanzia: gli Effetti dell'Ambiente) birth cohort study [14], we aimed at investigating the associations of person-level characteristics and item design factors with item nonresponse rate, as well as the associations of person-level characteristics with questionnaire breakoff.

Methods

Participants and Baseline Questionnaire

NINFEA is a Web-based birth cohort study, which started in Italy in 2005 [14]. Members of the cohort are children born to women who (1) had enough knowledge of the Italian language to complete internet-based questionnaires, (2) knew about the study, and (3) had access to the internet at the time of recruitment. Participants were recruited online through hyperlinks leading to the NINFEA website that were displayed on selected hospitals' home pages, pregnancy-related websites, and the NINFEA Facebook page, and offline using leaflets, face-to-face contacts, and posters placed in selected hospitals and clinics. The study was also advertised in local and national media in Italy. All selected hospitals and clinics for online and offline recruitment were located in the Piedmont and Tuscany regions, from which 82.87% (6391/7712) of the cohort originated. Pregnant women could enroll by registering at the study website [15] at any time during pregnancy. The ethical committees of the San Giovanni Battista Hospital and the Orthopedic Traumatology Center, Functional Re-education Center, Maria Adelaide Hospital, Turin, Italy (approval #0048362 and following amendments) approved the study, and all participants consented to participate. At enrollment, they completed a baseline questionnaire, and then were invited to fill in 5 follow-up questionnaires when their child turned 6 months, 18 months, 4 years, 7 years, and 10 years of age. This study focused on the baseline questionnaire.

In the period from 2005 to 2016, a total of 7712 pregnant women completed the NINFEA baseline questionnaire (database version 03.2017), and 1176 women participated during more than 1 pregnancy. The questionnaire was initially developed using the Hypertext Preprocessor scripting language [16]. After the first 1500 respondents, a major review of the questions was done and an updated version of the questionnaire was implemented using the Ruby programming language [17]. To avoid comparability issues, for this study we considered only the 5970 pregnant women who completed at least one section of the Ruby version.

The baseline questionnaire is composed of 18 sections investigating demographic factors, maternal general health, exposures before and during pregnancy, lifestyle, and reproductive history. Of these sections, 4 are supplementary and entirely dependent on answers given in the preceding section, and thus we did not consider them in the analyses. In total we included 244 items in the 14 analyzed sections; of these, 7 items were mandatory and therefore we excluded them from the analyses. We thus analyzed a total of 237 items, although the actual number of items presented to each participant at the time they completed the questionnaire varied due to filter questions that render sets of questions not applicable. For example, a negative answer to the filter question "Did you smoke during pregnancy?" would skip a series of questions about smoking. In contrast, a positive answer to the same filter question would present a set of applicable questions about smoking to the respondent.

Questionnaire Breakoff and Item Nonresponse

We analyzed 2 outcomes: questionnaire breakoff and item nonresponse. We considered a respondent to have broken off the questionnaire if she stopped answering the items before reaching the last section. If the last section was fully or partially completed and submitted, we considered the questionnaire not to be broken off, even if some items were left blank in the preceding sections. For this reason, no breakoff could have occurred in the last section of the questionnaire. For the analysis of questionnaire breakoff, the units of analysis were the 5970 women who completed at least one section of the questionnaire.

We based the analyses of item nonresponse on the 237 nonmandatory items from the 14 sections of the questionnaire. We assessed each of the 237 nonmandatory items, for each of the 5970 participants, and considered a blank as a nonresponse if the item was applicable. Item nonresponse was constructed

as a binary variable: 1=nonresponse, and 0=response. The units of analysis were the items of the questionnaire (at most 237 items × 5970 women = 1,414,890 items).

We analyzed the following person-level characteristics as predictors of questionnaire breakoff: age (≤30 years, 31-35 years, ≥36 years), university degree (yes, no), gestational trimester at enrollment (first trimester, second trimester, third trimester), first pregnancy (yes, no), employment status at the beginning of the pregnancy (employed, unemployed), type of recruitment (offline, online), Italian region of residence (Piedmont Region, Tuscany Region, other regions of Northern Italy, and other), and number of participations in the baseline questionnaire (1, ≥2). All the exposure variables were self-reported in the baseline questionnaire, except for the number of participations, which was constructed based on the total number of baseline questionnaires compiled by a woman. We assessed the type of recruitment from the first question, which asked about the way the participant had become aware of the study. We considered leaflets, posters, word-of-mouth, face-to-face invitation, and traditional media as offline recruitment methods, while we considered built-in links in websites and social media sites as online recruitment methods. Specifically, for the online recruitment, we advertised the study in selected forums or websites targeting pregnant women or health care workers, on the home pages of selected obstetric or pediatric hospitals or hospitals with a large number of deliveries, and on the NINFEA Facebook page. The number of involved websites, forums, and hospitals changed over time depending on the specific type of collaboration that was initiated. We conducted two small Facebook campaigns with advertisements targeting women in fertile age [18].

We assessed item nonresponse in association with the person-level characteristics analyzed for questionnaire breakoff, as well as in association with the design of the items themselves: (1) item type (checkbox, dropdown menu, radio button, text), (2) number of response options, and (3) whether the item was located in a grid (yes, no). Multimedia Appendix 1 provides examples of the item design characteristics. Specifically, radio button items can have only 1 answer selected among a set of predefined response options; dropdown menu items also have only 1 possible answer, but the list of response options is collapsed by default and has to be actively expanded to read the possible responses; checkboxes accept the selection of more than 1 answer from a set of predefined response options; and text items require the insertion of numeric or textual content. Some items in the questionnaire combined a radio button or a checkbox with a text item (eg, items with response options “Other, namely...”); these were considered as 2 individual items. We categorized the number of response options as 2, 3 to 5, and at least 6 options; we did not consider text items because they do not have any response option. An item was considered to be located in a grid if it was part of a group of items that shared the same set of response options and that required the respondents to link rows and columns in order to select an appropriate answer.

Statistical Analyses

We estimated the odds ratios (ORs) and 95% confidence intervals of breaking off the questionnaire according to person-level factors by using logistic regression with robust variance estimation to account for the correlation between the responses of mother who participated in the NINFEA cohort during more than 1 pregnancy.

To analyze the association of person-level and item design factors with item nonresponse, we used a 3-level hierarchical logistic regression model. The questionnaire items composed the first level, the questionnaire sections were the second level, and the women responding to the questionnaire were the third level. We fitted crude and adjusted models, by adjusting mutually for maternal age, university degree, employment status, gestational trimester, whether it was a first pregnancy, type of recruitment, region of residence, and number of participations.

As filters were used in the questionnaire, the total number of items to be responded to varied among participants. To account for these differences, we applied the inverse probability weighting (IPW) technique to deal with data missing by design [19]. In this study, we calculated the weights as the inverse of the probability of having a missing datum (by design) on every dependent item by considering only the women for whom that item was applicable. We estimated the weights using a logistic regression model that included the following person-level characteristics: age, university degree, gestational trimester at enrollment, whether it was a first pregnancy, employment status at the beginning of the pregnancy, and the type of recruitment. The underlying idea of IPW is to create weighted copies of the complete cases (dependent applicable items), according to selected person-level characteristics, to remove the selection bias introduced by the missing data. By doing so, we assumed that the nonresponse probability of women for whom the item was not applicable was equal to the nonresponse probability of women for whom the item was applicable, given that they had the same selected person-level characteristics. We did not truncate high-weight values, as, in sensitivity analyses, truncation at the 95th or 99th percentile did not affect the results more than marginally.

Analyses were conducted using the Stata 15.0 software (StataCorp LLC).

Results

Participant Characteristics

Table 1 lists the main characteristics of the 5970 women included in the analyses. Most of the NINFEA participants lived in the Piedmont Region, were recruited offline, and were in the third trimester of pregnancy. Two-thirds of women were younger than 35 years (n=4235), and more than half had a university degree (n=3605), were employed (n=5067), or were in their first pregnancy (n=3196). A total of 1176 women participated with more than 1 pregnancy in the NINFEA birth cohort.

Table 1. Characteristics of the study population (N=5970).

Participant characteristics	n (%)[a]
Age group (years)	
≤30	1735 (29.06)
31-35	2505 (41.96)
≥36	1730 (28.98)
University degree	
Yes	3605 (61.59)
No	2248 (38.41)
Employment status	
Unemployed	903 (15.13)
Employed	5067 (84.87)
Gestational trimester	
First	968 (16.41)
Second	1798 (30.48)
Third	3133 (53.11)
First pregnancy	
Yes	3196 (53.58)
No	2769 (46.42)
Type of recruitment	
Offline	4839 (83.71)
Online	942 (16.29)
Region of residence	
Piedmont Region	3328 (56.14)
Tuscany Region	1720 (29.01)
Other regions of North Italy	500 (8.43)
Other	380 (6.41)
Number of participations	
1	4794 (80.30)
≥2	1176 (19.70)

[a]Total numbers may vary due to missing values.

Questionnaire Breakoff and Item Nonresponse Characteristics

Table 2 shows the number of sections, item characteristics, and nonresponse percentage according to item design characteristics. We analyzed a total of 237 items from 14 sections in this study. Almost half of the items (n=116) were radio button type and included 3 to 5 response options. Of the 237 items, 39 (16.5%) were located in a grid. The highest nonresponse percentages among the applicable items were observed for filter questions, dropdown menu items, items containing 3 to 5 response options, and items located in grids.

Of the 5970 women, 519 (8.69%) did not finish the NINFEA baseline questionnaire. Breakoffs were spread over the 13 sections of the questionnaire. Table 3 shows the ORs of breakoff depending on the participants' characteristics. Women who at enrollment were in the first trimester of pregnancy had a threefold higher odds of questionnaire breakoff than did those who were in the third trimester of pregnancy (adjusted OR 3.01, 95% CI 2.31-3.92). Women without a university degree had 53% higher odds of questionnaire breakoff (95% CI 1.23-1.90) than did those with a higher education. Older age was also positively associated with questionnaire breakoff.

Table 2. Characteristics of the questionnaire items and frequency of nonresponse according to item characteristics.

Item characteristics	n (%)	Nonresponse, n (%)[a]
Sections	14	N/A[b]
Items	237	22,831 (2.15)
Filter question		
No	148 (62.4)	3900 (1.84)
Yes	89 (37.6)	18,931 (2.22)
Item type		
Checkbox	14 (5.9)	804 (1.48)
Dropdown menu	49 (20.7)	7454 (2.84)
Radio button	116 (48.9)	12,335 (2.17)
Text (open question)	58 (24.5)	2238 (1.26)
Number of response options[c]		
2	69 (38.6)	7606 (2.20)
3-5	85 (47.4)	11,827 (2.65)
≥6	25 (14.0)	1160 (1.27)
Item in a grid		
No	198 (83.5)	16,625 (1.96)
Yes	39 (16.5)	6206 (2.92)

[a]Calculated as the ratio between the total number of items not responded to and the total number of applicable items (n=1,062,519) for all participants.

[b]N/A: not applicable.

[c]Text items were not considered.

Table 3. Questionnaire breakoff according to participants' characteristics.

Participant characteristics	n (%)	Crude analyses, OR[a] (95% CI)	Adjusted analyses[b], OR (95% CI)
Age group (years)			
≤30	137 (7.9)	1.00	1.00
31-35	213 (8.5)	1.08 (0.87-1.36)	1.11 (0.84-1.44)
≥36	169 (9.8)	1.26 (1.00-1.60)	1.40 (1.05-1.88)
University degree			
Yes	220 (6.1)	1.00	1.00
No	206 (9.2)	1.55 (1.27-1.90)	1.53 (1.23-1.90)
Employment status			
Employed	363 (7.2)	1.00	1.00
Unemployed	156 (17.3)	2.71 (2.21-3.32)	0.99 (0.73-1.34)
Gestational trimester			
Third	189 (6.0)	1.00	1.00
Second	134 (7.5)	1.25 (1.00-1.58)	1.27 (0.98-1.65)
First	170 (17.6)	3.32 (2.65-4.15)	3.01 (2.31-3.92)
First pregnancy			
Yes	233 (7.3)	1.00	1.00
No	286 (10.3)	1.47 (1.22-1.76)	1.13 (0.90-1.43)
Type of recruitment			
Offline	389 (8.0)	1.00	1.00
Online	107 (11.4)	1.47 (1.17-1.84)	1.11 (0.82-1.51)
Region of residence			
Piedmont Region	236 (7.1)	1.00	1.00
Tuscany Region	170 (9.9)	1.44 (1.17-1.77)	1.06 (0.84-1.35)
Other regions of North Italy	49 (9.8)	1.42 (1.03-1.97)	1.14 (0.75-1.73)
Other	54 (14.2)	2.17 (1.58-2.99)	1.80 (1.21-2.66)
Number of participations			
1	387 (8.1)	1.00	1.00
≥2	132 (11.2)	1.44 (1.17-1.77)	1.19 (0.91-1.57)

[a]OR: odds ratio.

[b]Models adjusted for age, university degree, employment status, gestational trimester, first pregnancy, type of recruitment, region, and number of participations.

Of the 1,062,519 applicable items, 22,831 were not responded to, giving an overall item nonresponse rate of 2.15%. Table 4 presents the weighted crude and adjusted ORs of item nonresponse according to participants' characteristics. Similar to the findings for questionnaire breakoff, lower educational level, older age, and enrollment in the first trimester of pregnancy were positively associated with item nonresponse. In contrast, participating during 2 or more pregnancies (ie, responding to the questionnaires twice or more often) was associated with lower odds of item nonresponse. Number of pregnancies, employment status, and type of recruitment were not associated with item nonresponse in our study.

All the analyzed item design factors were associated with item nonresponse (Table 5). Items designed as a dropdown menu were 77% more likely to be left blank than were radio button items (95% CI 1.56-2.00). Text items had 30% lower odds of item nonresponse (95% CI 0.63-0.79) and checkboxes had 80% lower odds of item nonresponse (95% CI 0.16-0.25) than did radio button items. Items with 6 or more response options were 59% less likely to be left blank than were those with 2 response options (95% CI 0.35-0.47). Finally, items being located in a grid was positively associated with nonresponse (adjusted OR 1.69, 95% CI 1.49-1.91).

Table 4. Prevalence and crude and adjusted odds ratios (ORs) of item nonresponse according to participants' characteristics.

Participant characteristics	Prevalence (%)	Crude analyses, OR (95% CI)	Adjusted analyses[a], OR (95% CI)
Age group (years)			
≤30	2.1	1.00	1.00
31-35	2.0	1.03 (0.95-1.13)	1.07 (0.98-1.17)
≥36	2.4	1.25 (1.14-1.38)	1.25 (1.14-1.38)
University degree			
Yes	1.9	1.00	1.00
No	2.4	1.22 (1.14-1.31)	1.23 (1.14-1.33)
Employment status			
Employed	2.0	1.00	1.00
Unemployed	3.0	0.89 (0.78-1.01)	0.87 (0.77-0.98)
Gestational trimester			
Third	2.0	1.00	1.00
Second	2.1	1.04 (0.96-1.12)	1.00 (0.93-1.09)
First	2.6	1.17 (1.06-1.29)	1.18 (1.06-1.31)
First pregnancy			
Yes	2.2	1.00	1.00
No	2.1	1.05 (0.98-1.12)	1.03 (0.95-1.11)
Type of recruitment			
Offline	2.1	1.00	1.00
Online	2.4	1.12 (1.01-1.23)	1.07 (0.96-1.18)
Region of residence			
Piedmont Region	1.9	1.00	1.00
Tuscany Region	2.5	1.17 (1.08-1.27)	1.16 (1.07-1.25)
Other regions of North Italy	1.9	1.02 (0.90-1.15)	0.97 (0.85-1.11)
Other	2.8	1.37 (1.16-1.61)	1.14 (0.98-1.34)
Number of participations			
1	2.2	1.00	1.00
≥2	1.9	0.84 (0.77-0.92)	0.90 (0.82-0.99)

[a]Models adjusted for age, university degree, employment status, gestational trimester, first pregnancy, type of recruitment, region, and number of participations.

Table 5. Crude and adjusted odds ratios (ORs) of item nonresponse according to item design factors.

Item design factors	Crude analyses, OR (95% CI)	Adjusted analyses[a], OR (95% CI)
Item type		
Radio button	1.00	1.00
Checkbox	0.20 (0.17-0.25)	0.20 (0.16-0.25)
Dropdown menu	1.73 (1.53-1.94)	1.77 (1.56-2.00)
Text (open question)	0.70 (0.63-0.78)	0.70 (0.63-0.79)
Response options		
2	1.00	1.00
3-5	1.12 (1.04-1.21)	1.09 (1.01-1.18)
≥6	0.41 (0.35-0.47)	0.41 (0.35-0.47)
Item in a grid		
No	1.00	1.00
Yes	1.63 (1.44-1.83)	1.69 (1.49-1.91)

[a]Models adjusted for age, university degree, employment status, gestational trimester, first pregnancy, type of recruitment, region, and number of participations.

Discussion

Principal Findings

Our results showed that women enrolled in earlier stages of pregnancy had a higher probability of questionnaire breakoff than did women enrolled in the third trimester of pregnancy. Older and less-educated women were more likely to break off the questionnaire and to leave items blank. Dropdown menu items were associated with the lowest response rate among all types of items. Unexpectedly, text items were less likely to be left blank than were radio button items; similarly, items with 6 or more response options were less likely to be left blank than were those with 2 response options.

Our findings of higher breakoff and item nonresponse rates among women in the first trimester of pregnancy than among those enrolled in the third trimester could be explained by several factors, including participants' time available to answer the questionnaire. Women in later stages of pregnancy might have more time to complete the questionnaire, as they are already on maternity leave. Lower educational level was positively associated with questionnaire breakoff in the NINFEA Web-based cohort. This finding is consistent with other studies that included different populations (eg, men) [20,21] or used different data collection methods, such as postal questionnaires [22]. These consistencies are of particular interest, as the NINFEA study population includes self-selected volunteers having access to the internet; nevertheless, differences in completion of the questionnaire by educational level persist. Thus, regardless of the population or data collection method, epidemiologic studies that rely on self-administered questionnaires should identify incentives to motivate participation, specifically of individuals with low educational levels.

In contrast, there are determinants that are closely related to Web-based studies, such as whether the participants became aware of the study through online or offline channels. Few studies have investigated the associations between the type of recruitment and breakoff from internet-based questionnaires [23]. Our finding of no association is in line with the findings of an internet-based intervention that found no difference in questionnaire breakoff between online and offline recruitment methods [24].

The proportion of item nonresponse was low in our study, ranging from 1.3% to 2.9%. Another study that administered daily Web-based questionnaires also described low rates of item nonresponse, ranging from 0% to 7.4% [25]. In our study, online recruitment, older age, and lower educational levels were positively associated with item nonresponse. This is in line with findings of 3 quality-of-life Web-based surveys conducted in the United States [26]. The association between older age and lower educational levels with higher rates of item nonresponse is also consistent with other prior work [27,28]. Regardless of the data collection method used, these individuals have to expend a higher cognitive effort to respond to questions. In the case of a self-reported questionnaire responded to over the internet (with no support from an interviewer), the rates of nonresponse for these individuals can be even higher.

The number of times a woman participated in the NINFEA baseline questionnaire was not associated with breakoff, but it was associated with lower rates of item nonresponse. However, the confidence interval almost included the unit, and for this reason we believe this association might be due to residual confounding.

To analyze item nonresponse according to the type of item, we compared all items with the radio button items, since this was the most prevalent item in the NINFEA questionnaire. Our finding that checkbox items were associated with a lower item nonresponse than the radio button items is consistent with the literature and inherent in the logic of checkboxes [26,29]. The probability of checking at least 1 answer among several response options is likely higher than checking 1 answer among a pair of response options [29]. Our finding of lower item nonresponse

among items with 6 or more response options than among items with 2 response options supports this hypothesis. Text items were associated with a higher response than were radio button items in our study. The association of text items with item nonresponse is still controversial in the literature, as studies found text items to be positively or negatively associated with item nonresponse [26,30]. Dropdown menu items were positively associated with item nonresponse, as they require more actions to select an answer (3 actions for dropdown menu items vs 1 action for radio button items), and this can explain the higher item nonresponse rate [11,31].

As expected, items located in grids had higher odds of item nonresponse than did single items. Linking rows and columns of a grid to select an appropriate answer is more complex than choosing an answer of a single item; hence, if possible, grid items should be avoided [32,33].

Besides the design of the items, their content could also influence item nonresponse [26]. For instance, items asking about sensitive subjects could have higher nonresponse than items with nonsensitive content [34]. However, we did not perceive this behavior in our study. In the NINFEA baseline questionnaire, we considered only 3 of the 237 items to have sensitive content: alcohol consumption during pregnancy, use of soft drugs during pregnancy, and smoking during pregnancy. There were no missing responses for the first 2 items and 9 missing responses for the item asking about smoking.

Conclusion

We obtained our findings within the context of a longitudinal epidemiologic study: the NINFEA Web-based birth cohort. In this type of study, it is very important to avoid breakoffs and item nonresponse, since the presence of missing values in the baseline questionnaires makes analyses of future outcomes difficult. Using the IPW technique and multilevel modeling, we were able to comprehensively and concurrently analyze the association of person-level and item design factors with item nonresponse. By doing so, we were also able to adjust all analyses for the characteristics of the mothers.

To our knowledge, this is the first study evaluating determinants of questionnaire breakoff and item nonresponse in the context of e-epidemiology. Our study was based on only 1 internet-based epidemiologic study and included only pregnant women; thus, replications in other populations and settings are needed. It is crucial to understand the profile of nonresponders to develop personalized motivation methods and minimize item nonresponse and breakoffs. Personalized recruitment [35,36], use of reminders [37,38], incentives [39,40], and gamification [41] are only some of the strategies that can be used to keep participants motivated.

The low percentage of breakoffs in the baseline questionnaire of the NINFEA birth cohort demonstrates the feasibility of e-epidemiologic research, even when long questionnaires are applied. However, the questionnaires should be designed carefully. For instance, items with 1 and several radio button options should replace dropdown menu items and items located in grids, respectively, in order to reduce nonresponse. Also, we showed several person-level characteristics to be important determinants of breakoff and item nonresponse in internet-based questionnaires. For this reason, study coordinators should know their target population so as to employ focused motivation and recruitment techniques and to reduce breakoff and item nonresponse. Older and less educated individuals should be contacted directly (even by other means, such as telephone) in order to assist and encourage their participation in e-epidemiologic research.

Acknowledgments

This work was conducted during a PhD exchange period at the University of Turin, Italy, sponsored by the Capes Foundation within the Ministry of Education, Brazil (grant #88881.133234/2016-01).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of item characteristics.

[PDF File (Adobe PDF File), 68KB - jmir_v20i12e11046_app1.pdf]

References

1. Ekman A, Litton J. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
2. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
3. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006 Feb 01;163(3):197-203. [doi: 10.1093/aje/kwj036] [Medline: 16339049]
4. Sinclair M, O'Toole J, Malawaraarachchi M, Leder K. Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. BMC Med Res Methodol 2012;12:132 [FREE Full text] [doi: 10.1186/1471-2288-12-132] [Medline: 22938205]

5. van Gelder MM, Pijpe A. E-epidemiology: a comprehensive update. OA Epidemiol 2013 Jun 04;1(1). [doi: 10.13172/2053-079X--1-584]
6. Blumenberg C, Barros AJD. Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature. Int J Public Health 2018 Jul;63(6):765-773. [doi: 10.1007/s00038-018-1108-4] [Medline: 29691594]
7. Groves R, Dillman DA, Little RJA, Eltinge JL. Survey Nonresponse. New York, NY: John Wiley & Sons; 2001.
8. Peytchev A. Survey breakoff. Public Opin Q 2009 Apr 02;73(1):74-97. [doi: 10.1093/poq/nfp014]
9. Phillips AW, Reddy S, Durning SJ. Improving response rates and evaluating nonresponse bias in surveys: AMEE guide no. 102. Med Teach 2016 Dec;38(3):217-228. [doi: 10.3109/0142159X.2015.1105945] [Medline: 26648511]
10. Liu M, Wang Y. Comparison of face-to-face and web surveys on the topic of homosexual rights. J Homosex 2016 Jun;63(6):838-854. [doi: 10.1080/00918369.2015.1112587] [Medline: 26566766]
11. Dillman DA, Smyth JD, Christian LM. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, NJ: John Wiley & Sons; 2014.
12. Steinbrecher M, Rossmann J, Blumenstiel J. Why do respondents break off web surveys and does it matter? Results from four follow-up surveys. Int J Public Opin Res 2014 Sep 09;27(2):289-302. [doi: 10.1093/ijpor/edu025]
13. Rothman KJ, Mikkelsen EM, Riis A, Sørensen HT, Wise LA, Hatch EE. Randomized trial of questionnaire length. Epidemiology 2009 Jan;20(1):154. [doi: 10.1097/EDE.0b013e31818f2e96] [Medline: 19234404]
14. Richiardi L, Baussano I, Vizzini L, Douwes J, Pearce N, Merletti F. Feasibility of recruiting a birth cohort through the internet: the experience of the NINFEA cohort. Eur J Epidemiol 2007;22(12):831-837. [doi: 10.1007/s10654-007-9194-2] [Medline: 17955333]
15. Progetto NINFEA. NINFEA project. Turin, Italy: Azienda Ospedaliero-Universitaria, Citta della Salute e della Scienza di Torino; 2018. URL: https://www.progettoninfea.it/index_en [accessed 2018-11-05] [WebCite Cache ID 73i3f4rVu]
16. The PHP Group. PHP. 2015. URL: http://php.net/ [accessed 2015-05-14] [WebCite Cache ID 6YWJTpCvt]
17. Ruby Community. Ruby is.... 2018. URL: http://www.ruby-lang.org [accessed 2018-05-08] [WebCite Cache ID 6zGfJEOHB]
18. Richiardi L, Pivetta E, Merletti F. Recruiting study participants through Facebook. Epidemiology 2012 Jan;23(1):175. [doi: 10.1097/EDE.0b013e31823b5ee4] [Medline: 22157313]
19. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. Stat Methods Med Res 2013 Jun;22(3):278-295. [doi: 10.1177/0962280210395740] [Medline: 21220355]
20. Ekman A, Klint A, Dickman PW, Adami H, Litton J. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. Eur J Epidemiol 2007;22(5):293-300. [doi: 10.1007/s10654-006-9091-0] [Medline: 17206467]
21. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]
22. Stenhammar C, Bokström P, Edlund B, Sarkadi A. Using different approaches to conducting postal questionnaires affected response rates and cost-efficiency. J Clin Epidemiol 2011 Oct;64(10):1137-1143. [doi: 10.1016/j.jclinepi.2011.01.006] [Medline: 21524566]
23. Lane TS, Armin J, Gordon JS. Online recruitment methods for web-based and mobile health studies: a review of the literature. J Med Internet Res 2015;17(7):e183 [FREE Full text] [doi: 10.2196/jmir.4359] [Medline: 26202991]
24. Heffner JL, Wyszynski CM, Comstock B, Mercer LD, Bricker J. Overcoming recruitment challenges of web-based interventions for tobacco use: the case of web-based acceptance and commitment therapy for smoking cessation. Addict Behav 2013 Oct;38(10):2473-2476 [FREE Full text] [doi: 10.1016/j.addbeh.2013.05.004] [Medline: 23770645]
25. Romano MF, Sardella MV, Alboni F. Web health monitoring survey: a new approach to enhance the effectiveness of telemedicine systems. JMIR Res Protoc 2016 Jun 06;5(2):e101 [FREE Full text] [doi: 10.2196/resprot.5187] [Medline: 27268949]
26. Messer BL, Edwards ML, Dillman DA. Determinants of item nonresponse to web and mail respondents in three address-based mixed-mode surveys of the general public. Surv Pract 2012 Apr 01;5(2):1-9. [doi: 10.29115/SP-2012-0012]
27. de Leeuw E. Reducing missing data in surveys: an overview of methods. Qual Quant 2001 May;35(2):147-160. [doi: 10.1023/A:1010395805406]
28. Yan T, Curtin R. The relation between unit nonresponse and item nonresponse: a response continuum perspective. Int J Public Opin Res 2010 Oct 25;22(4):535-551. [doi: 10.1093/ijpor/edq037]
29. Smyth JD. Comparing check-all and forced-choice question formats in web surveys. Public Opin Q 2006 Mar 01;70(1):66-77. [doi: 10.1093/poq/nfj007]
30. Denscombe M. Item non‐response rates: a comparison of online and paper questionnaires. Int J Soc Res Methodol 2009 Oct;12(4):281-291. [doi: 10.1080/13645570802054706]
31. Bethlehem J, Biffignandi S. Handbook of Web Surveys. Hoboken, NJ: John Wiley & Sons; 2011.
32. Liu M, Cernat A. Item-by-item versus matrix questions. Soc Sci Comput Rev 2016 Nov 30. [doi: 10.1177/0894439316674459]
33. Couper MP, Tourangeau R, Conrad FG, Zhang C. The design of grids in web surveys. Soc Sci Comput Rev 2013 Jun;31(3):322-345 [FREE Full text] [doi: 10.1177/0894439312469865] [Medline: 25258472]

34. Jones MK, Calzavara L, Allman D, Worthington CA, Tyndall M, Iveniuk J. A comparison of web and telephone responses from a national HIV and AIDS survey. JMIR Public Health Surveill 2016 Jul 29;2(2):e37 [FREE Full text] [doi: 10.2196/publichealth.5184] [Medline: 27473597]
35. Muñoz-Leiva F, Sánchez-Fernández J, Montoro-Ríos F, Ibáñez-Zapata J. Improving the response rate and quality in Web-based surveys through the personalization and frequency of reminder mailings. Qual Quant 2009 Jun 6;44(5):1037-1052. [doi: 10.1007/s11135-009-9256-5]
36. Short CE, Rebar AL, Vandelanotte C. Do personalised e-mail invitations increase the response rates of breast cancer survivors invited to participate in a web-based behaviour change intervention? A quasi-randomised 2-arm controlled trial. BMC Med Res Methodol 2015 Aug 19;15:66 [FREE Full text] [doi: 10.1186/s12874-015-0063-5] [Medline: 26286486]
37. Van Mol C. Improving web survey efficiency: the impact of an extra reminder and reminder content on web survey response. Int J Soc Res Methodol 2016 May 17;20(4):317-327. [doi: 10.1080/13645579.2016.1185255]
38. Dokkum NFB, Koekenbier RH, van den Broek IV, van Bergen JE, Brouwers EE, Fennema JS, et al. Keeping participants on board: increasing uptake by automated respondent reminders in an internet-based Chlamydia screening in the Netherlands. BMC Public Health 2012;12:176 [FREE Full text] [doi: 10.1186/1471-2458-12-176] [Medline: 22404911]
39. Singer E, Ye C. The use and effects of incentives in surveys. Ann Am Acad Polit Soc Sci 2012 Nov 26;645(1):112-141. [doi: 10.1177/0002716212458082]
40. Olsen F, Abelsen B, Olsen JA. Improving response rate and quality of survey data with a scratch lottery ticket incentive. BMC Med Res Methodol 2012 Apr 19;12:52 [FREE Full text] [doi: 10.1186/1471-2288-12-52] [Medline: 22515335]
41. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Dec;71:31-48. [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]

Abbreviations

IPW: inverse probability weighting
NINFEA: Nascita e Infanzia: gli Effetti dell'Ambiente
OR: odds ratio

Edited by G Eysenbach; submitted 14.05.18; peer-reviewed by JM Carrillo de Gea, V Andreeva; comments to author 05.07.18; revised version received 22.08.18; accepted 25.08.18; published 07.12.18.

Please cite as:
Blumenberg C, Zugna D, Popovic M, Pizzi C, Barros AJD, Richiardi L
Questionnaire Breakoff and Item Nonresponse in Web-Based Questionnaires: Multilevel Analysis of Person-Level and Item Design Factors in a Birth Cohort
J Med Internet Res 2018;20(12):e11046
URL: https://www.jmir.org/2018/12/e11046/
doi:10.2196/11046
PMID:30530454

Original Paper

Mobile Health, Information Preferences, and Surrogate Decision-Making Preferences of Family Caregivers of People With Dementia in Rural Hispanic Communities: Cross-Sectional Questionnaire Study

Bo Xie[1,2], BS, Med, MSc, PhD; Jane Dimmitt Champion[1], DNP, PhD; Jung Kwak[1], PhD; Kenneth R Fleischmann[2], PhD

[1]School of Nursing, The University of Texas at Austin, Austin, TX, United States

[2]School of Information, The University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Bo Xie, BS, Med, MSc, PhD
School of Nursing
The University of Texas at Austin
1710 Red River Street
Austin, TX, 78712
United States
Phone: 1 512 232 5788
Email: boxie@utexas.edu

Abstract

Background: Mobile health (mHealth) technology holds promise for promoting health education and reducing health disparities and inequalities in underserved populations. However, little research has been done to develop mHealth interventions for family caregivers of people with dementia, particularly those in rural Hispanic communities, who often serve as surrogate decision makers for their relatives with dementia.

Objective: As part of a larger project to develop and test a novel, affordable, and easy-to-use mHealth intervention to deliver individually tailored materials in rural Hispanic communities, in this pilot study, we aimed to examine (1) characteristics of people with dementia and their family caregivers in rural Hispanic communities, (2) caregivers' preferences for types and amounts of health information and participation in surrogate decision making, and (3) caregivers' mobile device usage and their desire for receiving information via mobile devices.

Methods: This was a cross-sectional survey. A convenience sample of 50 caregivers of people with dementia was recruited from rural health care facilities in Southwest Texas during 3 weeks of April 2017 to May 2017 via word-of-mouth and flyers posted at the facilities.

Results: More women than men were in the patient group (χ^2_1=17.2, *P*<.001) and in the caregiver group (χ^2_1=22.2, *P*<.001). More patients were on Medicare and Medicaid; more caregivers had private insurance (*P*<.001 in all cases). Overall, 42% of patients did not have a power of attorney for their health care; 40% did not have a living will or advance directive. Caregivers were interested in receiving all types of information and participating in all types of decisions, although on subscales for diagnosis, treatment, laboratory tests, self-care, and complementary and alternative medicine, their levels of interest for decision-making participation were significantly lower than those for receiving information. On the psychosocial subscale, caregivers' desire was greater for surrogate decision-making participation than for information. Caregivers did not differ in their interests in information and participation in decision making on the health care provider subscale. All but 1 caregiver (98%) owned a mobile phone and 84% had a smartphone. Two-thirds wanted to receive at least *a little* dementia-related information via a smartphone or tablet. The amount of dementia-related information caregivers wanted to receive via a mobile device was significantly greater for women than for men (*U*=84.50, *P*=.029). Caregivers who owned a tablet were more likely to want to receive dementia-related information via a mobile device than those who did not own a tablet (*U*=152.0, *P*=.006).

Conclusions: Caregivers in rural Hispanic communities were interested in receiving a wide range of information as well as participating in making decisions for their relatives with dementia. There is much need for effective mHealth interventions that can provide information tailored to the needs and preferences of these caregivers.

(J Med Internet Res 2018;20(12):e11682) doi:10.2196/11682

KEYWORDS

mHealth; dementia; caregivers; rural communities

Introduction

Background

Mobile health (mHealth) technology has become an important tool for accessing health information, particularly among ethnic minorities; this new phenomenon presents ample opportunities for health researchers, practitioners, and educators to develop and implement health education interventions to improve health literacy and reduce health disparities and health inequities among ethnic minority groups [1]. mHealth has been used, alone or in combination with a traditional approach, to support health education or self-management for a wide range of health conditions, such as eating disorders [2], multiple sclerosis [3], cardiovascular disease [4], HIV [5], and mental illnesses [6], to name just a few. mHealth interventions have also been tested in a variety of age groups ranging from older adults [7] to pregnant or postpartum women [8] to young adults [9]. Preliminary evidence shows promise for the use of mHealth in chronic disease self-management and for improvements in many physical conditions; however, more systematic research is still needed to generate solid evidence for the efficacy of mHealth-based interventions [10].

Many mHealth interventions have targeted patients, but relatively few have focused on caregivers, and the latter have tended to focus on caregivers of children or youth [11-13]. Our own systematic review suggests that few mHealth interventions have been developed for family caregivers (hereafter *caregivers*) of people with dementia [14], with only a handful of exceptions published within the last few years [15-19].

The Need to Support Dementia Caregivers' Decision Making

Dementia has become a major public health concern worldwide. It is estimated that every 3 seconds someone somewhere in the world develops dementia [20]. Nearly 50 million people worldwide were estimated to be living with dementia in 2017, and this number is expected to reach 131.5 million by midcentury [20]. In the United States, Alzheimer disease, which represents the majority of dementia cases, has become the sixth leading cause of death overall and the fifth leading cause of death in older Americans aged 65 years and above [21]. The number of American people living with Alzheimer disease and related dementias (ADRD) is estimated to be 5.7 million in 2018, and this number is expected to increase to 13.8 million in 2050 [21]. The nature of this condition requires extensive care for people with dementia: it is estimated that in 2017, over 16 million informal caregivers in the United States, most of whom were family members, provided 18.4 billion hours of care [21].

Decision making in the treatment and care of people with dementia falls mostly on caregivers, who are expected to make informed decisions in the patient's best interest. However, caregivers often report being unprepared for their roles and responsibilities, uninformed about treatment options, uncertain about patients' preferences, and unsupported by professionals in their decision making [22-25]. A major challenge for caregivers is to obtain relevant information about treatment and care options so that they can evaluate the relative merits and risks of each option before making decisions [23,26]. Caring for people with dementia increases risks for caregivers' mental and physical well-being and deserves much attention [21]. In recognition of the need to support patients and families in making end-of-life (EOL) decisions and to improve EOL care, national projects and federal agencies such as the National Institutes of Health have prioritized advancing current knowledge about EOL decision making and developing innovative decision support interventions for patients with terminal illnesses and their families [27-29]. Given the projected growth of the population with advanced ADRD over the next 50 years [30], the significance of research designed to support caregivers in making EOL care decisions for their relatives will continue to grow.

Gaps in Existing Interventions for Dementia Caregivers

Interventions supporting caregivers' decision making are only beginning to emerge; in our recent systematic literature search [14], we found 5 published studies of decision aids for American caregivers in the last 10 years. These decision aids provided caregivers with information about treatment options, but our review identified major knowledge gaps: (1) all study samples were predominately white, (2) existing research has paid little attention to caregivers in rural areas, and (3) existing interventions have included no technology other than audio or video. Thus, no intervention has taken full advantage of recent technological developments to enable the provision of electronic contents tailored to caregivers' preferences for different types and amounts of information and participation in decision making [14].

These knowledge gaps must be addressed for several reasons. First, for people with dementia and their caregivers of racial or minority backgrounds, there may be special challenges to engaging in advance care planning or in accessing adequate EOL care; the literature has consistently documented cultural differences and disparities at EOL. A systematic review [31] has found that people with dementia from certain ethnic minority groups prefer different EOL treatments and are less likely to have advance directives because of disparities and differences in cultural values. African Americans, for example, are more likely to choose life-sustaining treatment than non-Hispanic whites based on factors such as fear that providers would undertreat, gaps in information and knowledge, and differences in cultural evaluations of the benefits and risks of some care options [31]. The low rate of advance care planning among various ethnic groups (eg, Hispanics, Japanese, Koreans, Chinese, and American Indians) has been attributed to cultural aversion to direct communication about serious illnesses and

poor prognosis as well as to preferences for group consensus and the family as a decision-making unit [32]. Another literature review [33] has reported differences in the experiences of caregivers from ethnic minority groups, including higher levels of depression and stress among Hispanic caregivers than among non-Hispanic whites, as well as different coping mechanisms. A meta-analysis found that people with dementia from ethnic minority groups were less able to access health and social services [34].

Second, approximately 15% of the US population, 46 million, lives in rural counties [35]. Rural caregivers face unique challenges [36]. Rural residents tend to be poorer, older, and sicker than their urban counterparts [37,38]. Rural services are often spread over long distances, and the cost of transportation and time often drastically decrease their use [39,40]. Rural nursing homes often lack a diversity of health services or health care professionals for people with dementia [41]. Fewer local health services and providers are available, including palliative care and hospice services [42-44]. Rural caregivers have fewer formal services for support and often rely more on informal services [39] and report greater financial burden [45]. Other unique barriers to the use of formal services include stigma of dementia, lack of privacy, beliefs and attitudes, lack of awareness of services, and less acceptability and accessibility of services [46]. Rural caregivers face different expectations of help and support than urban caregivers do: taking care of a family member with dementia might be seen as a part of life or a family responsibility rather than work, and an inability to provide help for a relative with dementia is more likely to be perceived as abandonment of a relative [45]. Moreover, one consistently identified need of caregivers in rural areas is the need for counseling and mental health services [41]. The coping styles of rural caregivers often differ from those of their urban counterparts, suggesting unique needs [47]. These characteristics and health disparities between urban and rural areas call for effective interventions tailored to the unique needs and circumstances of rural communities and caregivers.

Third, shared decision making and patient-centered care require serious attention to individual preferences [48-50]. However, existing research in this area has examined individual preferences for different types and amounts of health information and decision-making participation mainly from *the health care provider's perspective* —what *providers* think their patients need to know (typically to ensure compliance). Its focus is typically on a limited range of information and decision making (eg, information and decisions related to treatment). Preferences for other important types of information (eg, how to cope psychosocially) and decision making (eg, choosing which provider to go to) are understudied [51-54]. This trend has continued in interventions involving the use of mHealth technology, with existing interventions showing little consideration for individual preferences for the types and amounts of information received via mHealth. Of the systematic reviews we have examined [55-58], *how often* to receive messages) and 22% accommodated preferences for timing (*when* to receive messages) [56]. Such a trend is unfortunate because meta-analyses provide strong evidence that tailored health behavioral interventions outperform nontailored ones [59-61]; recent developments in mHealth offer unprecedented opportunities for providing tailored health behavioral interventions to hard-to-reach populations [62]. Research is much needed to help caregivers take advantage of new opportunities afforded by mHealth so that they can be better prepared to make informed decisions for their relatives.

Study Aims and Research Questions

This pilot study was part of a larger study plan to develop and test a novel, affordable, easy-to-use mHealth intervention to deliver individually tailored materials to rural Hispanic communities. We chose to focus on Hispanics because they are the second largest ethnic group after non-Hispanic whites and the fastest growing ethnic minority group in the United States and because they are overlooked in existing intervention studies for caregivers [14]. Our long-term goal is to help caregivers make good use of new technological advancements to be better prepared for the wide range of future care needs and care transitions for their relatives. Toward this end, we conducted our pilot study as a first step to understand community needs and determine the feasibility of the planned larger scale mHealth intervention. Specific aims of the pilot study were to understand (1) the characteristics of people with dementia and their family caregivers in rural Hispanic communities, (2) caregivers' preferences for different types and amounts of health information and decision-making participation, and (3) caregivers' mobile device usage and their desire for receiving information via mobile devices.

The primary research questions for this pilot study were as follows:

1. What are the main characteristics of people with dementia and their caregivers in rural communities?
2. What are caregivers' preferences for overall decision making in the family and for specific types of health information and decision-making participation?
3. What are caregivers' mobile device usage and desire for receiving information via mobile devices?

Methods

Design

This was a cross-sectional survey study.

Participants

A convenience sample of 50 caregivers was recruited from rural health care facilities in Southwest Texas. These facilities provide health care services, including services for people with dementia, for a 5-county rural area near the US-Mexico border. Participants were recruited during a 3-week period in April 2017 to May 2017 via word-of-mouth and flyers posted at the health care facilities. One of the researchers on the team, a family nurse practitioner who has been practicing in this rural community for over 30 years, identified community stakeholders, obtained permission to access facilities and post flyers for the study, and conducted the participant recruitment and data collection at the facilities. To be eligible, participants had to (1) be aged 18 years or older, (2) be able to read and write in English, and (3) self-identify as a family caregiver or have been caring for a

relative with dementia or memory problems by assisting with any activities of daily living (ADL) for at least 2 years. No one refused to participate in the study.

Procedure

Participants completed a survey instrument on paper while visiting a facility. Completion took approximately 20 to 25 min. Informed consent was obtained before any data collection. Each participant received a US $10 gift card after completing the instrument. The study was approved by the institutional review board of the authors' institution.

Materials

The instrument included the following:

- Demographics: 27 items about the patient and 8 items about the caregiver.
- ADL: 6 items, each item scored 1 to 4 with a scoring range of 6 to 24; the higher the score, the more dependent the relative.
- Instrumental Activities of Daily Living: 10 items, each item scored 1 to 4 with a scoring range of 10 to 40; the higher the score, the more dependent the relative;
- Health Information Wants Questionnaire (HIWQ): Preferences for health information and decision-making participation [63-66]; the 21-item HIWQ is a validated, self-administered instrument. It includes 2 parallel scales: the Information Preference Scale (IPS) and the Decision-making Preference Scale (DPS). Each scale contains 7 subscales with parallel items in 7 areas: diagnosis (4 items), treatment (3 items), laboratory tests (3 items), self-care (3 items), complementary and alternative medicine (CAM; 3 items), psychosocial aspects (3 items), and health care providers (2 items). On the IPS, participants indicate how much information they would like to have regarding each of the 7 health-related areas on a 5-point Likert scale (1=none, 2=a little, 3=some, 4=most, and 5=all). On the DPS, participants indicate their preferences for participation in each of the 7 parallel types of decision making on a 5-point Likert scale (1=the doctor alone, 2=mostly the doctor, 3=the doctor and myself equally, 4=mostly myself, and 5=myself alone).
- Technology usage: caregivers' cell phone and tablet usage and desire for receiving health information via mobile devices; 6 items.

Data Rescoring and Analysis Strategies

Data were entered into an IBM SPSS file by a research assistant (RA). A second RA independently evaluated the data for accuracy, missing data, and out-of-range values. With guidance from both an experienced biostatistician and the first author, any errors or discrepancies in the data were corrected. Descriptive statistics were used to provide a statistical profile of the sample, reporting frequencies and percentages for categorical data and means and SDs for continuous data. Paired sample *t* test and nonparametric tests (Chi-square and McNemar's) were used to compare the basic demographic characteristics of the patients and their caregivers. The original subscale scores of the HIWQ were calculated as means across relevant items. Using rescoring strategies that we had used in previous HIWQ studies [63-66], we rescaled the original scores to have a mean of 50 and range from 0 to 100 (100=the strongest desire for information or decision-making participation; 0=no desire). Correlational analyses (Spearman tests) were conducted, and Mann-Whitney tests determined whether there were significant differences between groups (with the dependent variables being at least ordinal).

Results

Main Demographic Characteristics of People With Dementia and Their Caregivers in Rural Hispanic Communities

Basic demographic characteristics (age, gender, education, race or ethnicity, and health insurance coverage) of the patients and their caregivers are presented in Table 1. Other key characteristics of the patients as reported by their caregivers are presented in Table 2. Caregivers were significantly younger than the patients: t_{45}=13.126, P<.001. More women than men were in the patient group (χ^2_1=17.2, P<.001) and in the caregiver group (χ^2_1=22.2, P<.001). The patient and caregiver groups did not differ in their group compositions in gender, race or ethnicity, or college or no college degree. More patients were on Medicare and Medicaid, whereas more caregivers had private insurance (P<.001 in all cases).

Caregivers' Preferences for Overall Decision Making in the Family and for Specific Types of Health Information and Decision-Making Participation

Caregivers' general decision-making patterns in the family and their expectations for who, in general, should make decisions related to their relative's condition are illustrated in Table 3. Caregivers' preferences for specific types of health information and decision-making participation are illustrated in Table 4. Caregivers had much interest in all 7 types of information. They also were interested in participating in all 7 types of decision making, although their levels of interest in surrogate decision-making participation were significantly less than their interests in receiving information on 5 of the 7 subscales: diagnosis, treatment, laboratory tests, self-care, and CAM. On the psychosocial subscale, caregivers' desire for decision-making participation was greater than that for information. Caregivers did not differ in their interests in information and decision-making participation on the health care provider subscale (Table 4). Mann-Whitney tests found no significant difference between women and men, Hispanics and whites, smartphone owners and nonowners, or tablet owners and nonowners in the amounts of specific types of information or decision-making participation they wanted.

Table 1. Basic demographic characteristics.

Variable	Patient	Caregiver
Age (years)		
Mean (SD)	78.96 (9.29)	52.62 (13.78)
Median (range)	81.00 (60-95)	53.00 (27-85)
Gender, n (%)		
Female	39 (78)	41 (82)
Education, n (%)		
8th grade (middle school) or less	24 (48)	4 (8)
Attended high school	2 (4)	3 (6)
Completed high school	6 (12)	7 (14)
Vocational training (after high school)	1 (2)	6 (12)
Attended college (did not graduate)	3 (6)	12 (24)
College graduate	10 (20)	15 (30)
Graduate school	0 (0)	0 (0)
Race or ethnicity, n (%)		
Hispanic or Latino	34 (68)	34 (68)
White	16 (32)	16 (32)
Asian	0 (0)	0 (0)
American Indian or Alaskan native	0 (0)	0 (0)
Black	0 (0)	0 (0)
Native Hawaiian or other Pacific Islander	0(0)	0 (0)
Health insurance coverage, n (%)		
Medicare	41 (82)	12 (24)
Medicaid	21 (42)	3 (6)
Private insurance	12 (24)	28 (56)
Veterans	1 (2)	1 (2)
None	0 (0)	12 (24)

Table 2. Other characteristics of the patients.

Other characteristics	Statistics
Whom the participant is caring for, n (%)	
Mother	29 (58)
Father	6 (12)
Husband	3 (6)
Mother-in-law	3 (6)
Grandmother	2 (4)
Friend	2 (4)
Brother	1 (2)
Cousin	1 (2)
Wife	1 (2)
Where relative lives, n (%)	
Alone in own home	14 (28)
In household with the participant	9 (18)
With another relative	15 (30)
In a group environment with assistance (eg, an assisted living facility or group home, but not a nursing home)	3 (6)
Nursing home	9 (18)
How long has been doing things for relative that he or she used to do for him or herself (month)	
Mean (SD)	33.16 (26.67)
Median (range)	24.00 (2-96)
Number of other family members or friends (not including participant) provide care routinely	
Mean (SD)	2.74 (2.17)
Median (range)	3.00 (0-7)
A professional home health person (paid or free) helps to care for relative, n (%)	
Yes	16 (32)
How long relative has been diagnosed with dementia or Alzheimer (month)	
Mean (SD)	41.31 (42.16)
Median (range)	24 (1-183)
Activities of daily living	
Mean (SD)	11.76 (5.79)
Median (range)	10.00 (6-24)
Instrumental activities of daily living	
Mean (SD)	26.98 (8.95)
Median (range)	28.00 (11-40)
Relative has made legal arrangements to have a health care power of attorney, n (%)	
Yes	29 (58)
Participant is the power of attorney	18 (36)
Relative has a living will or advance directive, n (%)	
Yes	30 (60)
Relative has shared the living will or advance directive with the participant	21 (42)

Table 3. General decision-making patterns and expectations.

Decision-making patterns and expectations	Statistics, n (%)
Within the family, who makes health care decisions for relative	
Relative alone	2 (4)
Mostly relative	3 (6)
Relative and myself or other family members equally	20 (40)
Mostly myself or other family members	11 (22)
Myself or other family members alone	11 (22)
Who participant thinks should make decisions related to relative's condition	
The health care provider alone	0 (0)
Mostly the health care provider	3 (6)
The health care provider and the family equally	30 (60)
Mostly the family	10 (20)
The family alone	5 (10)

Table 4. Preferences for 7 types of health information and participation in decision making.

Subscale	Information preference, mean (SD)	Decision-making preference, mean (SD)	*t* value (*df*)	*P* value
Diagnosis	68.58 (35.05)	39.32 (25.13)	4.760 (47)	<.001
Treatment	67.01 (38.21)	40.28 (24.75)	4.236 (47)	<.001
Laboratory tests	68.26 (35.60)	30.76 (27.36)	4.924 (47)	<.001
Self-care	67.38 (34.80)	54.97 (23.80)	2.139 (47)	.04
CAM[a]	70.83 (34.55)	47.74 (26.62)	3.109 (47)	.003
Psychosocial	55.56 (36.80)	68.92 (22.26)	−2.255 (47)	.03
Health care providers	60.16 (39.50)	54.69 (25.99)	0.786 (47)	.44

[a]CAM: complementary and alternative medicine.

Caregivers' Mobile Device Usage and Desire for Receiving Information via Mobile Devices

Descriptive results are presented in Table 5. Two-thirds of the caregivers wanted to receive at least *a little* dementia-related information via a smartphone or tablet. Mann-Whitney tests found the amount of dementia-related information caregivers wanted to receive via a mobile device was significantly greater for women than for men (U=84.50, P=.03). Caregivers who owned a tablet were more likely than those who did not own a tablet to want to receive dementia-related information via a mobile device (U=152.00, P=.006). No significant difference was found between caregivers who owned a smartphone and those who did not, or between Hispanics and whites, in how much dementia-related information they wanted to receive via a mobile device. Spearman tests found no significant correlation between how much dementia-related information caregivers wanted to receive via a mobile device and their age, education, cell phone usage duration, or tablet usage duration.

Table 5. Caregivers' mobile device usage and desire for receiving information via mobile devices.

Caregiver mobile device use	Statistics, n (%)
Own a cell phone	
Yes	49 (98)
No	1 (2)
Own a smartphone	
Yes	41 (82)
No	8 (16)
How long have used a smartphone	
Less than 1 year	2 (4)
At least 1 year but less than 3 years	3 (6)
At least 3 years but less than 5 years	6 (12)
At least 5 years but less than 10 years	16 (32)
At least 10 years	16 (32)
Own a tablet (eg, Apple iPad)	
Yes	31 (62)
No	18 (36)
How long have used a tablet	
Less than 1 year	7 (14)
At least 1 year but less than 3 years	3 (6)
At least 3 years but less than 5 years	8 (16)
At least 5 years but less than 10 years	14 (28)
At least 10 years	1 (2)
How much dementia-related information would like to receive via a smartphone or tablet	
None	17 (34)
A little	2 (4)
Some	7 (14)
Most	7 (14)
All	16 (32)

Discussion

Interpreting Our Study Participants' Basic Characteristics

This pilot study was part of a larger project to develop and test a novel, affordable, easy-to-use mHealth intervention to deliver individually tailored information to aid caregivers in rural Hispanic communities to make informed decisions for their relatives suffering from dementia. Our long-term goal is to help caregivers take advantage of new technological advancements to prepare for the wide range of future care needs and transitions for their relatives. This study was a first step taken to understand community characteristics and preferences and determine the feasibility of the larger mHealth intervention. We chose to focus on rural Hispanic communities because Hispanics are the second largest ethnic group and the fastest growing ethnic minority group in the United States and because rural communities face unique challenges [37-47]. Hispanics residing in rural areas are at double jeopardy in getting proper health care and services. Existing interventions for caregivers have largely overlooked the special needs and preferences of rural Hispanic residents [14]. More than two-thirds of the patients and caregivers in our study sample were Hispanic and the others were white. The sample contained no patients or caregivers of other racial or ethnic groups. The sample's race or ethnicity approximately reflects that of the population in the County where 71% of the population is Hispanic, with a few residents belonging to other ethnic minority groups [67].

Census data also show that 15% of people aged above 25 years in this rural Southwest Texas County had college degrees or higher [67]. In our sample, 20% (10/50) of the patients and 30% (15/50) of the caregivers had college degrees. These higher percentages for college education may have been due, at least in part, to the inclusion criterion that participants be able to read and write in English. The majority of the patients (78%; 39/50) were women, with over half of the patients (58%; 29/50) reported as mothers of the caregivers who completed the survey

instruments. The majority of the caregivers (82%; 41/50) were women as well. National data suggest that 60% of caregivers (including but not limited to caregivers of dementia patients) in the United States are female [68], almost two-thirds of people with Alzheimer disease in the United States are women, and approximately two-thirds of Alzheimer disease's caregivers in the United States are also women [21]. It appears that our study sample consisted of even higher percentages of female patients and female caregivers than the national data suggest. These differences in gender composition might be because of our small sample size, such that small differences in numbers could turn into rather large differences in percentages. However, they might also reflect characteristics of the rural Hispanic communities we studied (Hispanic women in rural areas might be even more likely than the general population to be caregivers of people with dementia). Future research should be conducted with larger, more representative samples.

On average, the mean of other family members or friends (not including the research participants themselves) providing care routinely was 2.74 (range 0-7), with a median of 3. Meanwhile, although the majority (76%; 38/50) of the patients lived in a home environment (alone or with a family member), over two-thirds (68%; 34/50) of the patients did not have a professional home health person, whether paid or free, to aid the family in caring for the patient. Together, these findings show a heavy reliance on informal care and an underutilization of formal care for people with dementia in rural Hispanic communities. Our findings suggest that patients and caregivers in these communities face unique challenges in accessing formal health and social services because of, as reported in the literature, their ethnic minority background and residence in rural areas [34,39-44]. Heavy reliance on informal care will likely become an even more serious challenge in the future. As the population ages, increasing numbers of older adults with dementia and/or other conditions will require care so that reliance on informal caregivers will not be sustainable in the long run [69]. Technological developments such as mHealth may be particularly promising for shifting dependency away from informal caregivers while meeting the care needs of the aging population.

All patients had at least some form of health insurance. The vast majority of the patients (82%; 41/50) were on Medicare and a large portion (42%; 21/50) was on Medicaid. The high percentage of patients on Medicaid is not surprising, because 25% of residents in this rural Texas County live in poverty [62]. A majority of caregivers had private health insurance (56%; 28/50), whereas 24% (12/50) of caregivers had no health insurance at all. These findings suggest additional challenges unique to patients and caregivers in rural ethnic minority communities, that is, they tend to be poorer, older, and sicker than their urban counterparts [37,38]. Notably, 42% (21/50) of the patients did not have a power of attorney for their health care and 40% (20/50) did not have a living will or advance directive. These findings also suggest unique challenges that rural Hispanic communities face, and they too are in line with those in the literature; ethnic minority groups, including Hispanics, have low rates of advance care planning [31,32]. These findings illustrate the need for effective interventions in rural Hispanic communities.

The majority (62%; 31/50) of health care decisions in the family were made with some form of shared decision making between the patient and family members; 10% (5/50) of decisions were made by the patient alone or mostly by the patient. However, 22% (11/50) of decisions were made by family members alone without involving the patient. Beyond the family, the majority (60%; 30/50) of caregivers felt that the health care provider and the family should play equal roles in making decisions. Another 30% (15/50) felt that the family mostly or the family alone should make all decisions. No caregiver thought that the health care provider should make decisions alone. In terms of caregivers' preferences for specific types of health information and decision-making participation, our data showed that caregivers were interested in a broad range of health information and decision-making participation, although their levels of interest varied across the 7 subscales and between information and decision-making preferences. Caregivers had a strong desire for all 7 types of information: on a 1 to 100 scale, where 1 indicated the least amount of information wanted and 100 the greatest amount of information wanted, participants scored from 55 to 71 on the 7 types of information wanted. They were also interested in all 7 types of decision-making participation, although their interest in decision-making participation was significantly lower than their interest in information on 5 of the 7 subscales (diagnosis, treatment, laboratory tests, self-care, and CAM). On the psychosocial subscale, however, caregivers' desire for decision-making participation was greater than that for information. On the health care provider subscale, no significant difference was found between caregivers' interests in information and decision-making participation. These findings are similar to those of earlier studies using the HIWQ in different samples [66,70], suggesting generalizability across populations in individual preferences for health information and decision-making participation.

Our data show that all but 1 (98%) of the participants had a mobile phone; however, 16% (8/50) lacked a smartphone. This is consistent with national data: as of January 2018, 95% of the US population had a mobile phone, whereas 17% lacked a smartphone [71]. Although the percentages of people with mobile phones in urban, suburban, and rural areas were approximately the same, rural areas had a greater percentage of people who had nonsmart mobile phone devices (26%) than urban areas (13%) [71]. In addition, those with less than high school education and those who made less than $30,000 a year had higher percentages of nonsmartphone use (33% and 25%, respectively) than the national average [71]. These findings further suggest unique challenges that rural communities often face (eg, poverty and lack of formal education), and they have implications for interventions targeting caregivers in rural areas. Specifically, although smartphones have many advantages over nonsmartphones, mHealth interventions that do not require smartphones (eg, short message service [SMS] text messages supported by all mobile phone devices, smart or nonsmart) may be the best way to reach the most caregivers in rural areas, particularly those who cannot afford smartphones and associated data plans.

Caregivers' Mobile Health Preferences

In terms of how much dementia-related information caregivers would like to receive via a mobile device, two-thirds of the caregivers went for an extreme: 34% (17/50) preferred to receive *no* information, whereas another 32% (16/50) preferred to receive *all* information via a mobile device. Caregivers' age, education, race or ethnicity, smartphone ownership, cell phone usage duration, or tablet usage duration did not seem to have a relationship with how much dementia-related information caregivers wanted to receive via a mobile device. Only 2 predicative variables were found. Women wanted to receive more dementia-related information via a mobile device than men. This is not surprising; research has consistently shown that women are more interested than men in obtaining health-related information [72-74]. Moreover, caregivers who owned a tablet wanted to receive more dementia-related information via a mobile device than those who did not own a tablet. This finding is particularly interesting given that no relationship was found between the other mobile device-related variables (smartphone ownership, cell phone usage duration, and tablet usage duration) and the amount of information caregivers wanted to receive via a mobile device.

Limitations and Future Directions

Due to limited resources, we were able to use only an English instrument; thus, our sample included only caregivers fluent in English and bilingual in English and Spanish. As such, the findings of this study might not be generalized to caregivers not fluent in English. The small sample might also limit the findings' generalizability. Future research should be conducted with larger and more representative samples. This pilot study focused specifically on mHealth and did not include interventions that were internet-based and that relied largely on computers (for a systematic review, see the study by Hopwood et al [75]). We chose to focus on mHealth mainly because those who live in rural Hispanic communities are more likely to have less formal education and higher levels of poverty and they might be more likely to use mobile devices than computers on a daily basis. However, it might be interesting to explore in future research whether or how mobile device and internet-based interventions might be perceived and used differently and/or similarly by rural caregivers.

Conclusions

This pilot study generated preliminary data about key characteristics of people with dementia and their family caregivers in rural Hispanic communities, including caregivers' preferences for different types and amounts of health information and decision-making participation and the needs of rural caregivers for mHealth-based interventions tailored to their unique circumstances. In particular, our data show that 42% (21/50) of the patients did not have a power of attorney for their health care and 40% (20/50) did not have a living will or advance directive. These findings illustrate the need for effective interventions to improve the rates of having a power of attorney and a living will or advance directive in rural Hispanic communities. Compared with the national data, our study found an even higher percentage of female caregivers, perhaps because Hispanic women in rural areas are even more likely than the general population to be caring for their families. Caregivers, women or men, were interested in a broad range of health information and decision-making participation; women, compared with their male counterparts, wanted to have even more dementia-related information via a mobile device. Together, these findings support a need for mHealth interventions that can provide relevant information for caregivers in rural Hispanic communities.

However, in developing mHealth interventions for these caregivers, it is important to bear in mind that although almost all caregivers in our study sample had a mobile phone, 16% lacked a smartphone. This is consistent with the findings for national samples. Thus, mHealth interventions that do not require smartphones (eg, SMS text messages supported by all mobile phone devices, smart or nonsmart) may be the best way to reach the most caregivers in rural areas, particularly those who cannot afford smartphones and associated data plans. Furthermore, caregivers' levels of interest in dementia-related information and decision-making participation varied across the 7 subscales. Thus, mHealth interventions, smartphone-based or not, should strive to provide information tailored to individual caregivers' specific preferences (eg, providing more self-care related information to caregivers who want more of such information, whereas providing more CAM-related information to those who want more CAM-related information).

The findings of this pilot study have implications for dementia research, practice, and policy making. Our study of the characteristics of people with dementia and their family caregivers in rural areas, especially those in racial or ethnical minority groups, supports a patient- and family-centered approach to address the significant need for interventions sensitive to underserved populations' unique situations. Affordable and easy-to-use mHealth interventions can help caregivers obtain desired health information to make informed decisions, even if they have limited technology experience and/or cannot afford the cost of smartphones and services. Such interventions should have a long-term and broad impact on EOL care for people with dementia and their caregivers in the rapidly evolving mHealth era.

Acknowledgments

This study was part of a pilot research grant from the population health initiative at the University of Texas at Austin, entitled *Values and Preferences in Dementia Family Caregivers' End-of-Life Decision-Making*; principal investigator (PI) was BX and co-PIs were KF and JK. The authors would like to thank the caregivers who participated in the study. The authors would like to thank the biostatistician Dr Adama Brown and graduate research assistant (GRA) Ya-Ching Huang in the Cain Center of the School of Nursing and GRA Kolina Koltai in the School of Information at the University of Texas at Austin for their assistance

in the data entry. Editorial support with manuscript development was provided by the Cain Center for Nursing Research and the Center for Transdisciplinary Collaborative Research in Self-Management Science (P30, NR015335) at the University of Texas at Austin School of Nursing.

Conflicts of Interest

None declared.

References

1. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
2. Anastasiadou D, Folkvord F, Lupiañez-Villanueva F. A systematic review of mHealth interventions for the support of eating disorders. Eur Eat Disord Rev 2018 Sep;26(5):394-416 [FREE Full text] [doi: 10.1002/erv.2609] [Medline: 29927004]
3. Giunti G, Guisado Fernández E, Dorronzoro Zubiete E, Rivera Romero O. Supply and demand in mHealth apps for persons with multiple sclerosis: systematic search in app Stores and scoping literature review. JMIR Mhealth Uhealth 2018 May 23;6(5):e10512 [FREE Full text] [doi: 10.2196/10512] [Medline: 29792295]
4. Feldman DI, Theodore Robison W, Pacor JM, Caddell LC, Feldman EB, Deitz RL, et al. Harnessing mHealth technologies to increase physical activity and prevent cardiovascular disease. Clin Cardiol 2018 Jul;41(7):985-991 [FREE Full text] [doi: 10.1002/clc.22968] [Medline: 29671879]
5. Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth interventions to support self-management in HIV: a systematic review. Open AIDS J 2017;11:119-132 [FREE Full text] [doi: 10.2174/1874613601711010119] [Medline: 29290888]
6. Biagianti B, Hidalgo-Mazzei D, Meyer N. Developing digital interventions for people living with serious mental illness: perspectives from three mHealth studies. Evid Based Ment Health 2017 Nov;20(4):98-101 [FREE Full text] [doi: 10.1136/eb-2017-102765] [Medline: 29025862]
7. Changizi M, Kaveh MH. Effectiveness of the mHealth technology in improvement of healthy behaviors in an elderly population-a systematic review. Mhealth 2017;3:51 [FREE Full text] [doi: 10.21037/mhealth.2017.08.06] [Medline: 29430455]
8. Chen H, Chai Y, Dong L, Niu W, Zhang P. ffectiveness and appropriateness of mHealth interventions for maternal and child health: systematic review. JMIR Mhealth Uhealth 2018 Jan 9;6(1):e7 [FREE Full text] [doi: 10.2196/mhealth.8998] [Medline: 29317380]
9. Slater H, Campbell JM, Stinson JN, Burley MM, Briggs AM. End user and implementer experiences of mHealth technologies for noncommunicable chronic disease management in young adults: systematic review. J Med Internet Res 2017 Dec 12;19(12):e406 [FREE Full text] [doi: 10.2196/jmir.8888] [Medline: 29233804]
10. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
11. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. JMIR Mhealth Uhealth 2017 Sep 26;5(9):e141 [FREE Full text] [doi: 10.2196/mhealth.7136] [Medline: 28951378]
12. Calderón TA, Martin H, Volpicelli K, Frasso R, Díaz Arroyo EC, Gozzer E, et al. Understanding potential uptake of a proposed mHealth program to support caregiver home management of childhood illness in a resource-poor setting: a qualitative evaluation. Mhealth 2017;3:19 [FREE Full text] [doi: 10.21037/mhealth.2017.04.03] [Medline: 28607905]
13. Kazi AM, Carmichael JL, Hapanna GW, Wangoo PG, Karanja S, Wanyama D, et al. Assessing mobile phone access and perceptions for texting-based mHealth interventions among expectant mothers and child caregivers in remote regions of Northern Kenya: a survey-based descriptive study. JMIR Public Health Surveill 2017 Jan 30;3(1):e5 [FREE Full text] [doi: 10.2196/publichealth.5386] [Medline: 28137702]
14. Xie B, Berkley AS, Kwak J, Fleischmann KR, Champion JD, Koltai KS. End-of-life decision making by family caregivers of persons with advanced dementia: a literature review of decision aids. SAGE Open Med 2018;6:2050312118777517 [FREE Full text] [doi: 10.1177/2050312118777517] [Medline: 29844911]
15. Davis BH, Nies MA, Shehab M, Shenk D. Developing a pilot e-mobile app for dementia caregiver support: lessons learned. Online J Nurs Informatics 2014 Feb 2;18:A [FREE Full text]
16. Davis BH, Shehab M, Shenk D, Nies MA. E-mobile pilot for community-based dementia caregivers identifies desire for security. Gerontechnology 2014;13(3):332-336 [FREE Full text]
17. Griffiths PC, Kovaleva M, Higgins M, Langston AH, Hepburn K. Tele-Savvy: an online program for dementia caregivers. Am J Alzheimers Dis Other Demen 2018 Jan 1:1533317518755331. [doi: 10.1177/1533317518755331] [Medline: 29544342]
18. Duggleby W, Ploeg J, McAiney C, Peacock S, Fisher K, Ghosh S, et al. Web-based intervention for family carers of persons with dementia and multiple chronic conditions (My Tools 4 Care): pragmatic randomized controlled trial. J Med Internet Res 2018 Jun 29;20(6):e10484 [FREE Full text] [doi: 10.2196/10484] [Medline: 29959111]

19. Barbabella F, Poli A, Andréasson F, Salzmann B, Papa R, Hanson E, et al. A web-based psychosocial intervention for family caregivers of older people: results from a mixed-methods study in three European countries. JMIR Res Protoc 2016 Oct 6;5(4):e196 [FREE Full text] [doi: 10.2196/resprot.5847] [Medline: 27713113]
20. Alzheimer's Disease International. 2015. Dementia statistics URL: https://www.alz.co.uk/research/statistics [accessed 2018-10-30] [WebCite Cache ID 73YcdjPFn]
21. Alzheimer's Association. 2018 Alzheimer's disease facts and figures. Alzheimers Dement 2018 Mar;14(3):367-429 [FREE Full text]
22. Collopy BJ. The moral underpinning of the proxy-provider relationship: issues of trust and distrust. J Law Med Ethics 1999;27(1):37-45 [FREE Full text] [Medline: 11657141]
23. Ditto PH, Danks JH, Smucker WD, Bookwala J, Coppola KM, Dresser R, et al. Advance directives as acts of communication: a randomized controlled trial. Arch Intern Med 2001 Feb 12;161(3):421-430 [FREE Full text] [Medline: 11176768]
24. Gessert CE, Forbes S, Bern-Klug M. Planning end-of-life care for patients with dementia: roles of families and health professionals. Omega (Westport) 2000;42(4):273-291. [Medline: 12569923]
25. Swigart V, Lidz C, Butterworth V, Arnold R. Letting go: family willingness to forgo life support. Heart Lung 1996;25(6):483-494 [FREE Full text] [Medline: 8950128]
26. Hawkins NA, Ditto PH, Danks JH, Smucker WD. Micromanaging death: process preferences, values, and goals in end-of-life medical decision making. Gerontologist 2005 Feb;45(1):107-117 [FREE Full text] [Medline: 15695421]
27. Lorenz K, Lynn J, Morton SC, Dy S, Mularski R, Shugarman L, et al. End-of-life care and outcomes. Evid Rep Technol Assess (Summ) 2004 Dec(110):1-6. [Medline: 15649108]
28. American Academy of Hospice Palliative Medicine, Center to Advance Palliative Care, HospicePalliative Nurses Association, Last Acts Partnership, National Hospice Palliative Care Organization. National consensus project for quality palliative care: clinical practice guidelines for quality palliative care, executive summary. J Palliat Med 2004 Oct;7(5):611-627. [doi: 10.1089/jpm.2004.7.611] [Medline: 15588352]
29. US Department of Health annd Human Services. NIH State Sci Statements. National Institutes of Health 2004;21(3):1-26 [FREE Full text] [Medline: 17308546]
30. Alzheimer's A. 2016 Alzheimer's disease facts and figures. Alzheimers Dement 2016;12(4):459-509 [FREE Full text]
31. Connolly A, Sampson EL, Purandare N. End-of-life care for people with dementia from ethnic minority groups: a systematic review. J Am Geriatr Soc 2012 Feb;60(2):351-360 [FREE Full text] [doi: 10.1111/j.1532-5415.2011.03754.x] [Medline: 22332675]
32. Kwak J, Haley WE. Current research findings on end-of-life decision making among racially or ethnically diverse groups. Gerontologist 2005 Oct;45(5):634-641 [FREE Full text] [Medline: 16199398]
33. Lines LM, Sherif NA, Wiener JM. Semantic Scholar. Research Triangle Park, NC: RTI Press; 2014 Dec. Racial and Ethnic Disparities among Individuals with Alzheimer's Disease in the United States: A Literature Review URL: https://pdfs.semanticscholar.org/7ed5/b9ed14cd3e002df546134e76766d01c5c4aa.pdf [accessed 2018-10-30] [WebCite Cache ID 73Yd5eA14]
34. Cooper C, Tandy AR, Balamurali TB, Livingston G. A systematic review and meta-analysis of ethnic differences in use of dementia treatment, care, and research. Am J Geriatr Psychiatry 2010 Mar;18(3):193-203 [FREE Full text] [doi: 10.1097/JGP.0b013e3181bf9caf] [Medline: 20224516]
35. Kusmin LD. US Department of Agriculture, Economic Research Service, Economic Information Bulletin 145. Rural America at a Glance, 2015 Edition URL: https://www.ers.usda.gov/webdocs/publications/44015/55581_eib145.pdf?v=42397 [accessed 2018-10-30] [WebCite Cache ID 73Yd7oSxL]
36. Moy E, Garcia MC, Bastian B, Rossen LM, Ingram DD, Faul M, et al. Leading causes of death in nonmetropolitan and metropolitan areas-United States, 1999-2014. MMWR Surveill Summ 2017 Jan 13;66(1):1-8 [FREE Full text] [doi: 10.15585/mmwr.ss6601a1] [Medline: 28081058]
37. Buckwalter KC, Davis LL. In: Talley RC, Chwalisz K, Buckwalter KC, editors. Elder caregiving in rural communities. Rural Caregiving in the United States: Research, Practice, Policy. New York, NYpringer; 2011:33-46.
38. Advisory Committee on Rural Health and Human Services. 2015 Oct. Mortality and Life Expectancy in Rural America: Connecting the Health and Human Service Safety Nets to Improve Health Outcomes over the Life Course: Policy Brief October 2015 URL: https://www.hrsa.gov/advisorycommittees/rural/publications/mortality.pdf [accessed 2018-10-30] [WebCite Cache ID 73YdB8Ukz]
39. Monahan DJ. Family caregivers for seniors in rural areas. J Fam Soc Work 2013;16(1):116-128 [FREE Full text] [doi: 10.1080/1052158.2012.747461]
40. Blusi M, Asplund K, Jong M. Older family carers in rural areas: experiences from using caregiver support services based on information and communication technology (ICT). Eur J Ageing 2013 Sep;10(3):191-199 [FREE Full text] [doi: 10.1007/s10433-013-0260-1] [Medline: 28804294]
41. Buckwalter K. Interventions for family caregivers of patients with Alzheimer's disease in community-based settings: items for consideration. Int Psychogeriatr 1996;8 Suppl 1:121-122 [FREE Full text] [Medline: 8934281]

42. Artnak KE, McGraw RM, Stanley VF. Health care accessibility for chronic illness management and end-of-life care: a view from rural America. J Law Med Ethics 2011;39(2):140-155 [FREE Full text] [doi: 10.1111/j.1748-720X.2011.00584.x] [Medline: 21561510]
43. Ashcraft AS, Owen DC. End-of-life planning in a rural elderly cohort. Geriatr Nurs 2016;37(1):71-74 [FREE Full text] [doi: 10.1016/j.gerinurse.2015.11.009] [Medline: 26708142]
44. Downing J, Jack BA. End-of-life care in rural areas: what is different? Curr Opin Support Palliat Care 2012 Sep;6(3):391-397 [FREE Full text] [doi: 10.1097/SPC.0b013e328356ab1f] [Medline: 22801467]
45. Ehrlich K, Boström AM, Mazaheri M, Heikkilä K, Emami A. Family caregivers' assessments of caring for a relative with dementia: a comparison of urban and rural areas. Int J Older People Nurs 2015 Mar;10(1):27-37 [FREE Full text] [doi: 10.1111/opn.12044] [Medline: 24433340]
46. Morgan DG, Semchuk KM, Stewart NJ, D'Arcy C. Rural families caring for a relative with dementia: barriers to use of formal services. Soc Sci Med 2002 Oct;55(7):1129-1142. [Medline: 12365526]
47. Sun F, Kosberg JI, Kaufman AV, Leeper JD. Coping strategies and caregiving outcomes among rural dementia caregivers. J Gerontol Soc Work 2010 Aug;53(6):547-567 [FREE Full text] [doi: 10.1080/01634372.2010.496823] [Medline: 20658420]
48. McNutt RA. Shared medical decision making: problems, process, progress. JAMA 2004 Nov 24;292(20):2516-2518 [FREE Full text] [doi: 10.1001/jama.292.20.2516] [Medline: 15562133]
49. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med 1999 Sep;49(5):651-661. [Medline: 10452420]
50. Epstein RM, Alper BS, Quill TE. Communicating evidence for participatory decision making. JAMA 2004 May 19;291(19):2359-2366. [doi: 10.1001/jama.291.19.2359] [Medline: 15150208]
51. Hack TF, Degner LF, Parker PA. The communication goals and needs of cancer patients: a review. Psychooncology 2005 Oct;14(10):831-846. [doi: 10.1002/pon.949] [Medline: 16200519]
52. Au A, Lam W, Tsang J, Yau T, Soong I, Yeo W, et al. Supportive care needs in Hong Kong Chinese women confronting advanced breast cancer. Psychooncology 2013 May;22(5):1144-1151. [doi: 10.1002/pon.3119] [Medline: 22715115]
53. So WK, Chan CW, Choi KC, Wan RW, Mak SS, Chair SY. Perceived unmet needs and health-related quality of life of Chinese cancer survivors at 1 year after treatment. Cancer Nurs 2013;36(3):E23-E32 [FREE Full text] [doi: 10.1097/NCC.0b013e318263f28e] [Medline: 22964864]
54. Qiu WL, Lin PJ, Ruan FQ, Wu X, Lin MP, Liu L, et al. Requests for health education from Chinese cancer patients during their recovery period: a cross-sectional study. J Cancer Educ 2013 Sep;28(3):428-434 [FREE Full text] [doi: 10.1007/s13187-013-0482-0] [Medline: 23728991]
55. 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] [Medline: 23235644]
56. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]
57. Nhavoto JA, Grönlund A. Mobile technologies and geographic information systems to improve health care systems: a literature review. JMIR Mhealth Uhealth 2014;2(2):e21 [FREE Full text] [doi: 10.2196/mhealth.3216] [Medline: 25099368]
58. Househ M. The role of short messaging service in supporting the delivery of healthcare: an umbrella systematic review. Health Informatics J 2014 Jul 18;22(2):140-150. [doi: 10.1177/1460458214540908] [Medline: 25038203]
59. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
60. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
61. 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]
62. Kim H, Xie B. Health literacy in the eHealth era: a systematic review of the literature. Patient Educ Couns 2017 Jun;100(6):1073-1082. [doi: 10.1016/j.pec.2017.01.015] [Medline: 28174067]
63. Xie B, Wang M, Feldman R. Preferences for health information and decision-making: development of the Health Information Wants Questionnaire. In: Proceedings of the 2011 iConference. New York: ACM; 2011 Presented at: iConference; February 8-11, 2011; Seattle, WA p. 273-280 URL: https://doi.org/10.1145/1940761.1940799
64. Xie B, Wang M, Feldman R, Zhou L. Health information and decision-making preferences in the internet age: a pilot study using the Health Information Wants (HIW) Questionnaire. In: Proceedings of the 1st ACM International Health Informatics Symposium. New York: ACM; 2010 Presented at: The 1st ACM International Health Informatics Symposium; November 11-12, 2010; Arlington, VA p. 610-619 URL: https://doi.org/10.1145/1882992.1883090
65. Xie B, Wang M, Feldman R, Zhou L. Internet use frequency and patient-centered care: measuring patient preferences for participation using the health information wants questionnaire. J Med Internet Res 2013;15(7):e132 [FREE Full text] [doi: 10.2196/jmir.2615] [Medline: 23816979]

66. Xie B, Wang M, Feldman R, Zhou L. Exploring older and younger adults' preferences for health information and participation in decision making using the Health Information Wants Questionnaire (HIWQ). Health Expect 2014 Dec;17(6):795-808 [FREE Full text] [doi: 10.1111/j.1369-7625.2012.00804.x] [Medline: 22783800]
67. United States Census Bureau. 2017. Quick Facts URL: https://www.census.gov/quickfacts [accessed 2018-10-30] [WebCite Cache ID 73YdWtSWP]
68. National Alliance for Caregiving the AARP Public Policy Institute. Caregiving. 2015. 2015 Report Caregiving in the U.S URL: http://www.caregiving.org/wp-content/uploads/2015/05/2015_CaregivingintheUS_Final-Report-June-4_WEB.pdf [accessed 2018-10-30] [WebCite Cache ID 73YdiulLv]
69. Kingston A, Comas-Herrera A, Jagger C, MODEM project. Forecasting the care needs of the older population in England over the next 20 years: estimates from the Population Ageing and Care Simulation (PACSim) modelling study. Lancet Public Health 2018 Sep;3(9):e447-ee55 [FREE Full text] [doi: 10.1016/S2468-2667(18)30118-X] [Medline: 30174210]
70. Xie B, Su Z, Liu Y, Wang M, Zhang M. Health information wanted and obtained from doctors/nurses: a comparison of Chinese cancer patients and family caregivers. Support Care Cancer 2015 Oct;23(10):2873-2880 [FREE Full text] [doi: 10.1007/s00520-015-2651-7] [Medline: 25855039]
71. Pew Research Center. Pew Internet. 2018. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-10-30] [WebCite Cache ID 73YdkJESJ]
72. Bidmon S, Terlutter R. ender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. J Med Internet Res 2015;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]
73. Fox S. Pew Internet. Profiles of Health Information Seekers URL: http://www.pewinternet.org/2011/02/01/profiles-of-health-information-seekers/ [accessed 2018-10-30] [WebCite Cache ID 73Ydm9LNT]
74. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
75. Hopwood J, Walker N, McDonagh L, Rait G, Walters K, Iliffe S, et al. internet-based interventins aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216 [FREE Full text] [doi: 10.2196/jmir.9548] [Medline: 29895512]

Abbreviations

ADL: activities of daily living
ADRD: Alzheimer disease and related dementias
CAM: complementary and alternative medicine
DPS: decision-making preference scale
EOL: end-of-life
GRA: graduate research assistant
HIWQ: Health Information Wants Questionnaire
IPS: Information Preference Scale
SMS: short message service
PI: principal investigator
mHealth: mobile health
RA: research assistant

Edited by G Eysenbach; submitted 24.07.18; peer-reviewed by B Davis, N Davies; comments to author 05.09.18; revised version received 11.09.18; accepted 12.09.18; published 10.12.18.

Please cite as:
Xie B, Champion JD, Kwak J, Fleischmann KR
Mobile Health, Information Preferences, and Surrogate Decision-Making Preferences of Family Caregivers of People With Dementia in Rural Hispanic Communities: Cross-Sectional Questionnaire Study
J Med Internet Res 2018;20(12):e11682
URL: https://www.jmir.org/2018/12/e11682/
doi:10.2196/11682
PMID:30530450

Original Paper

Continuous Versus Intermittent Vital Signs Monitoring Using a Wearable, Wireless Patch in Patients Admitted to Surgical Wards: Pilot Cluster Randomized Controlled Trial

Candice Downey[1], MBChB, BSc, PGDip, MRCS; Rebecca Randell[2], BSc, PhD; Julia Brown[3], BSc, MSc; David G Jayne[1], MBBCh, BSc, MD, FRCS

[1]Leeds Institute of Biomedical & Clinical Sciences, University of Leeds, Leeds, United Kingdom

[2]School of Healthcare, University of Leeds, Leeds, United Kingdom

[3]Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom

Corresponding Author:
Candice Downey, MBChB, BSc, PGDip, MRCS
Leeds Institute of Biomedical & Clinical Sciences
University of Leeds
Clinical Sciences Building
St James's University Hospital
Leeds, LS9 7TF
United Kingdom
Phone: 44 7544783204
Fax: 44 113 2065281
Email: c.l.downey@leeds.ac.uk

Related Article:

Comment in: https://www.jmir.org/2021/3/e14042

Abstract

Background: Vital signs monitoring is a universal tool for the detection of postoperative complications; however, unwell patients can be missed between traditional observation rounds. New remote monitoring technologies promise to convey the benefits of continuous monitoring to patients in general wards.

Objective: The aim of this pilot study was to evaluate whether continuous remote vital signs monitoring is a practical and acceptable way of monitoring surgical patients and to optimize the delivery of a definitive trial.

Methods: We performed a prospective, cluster-randomized, parallel-group, unblinded, controlled pilot study. Patients admitted to 2 surgical wards at a large tertiary hospital received either continuous and intermittent vital signs monitoring or intermittent monitoring alone using an early warning score system. Continuous monitoring was provided by a wireless patch, worn on the patient's chest, with data transmitted wirelessly every 2 minutes to a central monitoring station or a mobile device carried by the patient's nurse. The primary outcome measure was time to administration of antibiotics in sepsis. The secondary outcome measures included the length of hospital stay, 30-day readmission rate, mortality, and patient acceptability.

Results: Overall, 226 patients were randomized between January and June 2017. Of 226 patients, 140 were randomized to continuous remote monitoring and 86 to intermittent monitoring alone. On average, patients receiving continuous monitoring were administered antibiotics faster after evidence of sepsis (626 minutes, n=22, 95% CI 431.7-820.3 minutes vs 1012.8 minutes, n=12, 95% CI 425.0-1600.6 minutes), had a shorter average length of hospital stay (13.3 days, 95% CI 11.3-15.3 days vs 14.6 days, 95% CI 11.5-17.7 days), and were less likely to require readmission within 30 days of discharge (11.4%, 95% CI 6.16-16.7 vs 20.9%, 95% CI 12.3-29.5). Wide CIs suggest these differences are not statistically significant. Patients found the monitoring device to be acceptable in terms of comfort and perceived an enhanced sense of safety, despite 24% discontinuing the intervention early.

Conclusions: Remote continuous vital signs monitoring on surgical wards is practical and acceptable to patients. Large, well-controlled studies in high-risk populations are required to determine whether the observed trends translate into a significant benefit for continuous over intermittent monitoring.

Trial Registration: International Standard Randomised Controlled Trial Number ISRCTN60999823; http://www.isrctn.com/ISRCTN60999823 (Archived by WebCite at http://www.webcitation.org/73ikP6OQz)

(J Med Internet Res 2018;20(12):e10802) doi:10.2196/10802

KEYWORDS

general surgery; monitoring; physiological; randomized controlled trial; vital signs

Introduction

Perioperative complications are unfortunately common following surgical procedures. Postoperative mortality is the third leading cause of death in the United States [1]. The International Surgical Outcomes Study found that 17% of patients undergoing inpatient surgery developed at least one complication [2]. This figure rose to 27% in patients undergoing major surgery. In addition, 2.8% of patients who developed a postoperative complication died before discharge from hospital. Monitoring patients beyond the operating room is important to allow early detection of clinical deterioration and timely intervention [3].

The early warning score system is predicated on the idea that derangements in simple physiological observations can identify hospital inpatients at high risk of deterioration [4]. Prodromal warning signs, such as increased respiratory rate or decreased blood pressure, precede critical illness [5], allowing early recognition and management of patients to reverse the abnormal physiological decline or prompt admission to a critical care area.

A critical limitation of early warning score systems is their intermittent nature [6]. Clinical deterioration on general wards may remain undetected for hours before clinicians are alerted [3]. One solution may be continuous vital signs monitoring, which until now has been limited to use on critical care wards owing to prohibitive cost and implications for patient mobility and recovery.

The development of wireless and wearable sensors allows continuous monitoring of ambulatory patients. A number of such tools have already received the Food and Drug Administration clearance, but clinical studies are required to demonstrate their clinical utility in the postsurgical setting [3,7].

A recent systematic review identified 9 studies assessing the effect of continuous vital signs monitoring on general wards [8]. The authors found no evidence of a marked reduction in intensive care unit transfers or other adverse events with continuous monitoring but recognized heterogeneous methods, study populations, and outcome measures. Efficient, well-designed pilot studies are vital to ensure the robust design and implementation of large-scale clinical trials [9].

This study aims to evaluate whether continuous remote vital signs monitoring is a practical way of monitoring surgical patients outside of the critical care setting and whether its use is acceptable to patients. The pilot data will be used to inform a further definitive trial to optimize recruitment, treatment compliance, and follow-up protocols.

Methods

Study Design

The study was designed as a pilot cluster-randomized, prospective, parallel-group, controlled single-center pilot study, comparing remote continuous vital signs monitoring and intermittent monitoring with intermittent monitoring alone.

Ethical approval was granted on November 30, 2016, by the Yorkshire & The Humber—Bradford Leeds Research Ethics Committee (ref: 16/YH/0426). The study was prospectively registered on the International Standard Randomised Controlled Trial Number registry (ISRCTN60999823). No changes were made to the registered protocol. The trial was performed in accordance with the Good Clinical Practice guidelines and the Declaration of Helsinki and is presented according to the Consolidated Standards of Reporting Trials (CONSORT) statement principles [10] and the CONSORT-EHEALTH checklist (Multimedia Appendix 1) [11].

The study population comprised patients admitted to 2 elective general surgery wards (male and female) at a single tertiary center in Leeds, United Kingdom. Patients aged ≥18 years who were able to provide informed consent to participate were included. Patients with a known allergy to the electrode adhesive and those with a cardiac pacemaker *in situ* were excluded. Patients were approached face-to-face by a research nurse or clinical fellow as soon as possible after their admission onto the wards. After consideration of a patient information sheet (see Multimedia Appendix 2), participants gave informed consent to enter the study.

Randomization

Consenting participants were allocated to one of the two monitoring arms for the length of their admission, according to the ward bay they were first arbitrarily admitted to. Each ward has 4 bays, containing 6 beds each.

Of the 4 bays on each ward, 3 were randomly allocated to one of the monitoring arms; 2 bays were allocated to receive the patch and one to receive usual intermittent monitoring. Each bay was independently block randomized to an intervention arm by the primary investigator (CD) using Web-based software: Sealed Envelope [12].

The 2 remaining bays (one on each ward) could not be randomized because they did not have the required hardware

installed. Patients in these bays were therefore allocated to receive usual intermittent monitoring alone.

The allocation of patients to each bay was performed by hospital bed managers, who were independent of the trial and unaware of the bay allocations. Owing to the nature of the intervention, neither patients nor their nurses were blinded to the allocated monitoring arm.

Control

All patients in the study received usual intermittent vital signs monitoring. In our institution, this is the National Early Warning Score, which involves intermittent manual charting of vital signs and the calculation of a combined score, indicating the patient's status. The control arm received intermittent monitoring alone. For postoperative patients, this typically consisted of an hourly recording of blood pressure, pulse, temperature, respiratory rate, and oxygen saturation until the patient's condition was stable when the frequency of observations was decreased to 2 hourly and, then, 4 hourly. For patients not undergoing an operation, the frequency of monitoring was tailored to their condition.

Intervention

Patients admitted to an intervention bay received usual intermittent monitoring, in addition to continuous vital signs monitoring through the SensiumVitals (Sensium, Abingdon, United Kingdom) system. This system consists of a Conformité Européene–marked wireless patch (Figure 1) worn on the chest of a patient, which continuously monitors heart rate, respiratory rate, and temperature. The data are transmitted wirelessly every 2 minutes to a central monitoring station or a mobile device carried by the patient's nurse. The nurse is alerted when there is any deviation from preset physiological norms. The alert prompts an acknowledgment of the notification, after which nurses are free to act according to their clinical discretion. Reminders were sent every 14 minutes until acknowledgment, and levels of engagement were monitored through daily ward visits. All other clinical care remained as normal in the intervention group.

The monitoring system was set up in the wards over a period of 6 weeks, during which a number of stakeholders were engaged with the project. Early on, permission from the Estates and Information Technology departments was obtained. The ceiling-mounted bridges were installed by the hospital Estates department using existing electrical wiring circuits to ensure compliance with local policies. The monitoring software was integrated with the hospital admissions data system so that patients could easily be added to the remote monitoring system. All data were stored and retained on the hospital network, alleviating initial concerns about data security by inheriting all hospital security procedures and data backup policies.

General surgeons were informed of the project at local audit meetings so that they would understand potential escalations of care prompted by the device, although they were not expected to apply the patches or carry the mobile devices. Nursing staff were trained face-to-face to use the system over a period of 1 week, after which ad-hoc refresher training was available on request.

Primary Outcome Measure

The primary outcome measure was time to antibiotics after the first evidence of sepsis, defined according to a revised consensus conference definition in 2001 by the presence of a likely source of infection and ≥2 of the following criteria [13]: (1) temperature >38.3°C or <36.0°C; (2) tachycardia >90 beats/minute; (3) tachypnea >20 breaths/minute; (4) partial pressure of carbon dioxide <4.3 kPa; (5) hyperglycemia (blood glucose >6.6 mmol/L) in the absence of diabetes mellitus; (6) acutely altered mental status; and (7) white blood cells count>12×109/L or <4×109/L.

The decision to prescribe antibiotics was usually made by a junior doctor on the ward, based on local protocols and clinical discretion. The time to antibiotics was determined by review of the observations chart, SensiumVitals data, electronic medications record, and medical notes of patients during their hospital admission.

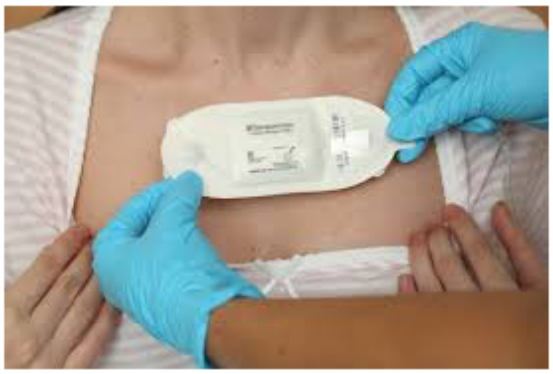
Secondary Outcome Measures

The secondary outcome measures were in-hospital mortality, length of hospital stay, number of admissions to Level II or III care, and readmission to hospital within 30 days of discharge.

Patient Acceptability and Compliance

Patient compliance was determined by the number of patients not wearing a patch for at least 5 days. To assess the acceptability, patients in the continuous monitoring group were asked to complete a short 2-question questionnaire at the bedside on the day of discharge from hospital. Patients were asked to rank the comfort and sense of safety they perceived from wearing the patch on a scale from "Strongly Agree" to "Strongly Disagree."

Figure 1. The SensiumVitals patch. Photograph used with permission from Sensium (Abingdon, United Kingdom).



Data Collection

The data collection was performed daily by a research nurse and a clinical fellow. This allowed any harms or benefits to be collected in real time. The data were taken from the clinical records made by patients' usual care teams, including a succession of junior medical staff on rotation, who were unaware of the study. The objective methods of collecting the outcome data minimized the risk of bias. In addition, the predefined criteria for the outcome measures provided minimal scope for interpretation of their presence or absence by the data collection team.

Statistical Analysis

A formal sample size calculation was not possible given the lack of data surrounding the primary outcome measure; thus, assumptions were used to calculate an appropriate sample size. A sample size of 325-625 was suggested as an appropriate target based on the assumed eligibility rate (90%), consent rate (30%-50%), and patient turnover (4 patients per bed per calendar month).

The analysis was on an intention-to-treat basis at the individual patient level. Each of the outcome measures was summarized by the intervention or control group using descriptive statistics. As there was no formal sample size calculation, no statistical comparison between trial arms was made.

Exploratory Analysis

The primary analysis included only the 6 randomized bays. The 2 nonrandomized bays were included in a separate exploratory analysis.

Progression Criteria

Although no formal progression criteria were defined in the protocol, considerations for the progression to a definitive trial included:

- Recruitment rate (at least 325 patients within the 9-month recruitment period)
- Protocol adherence (proportion of patients wearing the patch for at least 5 days)
- Suitability of primary outcome measure to inform the sample size of a definitive trial.

Results

Principal Results

In this study, 226 patients were randomized between January and June 2017. Figure 2 presents the patient flowchart, and Table 1 presents patients' characteristics.

While 140 patients were allocated to receive continuous monitoring alongside standard care, 86 patients were randomized to the control group. A further 124 patients from nonrandomized bays were included in the exploratory analysis.

Two patients in the control arm (both from randomized bays) were given the continuous monitoring intervention at the request of the direct care team.

Overall, 73% (257/350) of patients underwent a surgical intervention during their admission; these were mostly colorectal resections (n=132), stoma formations (n=23), stoma reversals (n=12), hernia repairs (n=20), and other colorectal laparotomies, including fistula exploration (n=23). Less common procedures were hepatobiliary (n=14), urological (n=9), appendicectomy (n=7), and abdominal wall repair (n=5). Of note, 8 procedures were classified as Other.

A similar number of complications and sepsis events occurred across both arms of the study (see Multimedia Appendix 3), indicating that both groups had similar baseline risk factors.

One patient died of alcoholic liver disease during their participation in the study. This patient was allocated to receive continuous monitoring.

Primary Outcome Measure

Table 2 summarizes the main results of the study. In the intervention arm, 17.1% (24/140) of patients experienced a sepsis event; this figure was 14% (12/86 patients) in the control arm. Of 36 sepsis events recorded in randomized bays, there were sufficient data to analyze the time to antibiotics in 34 cases. The average time from the first evidence of sepsis to the first administration of antibiotics was 626 minutes in the intervention group (n=22, 95% CI 431.7-820.3 minutes). The average time to antibiotics in the control group was 1012.8 minutes (n=12, 95% CI 425.0-1600.6 minutes; Figure 3). Of 36 sepsis events, 34 cases were triggered by derangements in the heart rate, respiratory rate, or temperature—heart rate alone (n=1); temperature alone (n=1); heart rate and temperature (n=23);

respiratory rate and temperature (n=2); heart rate, respiratory rate, and temperature (n=7); and unknown (n=2).

Secondary Outcome Measures

There were very few inpatient deaths (n=1) and admissions to level II/III care (n=5) across both study arms. The length of hospital stay was on average 1.3 days shorter in patients who had continuous monitoring (13.3 days, 95% CI 11.3-15.3 days vs 14.6 days, 95% CI 11.5-17.7 days). The rate of readmissions within 30 days of discharge was lower in the continuous monitoring group (11.4%, 95% CI 6.16-16.7 vs 20.9, 95% CI 12.3-29.5; Figure 3).

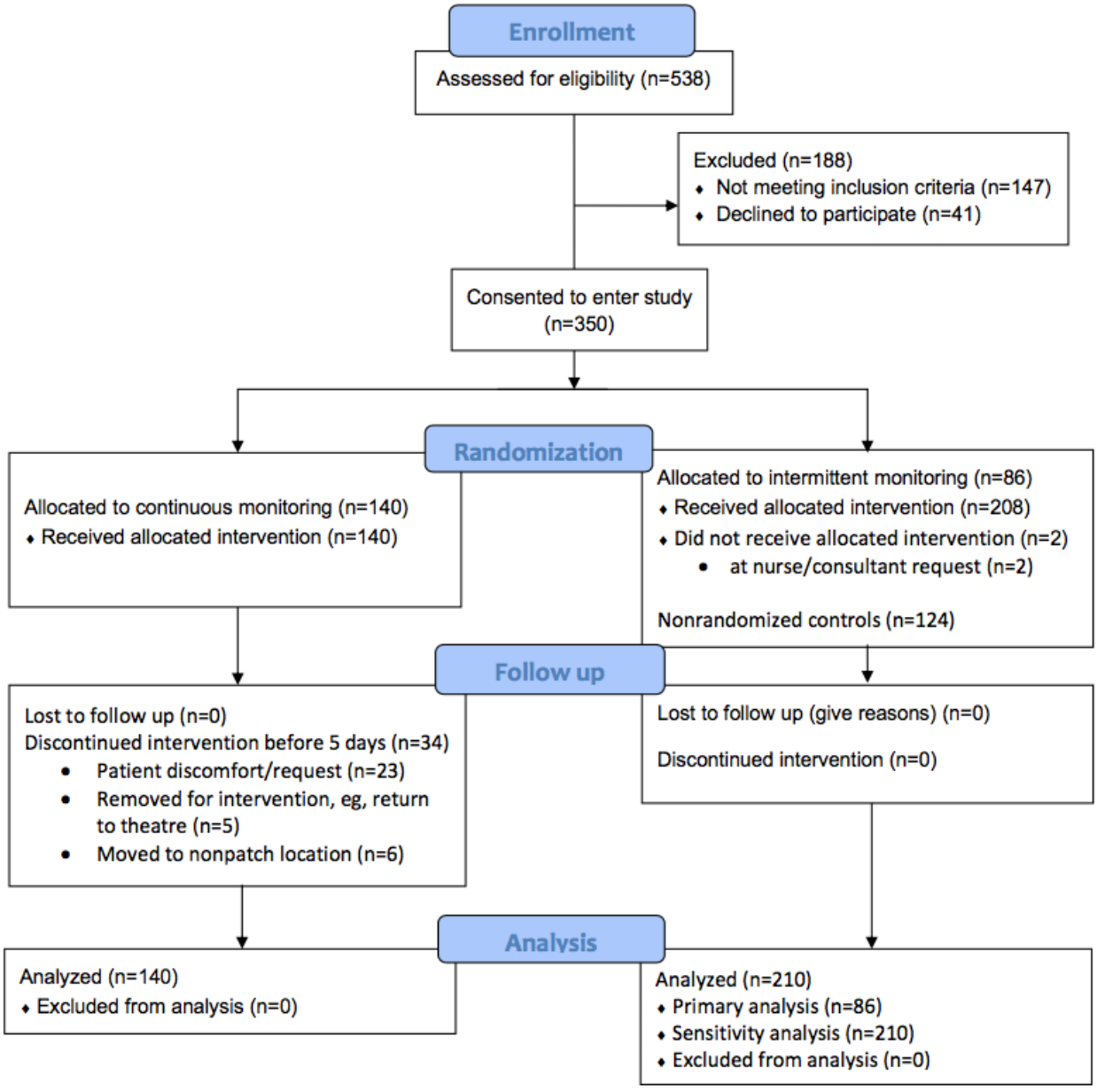
Exploratory Analysis

When the 2 nonrandomized bays were analyzed alongside the 6 randomized bays, the results were very comparable with narrower CIs (see Multimedia Appendix 3).

Patient Acceptability and Compliance

Overall, 41% (58/140) patients in the continuous monitoring group returned a short questionnaire; the results are shown in Figure 4. The majority of patients found the patch to be comfortable (47/57, 82%) and reported feeling safer while wearing the patch (46/56, 82%).

Figure 2. The Consolidated Standards of Reporting Trials flow diagram for the trial.



Patients in the continuous monitoring group wore the patch for an average of 5 (range: 1-24) days. Of 142 patients who wore the monitoring patch, 34 had the continuous monitoring discontinued early (Figure 2); 23 of these were at patient request. Two patients developed a rash under the electrodes, 18 patients found it itchy or bothersome, and 3 patients did not offer a reason for removing the patch.

Progression Criteria

In the pilot trial, 350 patients were recruited within 7 months, which is well within the recruitment target. Adherence to protocol was acceptable; 75.7% (106/140) of patients in the intervention arm wore the patch for at least 5 days.

The low rate of sepsis events across both arms of the study has meant that the CIs around the mean time to antibiotics are wide, and it has not been possible to accurately estimate the intercluster correlation coefficient for this endpoint from the study data. As such, it is unlikely that the time to antibiotics in cases of sepsis is a suitable outcome measure to inform the sample size of a definitive trial using the same protocol.

Table 1. Baseline patients' characteristics.

Characteristics	SensiumVitals + intermittent monitoring (n=140)	Intermittent monitoring alone (n=86)
Males, n (%)	76 (54.3)	39 (45.4)
Females, n (%)	64 (45.7)	47 (54.6)
Age (years), mean, range	65.2, 24-94	63.7, 21-92
American Society of Anaesthesiologists score for preoperative functional status, n (%)		
1	9 (6.4)	9 (10.5)
2	62 (44.3)	35 (40.7)
3	42 (30.0)	22 (25.6)
4	3 (2.1)	3 (3.5)
Not documented	24 (17.1)>	17(19.8)
Emergency admissions, n (%)	70 (50)	44 (51.2)
Elective admissions, n (%)	70 (50)	42 (48.8)
Surgical intervention, n (%)	103 (73.6)	62 (72.1)
Medical outliers, n (%)	19 (13.6)	14 (16.3)

Table 2. A summary of outcome measures.

Outcome measures	SensiumVitals + intermittent monitoring (n=140)	Intermittent monitoring alone (n=86)
Complications[a], n (%)	102 (72.9)	57 (66.3)
Major complications[b], n (%)	8 (5.7)	5 (5.8)
Sepsis events, n (%)	24 (17.1)	12 (14.0)
Time (min) to antibiotics in cases of sepsis[c], mean (95% CI)	626.0 (431.7-820.3)	1012.8 (425.0-1600.6)
Level II or III admissions		
n (%)	3 (2.1)	2 (2.3)
95% CI	0-4.54	0-5.51
Length of stay (in days) in hospital, mean (95% CI)	13.3 (11.3-15.3)	14.6 (11.5-17.7)
Readmissions		
n (%)	16 (11.4)	18 (20.9)
95% CI	6.16-16.7	12.3-29.5
Inpatient deaths, n (%)	1 (0.7)	0 (0)

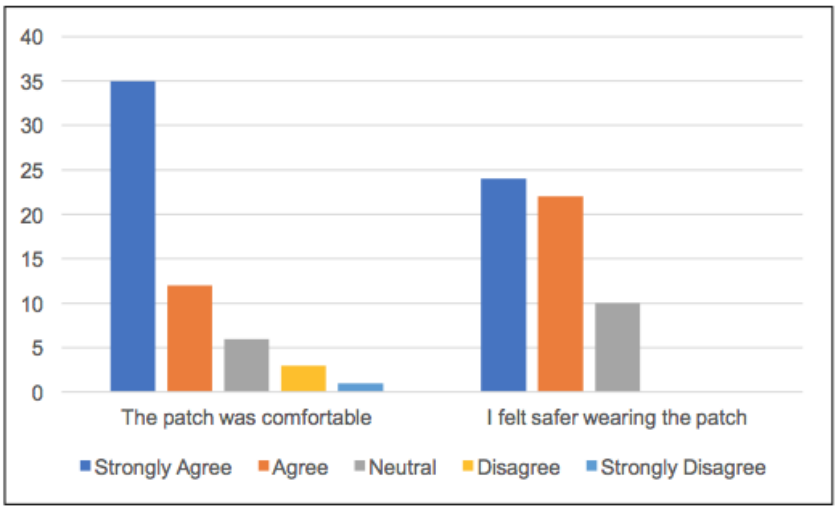
[a]All.

[b]Clavien-Dindo>2.

[c]SensiumVitals + intermittent monitoring (n=22); intermittent monitoring alone (n=12)

Figure 3. Scatter graphs to show mean (x) and 95% CIs between trial arms for time to antibiotics in sepsis, length of hospital stay, and 30-day readmission rate. NEWS: National Early Warning Score.

Figure 4. Patient responses to the questionnaire.



Discussion

In this single-center, randomized controlled pilot trial, surgical patients with evidence of sepsis tended to receive antibiotics faster if they received continuous vital signs monitoring compared with those receiving usual intermittent monitoring alone. Patients receiving continuous vital signs monitoring had a shorter average length of hospital stay and were less likely to require readmission within 30 days of discharge. Patients found the monitoring device to be acceptable in terms of comfort and perceived safety.

The findings must be interpreted within the limitations of the study. A formal sample size calculation was not possible given the lack of data surrounding the primary outcome measure, and so the findings were limited to descriptive statistics; no formal statistical comparison was possible [14]. Although the wide, overlapping CIs suggest that a statistically significant difference between the 2 groups is unlikely, with a larger sample size and increased study power, it is possible that the observed trends might become statistically significant. In addition, the relatively small number of sepsis cases means there is likely to be an imbalance in prerandomization variables, which would require an adjustment in a formal analysis.

There were very few cases of inpatient mortality or admission to level II/III care, making comparisons between the monitoring arms difficult. One explanation for this low event rate is that the population contained a high proportion of low-risk patients—medical outliers and those who did not undergo surgery during their admission. A more striking effect might be evident in a higher-risk population.

The limitations of the randomization technique must also be considered. Ideally, the study data would have been analyzed at the cluster level, but small numbers of patients within each bay necessitated analysis at the individual level. The cluster-randomization methodology led to differences in the baseline demographics of the treatment arms. One of the female bays allocated to receive continuous monitoring had a proportionally lower turnover of patients than the other bays. This led to an imbalance in the male:female ratio between the 2 arms. The fact that the control arm was, on average, 1.5 years younger than the treatment arm may have conferred an advantage to this group.

The potential benefits of continuous monitoring may have been underestimated in this study owing to the exposure to the patch in the intervention arm. Nearly a quarter of the patients who were allocated to receive continuous monitoring did not wear the patch for their entire admission; however, this may reflect what can be truly expected in the clinical environment. The patient-reported acceptability of the device was high in the questionnaire results. This result may be subject to selection bias. A number of patients were missed when they were discharged from hospital outside normal working hours, and enthusiastic patients may have been more likely to complete the questionnaire.

There were other challenges to implementing the technology. There was initially an unacceptably high number of alerts sent to nursing staff; these were reduced by 90% by adjusting the alarm thresholds to more clinically appropriate levels and increasing the intervals between reminder alerts. Engagement with the new system varied between nursing staff but was aided by support from senior ward nurses. Engagement was further increased with the implementation of changes suggested by the nursing staff themselves such as smaller devices and louder alert tones.

There are few clinical evaluations of continuous vital signs monitoring in the literature [8]. The preponderance of observational studies means that causal associations between interventions and patient outcomes have to be interpreted with care. The 3 largest randomized controlled trials of continuous monitoring report conflicting results, illustrating the difficulties in evaluating such complex interventions. The potential benefit of the additional monitoring may be negated by inadequacies in other areas, such as staffing levels, escalation protocols, and nursing compliance [15]. Demonstrating clinical benefit will likely require large, well-controlled studies in high-risk populations to find significant differences in clinical outcomes such as critical care admissions. This is important as these systems are not without financial cost. System prices are around US $1500, and the cost of disposable patches varies [7]. Further

research is required to determine with certainty whether continuous postoperative monitoring offers a significant benefit over intermittent monitoring and can be justified for routine care in terms of cost-effectiveness.

In conclusion, this study has demonstrated the practicability and acceptability of implementing a remote continuous monitoring system in the general surgical ward setting. There is a trend toward clinical benefit. The findings of this study will be used to inform the protocols for further evaluations. Follow-up studies should be individually randomized and stratified to minimize the baseline differences between the 2 treatment arms and include a high-risk population with a high rate of adverse events. Furthermore, rare outcomes, such as mortality, should be avoided in preference of endpoints that are common to all participants such as the length of hospital stay.

Acknowledgments

The research is supported by the National Institute for Health Research (NIHR) infrastructure at Leeds. The study is supported by the Health Foundation. The Health Foundation is an independent charity committed to bringing about better health and health care for people in the UK.

CD is in possession of a Doctoral Research Fellowship (CD, DRF-2016-09-037) supported by the NIHR. DGJ received funding support through an NIHR Research Professorship. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, Health Education England, or the Department of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 578KB - jmir_v20i12e10802_app1.pdf]

Multimedia Appendix 2

Patient information sheet.

[PDF File (Adobe PDF File), 77KB - jmir_v20i12e10802_app2.pdf]

Multimedia Appendix 3

Exploratory analyses and questionnaire.

[PDF File (Adobe PDF File), 83KB - jmir_v20i12e10802_app3.pdf]

References

1. Bartels K, Karhausen J, Clambey ET, Grenz A, Eltzschig HK. Perioperative organ injury. Anesthesiology 2013 Dec;119(6):1474-1489 [FREE Full text] [doi: 10.1097/ALN.0000000000000022] [Medline: 24126264]
2. International Surgical Outcomes Study group. Global patient outcomes after elective surgery: prospective cohort study in 27 low-, middle- and high-income countries. Br J Anaesth 2016 Oct 31;117(5):601-609 [FREE Full text] [doi: 10.1093/bja/aew316] [Medline: 27799174]
3. Michard F, Gan TJ, Kehlet H. Digital innovations and emerging technologies for enhanced recovery programmes. Br J Anaesth 2017 Jul 01;119(1):31-39 [FREE Full text] [doi: 10.1093/bja/aex140] [Medline: 28605474]
4. Goldhill DR, McNarry AF. Physiological abnormalities in early warning scores are related to mortality in adult inpatients. Br J Anaesth 2004 Jun;92(6):882-884 [FREE Full text] [doi: 10.1093/bja/aeh113] [Medline: 15064245]
5. Ridley S. The recognition and early management of critical illness. Ann R Coll Surg Engl 2005 Sep;87(5):315-322 [FREE Full text] [doi: 10.1308/003588405X60669] [Medline: 16176687]
6. Downey CL, Tahir W, Randell R, Brown JM, Jayne DG. Strengths and limitations of early warning scores: A systematic review and narrative synthesis. Int J Nurs Stud 2017 Nov;76:106-119. [doi: 10.1016/j.ijnurstu.2017.09.003] [Medline: 28950188]
7. Hofmann B, Welch HG. New diagnostic tests: more harm than good. BMJ 2017 Dec 18;358:j3314. [doi: 10.1136/bmj.j3314] [Medline: 28720607]
8. Cardona-Morrell M, Prgomet M, Turner RM, Nicholson M, Hillman K. Effectiveness of continuous or intermittent vital signs monitoring in preventing adverse events on general wards: a systematic review and meta-analysis. Int J Clin Pract 2016 Oct;70(10):806-824. [doi: 10.1111/ijcp.12846] [Medline: 27582503]

9. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]
10. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
11. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
12. Sealed Envelope Ltd. 2017. Create a blocked randomisation list URL: https://www.sealedenvelope.com/simple-randomiser/v1/lists [accessed 2018-11-06] [WebCite Cache ID 73irpQkpL]
13. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, International Sepsis Definitions Conference. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Intensive Care Med 2003 Apr;29(4):530-538. [doi: 10.1007/s00134-003-1662-x] [Medline: 12664219]
14. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j.2002.384.doc.x] [Medline: 15189396]
15. Watkinson PJ, Barber VS, Price JD, Hann A, Tarassenko L, Young JD. A randomised controlled trial of the effect of continuous electronic physiological monitoring on the adverse event rate in high risk medical and surgical patients. Anaesthesia 2006 Nov;61(11):1031-1039 [FREE Full text] [doi: 10.1111/j.1365-2044.2006.04818.x] [Medline: 17042839]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 17.04.18; peer-reviewed by T Bonnici, M Mohammed; comments to author 02.08.18; revised version received 22.08.18; accepted 24.09.18; published 11.12.18.

Please cite as:
Downey C, Randell R, Brown J, Jayne DG
Continuous Versus Intermittent Vital Signs Monitoring Using a Wearable, Wireless Patch in Patients Admitted to Surgical Wards: Pilot Cluster Randomized Controlled Trial
J Med Internet Res 2018;20(12):e10802
URL: https://www.jmir.org/2018/12/e10802/
doi:10.2196/10802
PMID:30538086

Original Paper

Continuous Temperature-Monitoring Socks for Home Use in Patients With Diabetes: Observational Study

Alexander M Reyzelman[1], DPM; Kristopher Koelewyn[2], DPM; Maryam Murphy[3], BS; Xuening Shen[4], MS; E Yu[4], BS; Raji Pillai[5], PhD; Jie Fu[4], BS; Henk Jan Scholten[3], LLM; Ran Ma[3], BS

[1]California School of Podiatric Medicine, Samuel Merritt University, San Francisco, CA, United States

[2]Northport VA Medical Center, Northport, CA, United States

[3]Siren Care Inc, San Francisco, CA, United States

[4]Siren Care (Shanghai) Information Technology Co Ltd, Shanghai, China

[5]Medical Affairs Consulting Inc, Menlo Park, CA, United States

Corresponding Author:
Alexander M Reyzelman, DPM
California School of Podiatric Medicine
Samuel Merritt University
2299 Post Street
San Francisco, CA, 94115
United States
Phone: 1 415 345 1195
Email: areyzelman@samuelmerritt.edu

Abstract

Background: Over 30 million people in the United States (over 9%) have been diagnosed with diabetes. About 25% of people with diabetes will experience a diabetic foot ulcer (DFU) in their lifetime. Unresolved DFUs may lead to sepsis and are the leading cause of lower-limb amputations. DFU rates can be reduced by screening patients with diabetes to enable risk-based interventions. Skin temperature assessment has been shown to reduce the risk of foot ulceration. While several tools have been developed to measure plantar temperatures, they only measure temperature once a day or are designed for clinic use only. In this report, wireless sensor-embedded socks designed for daily wear are introduced, which perform continuous temperature monitoring of the feet of persons with diabetes in the home environment. Combined with a mobile app, this wearable device informs the wearer about temperature increases in one foot relative to the other, to facilitate early detection of ulcers and timely intervention.

Objective: A pilot study was conducted to assess the accuracy of sensors used in daily wear socks, obtain user feedback on how comfortable sensor-embedded socks were for home use, and examine whether observed temperatures correlated with clinical observations.

Methods: Temperature accuracy of sensors was assessed both prior to incorporation in the socks, as well as in the completed design. The measured temperatures were compared to the reference standard, a high-precision thermostatic water bath in the range 20°C-40°C. A total of 35 patients, 18 years of age and older, with diabetic peripheral neuropathy were enrolled in a single-site study conducted under an Institutional Review Board–approved protocol. This study evaluated the usability of the sensor-embedded socks and correlated the observed temperatures with clinical findings.

Results: The temperatures measured by the stand-alone sensors were within 0.2°C of the reference standard. In the sensor-embedded socks, across multiple measurements for each of the six sensors, a high agreement (R^2=1) between temperatures measured and the reference standard was observed. Patients reported that the socks were easy to use and comfortable, ranking them at a median score of 9 or 10 for comfort and ease of use on a 10-point scale. Case studies are presented showing that the temperature differences observed between the feet were consistent with clinical observations.

Conclusions: We report the first use of wireless continuous temperature monitoring for daily wear and home use in patients with diabetes and neuropathy. The wearers found the socks to be no different from standard socks. The temperature studies conducted show that the sensors used in the socks are reliable and accurate at detecting temperature and the findings matched clinical observations. Continuous temperature monitoring is a promising approach as an early warning system for foot ulcers, Charcot foot, and reulceration.

(J Med Internet Res 2018;20(12):e12460) doi:10.2196/12460

KEYWORDS

diabetes; diabetic foot ulcer; continuous temperature monitoring; Charcot arthropathy; digital health; wearable; neurofabric; mobile phone; wireless; Bluetooth; neuropathy; home use

Introduction

An estimated 30.3 million people in the United States (ie, 9.4% of the US population) have been diagnosed with diabetes, according to the Centers for Disease Control and Prevention [1].

Complications of Diabetes, Foot Ulcers, and Prognosis

Diabetes damages blood vessels and nerves, particularly in the feet, and can lead to severe infections that are difficult to treat. About 25% of people with diabetes will experience a diabetic foot ulcer (DFU) in their lifetime [2-4]. When circulation is so poor that a foot ulcer fails to heal or when treatment fails to stop the spread of an infection, sepsis can result. In such cases, amputation is often necessary. Diabetes is the leading cause of lower-limb amputations; DFUs precede approximately 84% of nontraumatic major amputations among people with diabetes [5-7]. The rates of recurrence of foot ulcers are very high, being greater than 40% after one year and 60% within three years [2,8].

Charcot foot, also called Charcot arthropathy, is one of the most debilitating outcomes of diabetes [9]. The condition causes increased blood flow to the foot and increased bone resorption. Immediately keeping all weight off, or off - loading, of incipient Charcot foot appears to minimize fractures and incapacitating deformities. However, there is a potential for delayed diagnosis and therapeutic intervention as plain x - rays may not show fractures at the early stages [10].

Cost to the Health Care System

Diabetic foot ulcers result in considerable cost to the health care system when immediate ulcer episodes, social services, home care, and subsequent ulcer episodes are taken into consideration. Patients with a DFU were seen by their outpatient health care provider about 14 times per year and were hospitalized about 1.5 times per year. The cost of care for these patients was substantial, at about US $33,000 for total reimbursement of all Medicare services per year [11].

The total direct and indirect estimated cost of diagnosed diabetes in the United States in 2012 was US $245 billion. After adjusting for age group and sex, average medical expenditures among people with diagnosed diabetes were about 2.3 times higher than expenditures for people without diabetes [12] .

The national inpatient and emergency department bill summed to US $8.78 billion per year, averaging US $115,957 per case for major amputations [13].

Diagnosing Foot Ulcers

Screening patients with diabetes to identify those at risk for foot ulceration has been shown to be beneficial [3]. DFU rates can be reduced with screening and appropriate intervention [4,14]. Self-care is a critical factor in detecting early signs of ulcers and injury. However, visual inspection has limitations (eg, patients with obesity or visual impairment cannot see their feet easily); hence, it is not very effective to identify the early signs. A recent study using a remote foot-temperature-monitoring system showed the ability to detect 97% of nontraumatic DFUs five weeks before they presented to the participant and/or clinician [15,16]. Additional options for detecting ulcers early in order to treat and heal ulcer wounds successfully may help prevent lower-limb amputations.

In diabetic foot complications such as foot ulcers and osteomyelitis, elevated temperatures in regions of the foot have been shown to be a precursor for ulceration [17]. In Charcot foot cases, increased plantar foot temperature is observed and is strongly correlated with the location of arthropathy. Temperatures decreased in a predictable manner as acute arthropathy resolved [18,19].

Thus, skin temperature assessment in persons with diabetes is a valuable tool for assessing inflammation in diabetic feet, as well as its resolution [20,21]. Home temperature monitoring has been shown to be an effective approach as an early warning system, to provide patients with objective feedback so they can modify their activity and protect their foot before ulcers develop; such monitoring is included in the International Working Group on Diabetic Foot clinical practice guidelines [16,22].

A handheld, infrared, dermal thermometer was designed to take temperatures on the bottom of both feet at six different spots each morning and compare these temperatures from spot to spot. Temperature differences of 4°F (2.22°C) or higher observed at comparable spots between the feet serve as an early sign of DFUs [16,23-25]. However, this tool has shown limited adoption. A reported shortcoming is that the manual temperature measurement on specific spots on the foot is subjective: asymmetric analysis tended to find false abnormal areas when the left and right feet had different sizes and shapes.

Digital health is a vast and burgeoning field and spans several aspects of health management. With the advent of the Internet of Things and the Internet of Medical Things coupled with smart devices, the potential for improved home care for medical applications is fast becoming a reality. Such devices can facilitate the management of chronic conditions at home, including the effective and timely management of DFUs. Diabetic foot scanners and voice-enabled scales are in development [26]. A “smart mat” allows daily measurement of plantar temperature, compares the temperature profile of the two feet, and aims to identify regions with increased temperature, detecting potential ulcer formation at an early stage [15,27].

Innovation in wearables has led to the development of “smart socks,” with embedded sensors for measuring temperature and reporting increases. A recent report describes socks made

entirely of optical fiber [28]. While laboratory testing verified the accuracy of the sensors within, the fragility of the optical fiber resulted in limited usability. Another report describes socks that measure temperature at 10-minute intervals and contain wires [29].

Continuous Temperature Monitoring and Goal of This Study

All the tools described above are designed to measure temperatures once a day or at long intervals, are for clinic use only, or include wired data transmission. Once-a-day measurements present a risk of giving false positives. Continuous monitoring allows the assessment of temperature over longer periods, taking into consideration varying levels of activity over time, and thus has a greater potential to report consistent and clinically relevant temperature increases [30]. Continuous temperature monitoring can reduce false positives and has the potential to further improve home care and early detection.

Here, we introduce wireless sensor-embedded socks, made of neurofabric textile with microsensors embedded directly into the fabric—for continuous temperature monitoring of the feet of people with diabetes—and wireless reporting. They are designed to be easy to use and are washable as well as reusable.

A pilot study was conducted to assess (1) how comfortable sensor-embedded socks were for daily use and (2) whether observed temperatures correlated with clinical observations. Illustrative cases are presented.

Methods

Sensor-Embedded Socks

The socks are made of “smart textile”: textile with microsensors woven directly into the fabric (Siren Diabetic Socks, Neurofabric, Siren Care Inc, San Francisco, CA). These virtually invisible sensors are seamlessly integrated into the socks to monitor temperature changes on the bottom of the feet. The sensor-embedded socks are designed to be reusable and are machine washable and dryable.

The sensors embedded in the socks are connected to a small tag on the sock, which encases a microcontroller unit, battery, and Bluetooth chip (see Figure 1A). The six sensors take temperature measurements at 10-second intervals to track temperature increases at the bottom of the user's feet, specifically at the hallux; metatarsal points (MTPs) 1, 3, and 5; midfoot; and heel (see Figure 1B). The data are stored in the tag and sent via Bluetooth to the phone paired with each pair of socks.

The mobile phone app can be programmed to generate alerts when the user’s feet show temperature increases that could be a warning sign of a potential ulcer (see Figure 2). In this study, the mobile phone app displayed temperature readings to the user, but alerts were not generated.

Assessment of Accuracy of Sensors Embedded in Socks in Detecting Temperature

Sensors were tested prior to and after incorporation in the socks using a high-precision, thermostatic water bath (Zhejiang Jinbo Electronic Co, Ltd, China) and verified with a 0.01°C high-precision mercury thermometer. The stand-alone sensors were tested at four temperatures: 20°C, 25°C, 37°C, and 45°C.

The sensors woven into the socks were tested in the range 20°C-40°C. Three pairs of socks were tested by immersion in the thermostatic water bath for 10 seconds. The temperature recorded in the sock tag was compared with the reference standard.

Figure 1. (A) Image of socks with tag (circled) containing battery, microcontroller unit, and Bluetooth chip. (B) Bottom of socks where sensors are located at the hallux (sensor 1), metatarsal points 1,3, and 5 (sensors 2-4), midfoot (sensor 5), and heel (sensor 6).

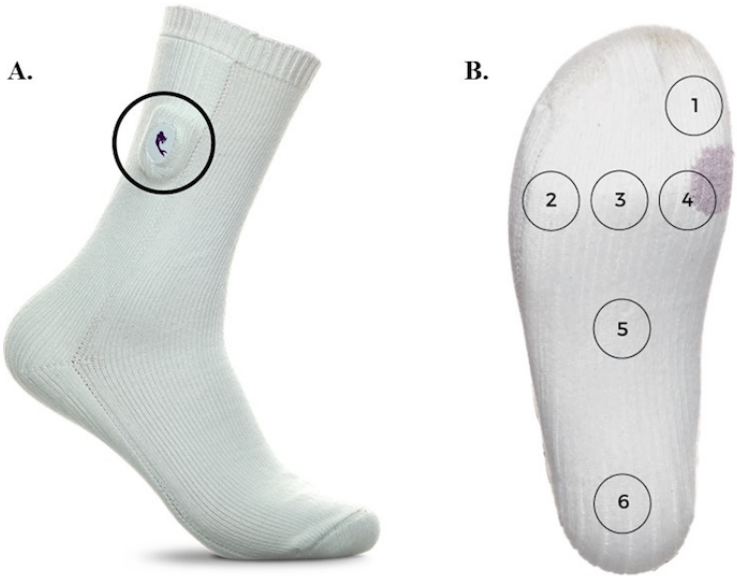


Figure 2. Typical workflow for sensor-embedded socks. The left-hand image shows a temperature-sensing sock: passive continuous monitoring of six key locations occur on the foot. The center image represents continuous monitoring: algorithms monitor temperature reading and generate alerts. The right-hand image displays the patient interface of the app, used for viewing the alerts.

Assessment of Sensor-Embedded Socks Worn by Patients With Diabetes

A single-site study was conducted under an Institutional Review Board-approved protocol to evaluate the usability of the smart socks for patients with diabetic peripheral neuropathy (DPN). Informed consent was obtained from all patients.

A total of 35 patients, 18 years of age and older, from a private clinical practice were enrolled into three groups based on patient-reported medical history and/or medical documents. The groups were as follows: (1) Group 1 included subjects with DPN and no previous history of ulcers (n=11), (2) Group 2 included subjects with DPN and a previous history of ulcers (n=13), and (3) Group 3 included subjects with DPN and a current preulcer as determined by the investigator (n=11).

Subjects participated in two clinic visits. In the first visit, screening procedures were conducted, which included the following: a general physical exam performed by a board-certified podiatrist, visual foot inspection, digital photographs of both feet, and medical history intake. Subjects were provided with the socks and were given an Android mobile phone with the app needed for temperature monitoring. The socks were wirelessly connected with the mobile phone via Bluetooth. The patients were instructed to wear the socks continuously for 6 hours, after which the socks could be removed. The data were streamed via Bluetooth directly to the Android app installed on the phone provided to the patient during the screening and initiation visit. All data were stored in the sock tag and sent via Bluetooth to the phone paired to the socks that were assigned to the enrolled patient.

At the second or end-of-study visit to the clinic—7 days plus or minus 2 days from the screening and initiation visit—the socks were returned to the investigator and the patient was examined for potential adverse reactions. An exit questionnaire was completed to obtain usability information from the patient on the comfortableness of the socks, the ease of Android app use, and the practicality of integrating this specific system into the patient's everyday life. Upon exit from the trial, data were exported from the Android phone to a secure laptop for analysis. All data collected were deidentified and only subject numbers were used for the duration of the trial.

Results

Accuracy of Sensor-Embedded Socks

Testing of Sensors not Incorporated in Socks

A total of 36 stand-alone sensors were tested in a high-precision thermostatic water bath for 10 seconds at four temperatures: 20°C, 25°C, 37°C, and 45°C. The results are shown in Table 1. The temperatures measured by the sensors were within 0.2°C of the reference standard, demonstrating the high accuracy of the sensors used in the socks.

Table 1. Stand-alone sensor temperature measurements at four water bath temperatures.

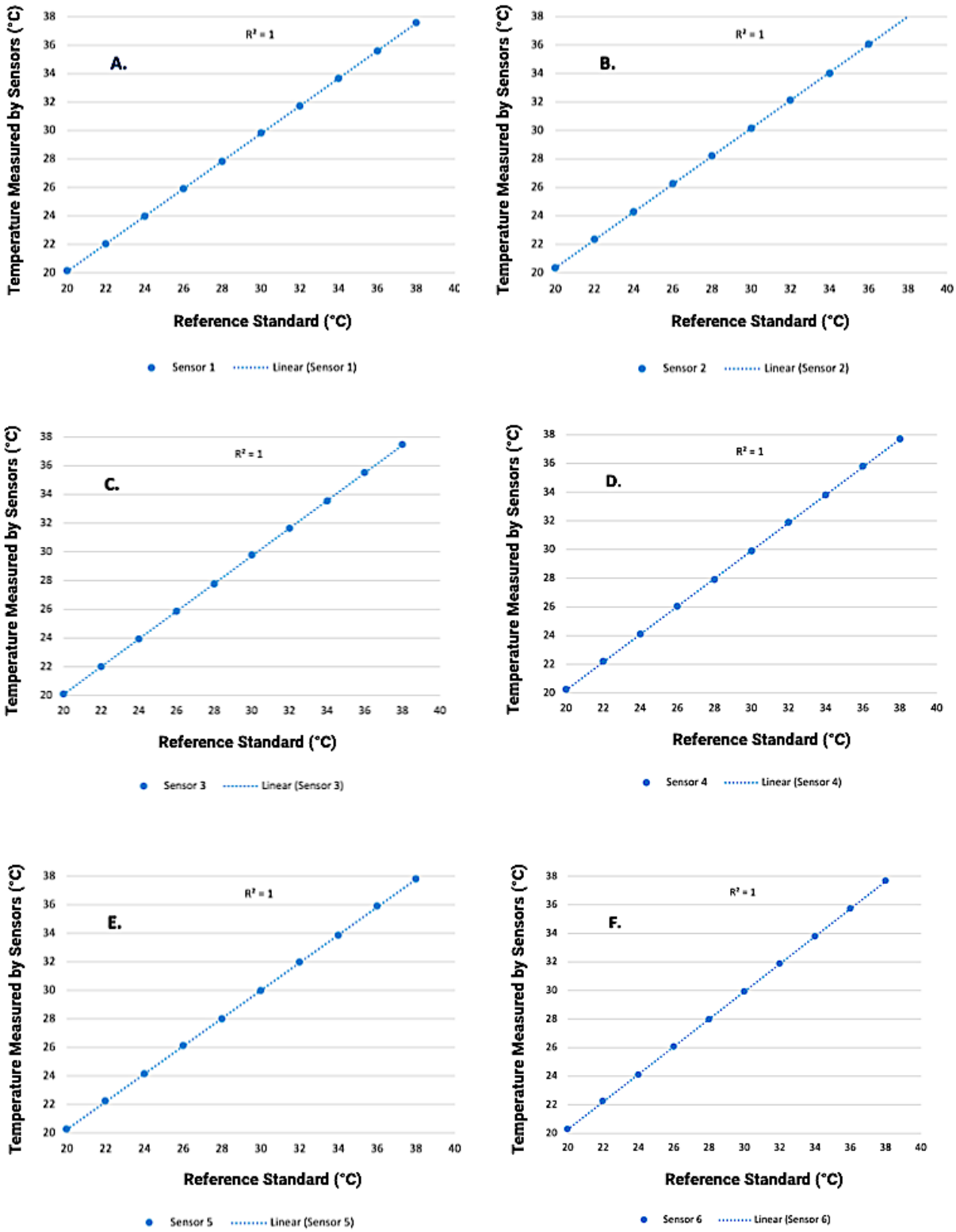
Temperature of water bath (reference standard), °C	Stand-alone sensor temperature, °C		
	Mean (SD)	Minimum	Maximum
20	20.08 (0.04)	19.99	20.13
25	25.11 (0.06)	25.00	25.18
37	37.03 (0.07)	36.91	37.10
45	44.92 (0.08)	44.81	45.01

Testing of Sensors Embedded in Socks

Three pairs of socks with embedded sensors were tested by immersion in a thermostatic water bath for 10 seconds, in the range 20°C-40°C. The temperature recorded in the sock tag was compared with the reference standard. The average of 18 measurements for each of six sensors is displayed in Figure 3.

For each of the six sensors, there was a high agreement (R^2=1) between temperatures measured and the reference standard, establishing that sensor-embedded socks can accurately measure temperature across the wide range tested of 20°C-40°C.

Figure 3. Panels A-F show data for sensors 1-6, positioned at the hallux, metatarsal point 1, metatarsal point 3, metatarsal point 5, arch (midfoot), and heel, respectively. The x-axis shows the reference standard and the y-axis shows the temperatures measured by the sensors embedded in the socks.



Patient Experience of Sensor-Embedded Socks

Patient Information

A total of 35 patients with diabetes assigned to three groups were included in the study as summarized in Table 2.

User Experience

Patients wore the socks at home for 3-21 hours (median 7). Some patients wore the socks at night and slept in them. Upon their return visit to the clinic, they returned the socks and provided feedback via the exit questionnaire on different aspects of the socks, such as design, usefulness, and comfortableness. The results are shown in Figure 4. In these patients' experience while wearing the socks at home, they found the socks to be safe, comfortable, useful, and well-designed. For each question, the median response was 9 or 10 on a 10-point scale, indicating a high level of satisfaction.

Patients also provided feedback on a different scale for a separate set of questions (see Figure 5), some overlapping with the earlier questions. The findings were consistent with those from the 10-point scale. The median value for each of the responses was 4 or 5 on a 5-point scale, confirming a positive experience. Patients reported that the socks felt just like their normal, everyday socks. Notably, patients did not feel they were wearing sensor-embedded socks or that they had to use the socks carefully. Their stated willingness to wear the socks every day underscores the socks' suitability for home use. This suggests that smart textiles, which are used to make the sensor-embedded socks, can seamlessly integrate into the life of the wearer and are suitable for home use.

Table 2. Distribution of patients with diabetes included in the study.

Patient characteristics	Group 1[a] (n=11)	Group 2[b] (n=13)	Group 3[c] (n=11)	Overall (n=35)
Female, n (%)	5 (45)	4 (31)	1 (9)	10 (29)
Male, n (%)	6 (55)	9 (69)	10 (91)	25 (71)
Age (years), median (range)	50 (37-80)	61 (40-71)	64 (50-73)	62 (37-80)
Age when diagnosed with diabetes (years), median (range)	46 (29-70)	46 (22-61)[d]	46 (32-65)	46 (29-65)
Length of time living with diabetes (years), median (range)	8 (1.5-30)	15 (5-47)[d]	13 (4-40)	11 (1.5-47)
Patients with type 1 diabetes, n (%)	0 (0)	0 (0)	2 (18)	2 (6)
Patients with type 2 diabetes, n (%)	11 (100)	13 (100)	9 (82)	33 (94)

[a]Group 1 included subjects with diabetic peripheral neuropathy (DPN) and no previous history of ulcers.

[b]Group 2 included subjects with DPN and a previous history of ulcers.

[c]Group 3 included subjects with DPN and a current preulcer as determined by the investigator.

[d]Age at diagnosis not available for two subjects.

Figure 4. Patients reported on their experience on a scale of 1 to 10 where 10 is "Good" and 1 is "Bad." In the box and whisker plot, the line within the box represents the median, the x in the box represents the mean, the bounds of the box are at the 1st and 3rd quartiles (25% and 75%), the whisker (vertical line) extends to the minimum value, and the dots are outliers.

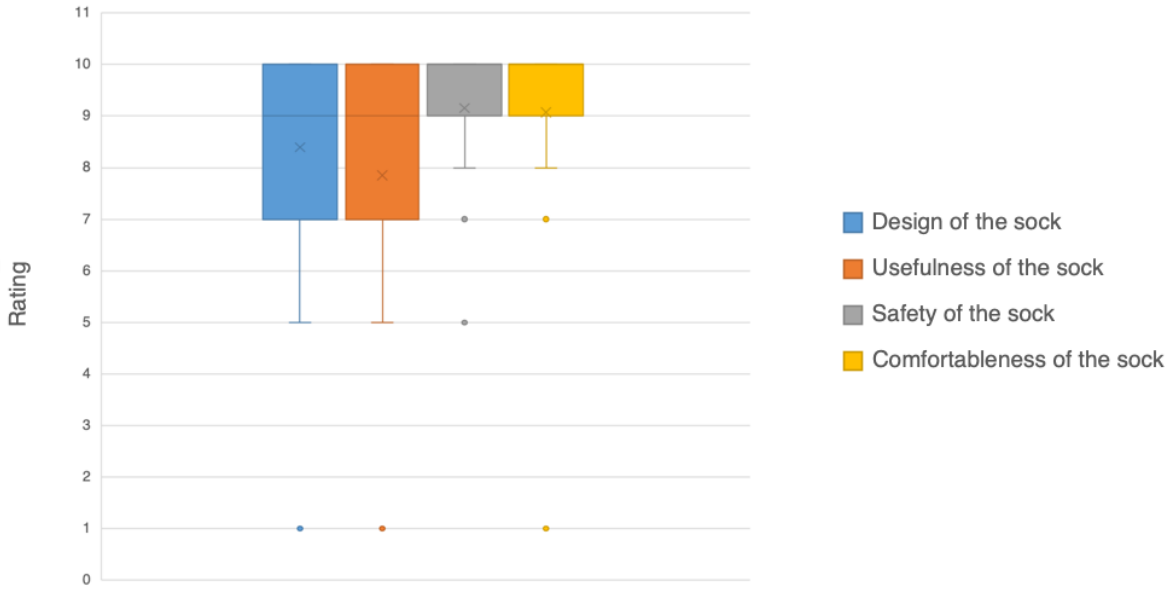


Figure 5. Patients reported on their experience on a scale of 1-5, where 5 is "Completely Agree" and 1 is "Completely Disagree." In each box and whisker plot, the line within the box represents the median, the x in the box represents the mean, the bounds of the box are at the 1st and 3rd quartiles (25% and 75%), the whisker (vertical line) extends to the minimum values, and the dots are outliers.

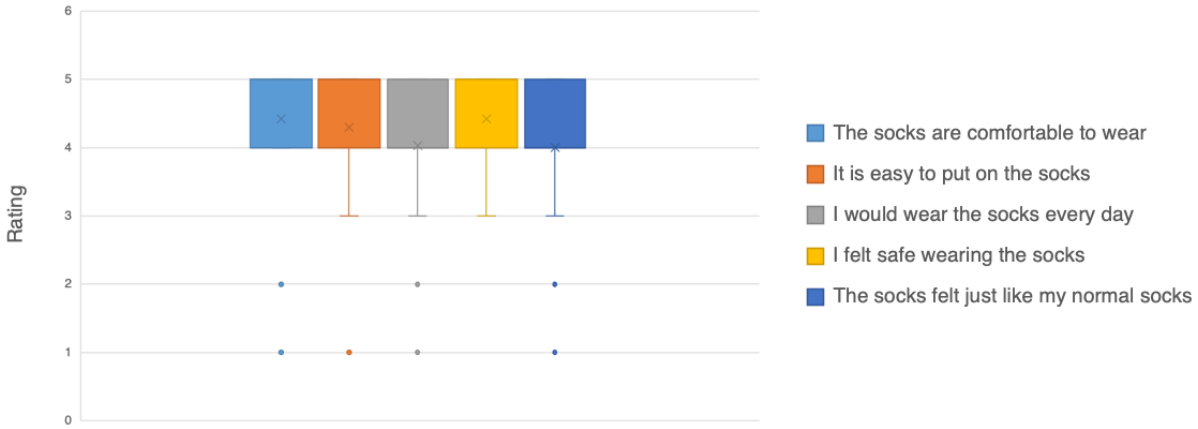


Table 3. Patient responses to statements on the mobile app.

Statements	Response score[a], mean (SD)	Response score, median (IQR[b])
It is easy to connect the sock to the mobile app	3.4 (0.89)	3 (0.75)
The mobile app gives useful information about my feet	3.5 (1.10)	4 (1)
The overview of the feet in the mobile app is easy to understand	3.6 (1.04)	4 (1)
I would use the app every day	3.9 (0.91)	4 (2)

[a]Responses were on a 5-point scale where 5 is "Completely Agree" and 1 is "Completely Disagree."

[a]IQR: interquartile range.

The mobile app was found to be useful and easy to use. The mean and the median responses to key statements on the mobile app are shown in Table 3. Users found it easy to navigate the mobile app and they found the information provided to be informative about their feet.

Case Studies

A few illustrative cases are shown below, one from each of the three study groups.

Case Study 1

Patient 14 (Group 1) is a 64-year-old male diagnosed with type 2 diabetes at 53 years of age. He has no history of foot ulceration or amputation and has experienced neuropathic pain for the past 8 years. His feet showed no visible signs of injury (see Figure 6). Patient 14 wore the socks for 6 hours, during which minor variations in temperature between the contralateral locations were observed with differences of less than 2.2°C or 4°F (see Figure 7). The continuous monitoring of the temperature by the socks show minor variations over the 6-hour period. Consistent with the initial observations and medical history, no temperature elevations were found.

Case Study 2

Patient 30 (Group 2) is a 63-year-old male diagnosed with type 2 diabetes at 45 years of age. He has a history of ulcers and was diagnosed with Charcot arthropathy of the right foot at 57 years of age. Intake photographs (see Figure 8) and examination showed Charcot of the right foot with a collapsed midfoot. Patient 30 wore the sensor-embedded socks for 8 hours, during which period the right foot was consistently warmer than the left foot. The temperatures on the right hallux, MTP 1, MTP 3, MTP 5, and midfoot (arch) were elevated more than 2.2°C or 4°F, up to 8°C (see Figure 9).

Thus, the findings from continuous temperature monitoring are consistent with the patient's medical history and intake evaluation of Charcot of the right foot. This suggests that the clinical assessment of Charcot arthropathy may benefit from this monitoring system, as it provides a temperature map of the entire foot over a long period of time, rather than static and local temperature changes.

Case Study 3

Patient 16 (Group 3) is a 73-year-old male diagnosed with type 2 diabetes at 65 years of age. He has a history of preulcerative lesions. Intake photographs (see Figure 10) and exam indicated a current preulcerative lesion at the right plantar region between the second and third metatarsal. Patient 16 wore the socks for 9 hours, during which higher temperatures were recorded by two of the six sensors, at the positions of the right metatarsals 3 and 5 (see Figure 11). This observation is consistent with the patient's medical records indicating a preulcerative lesion in the MTP 2-3 area. These data suggest that in high-risk patients, continuous monitoring may be able to pick up an injury or preulcerative lesion.

Figure 6. A digital photograph of patient 14's feet show no visible signs of injury.

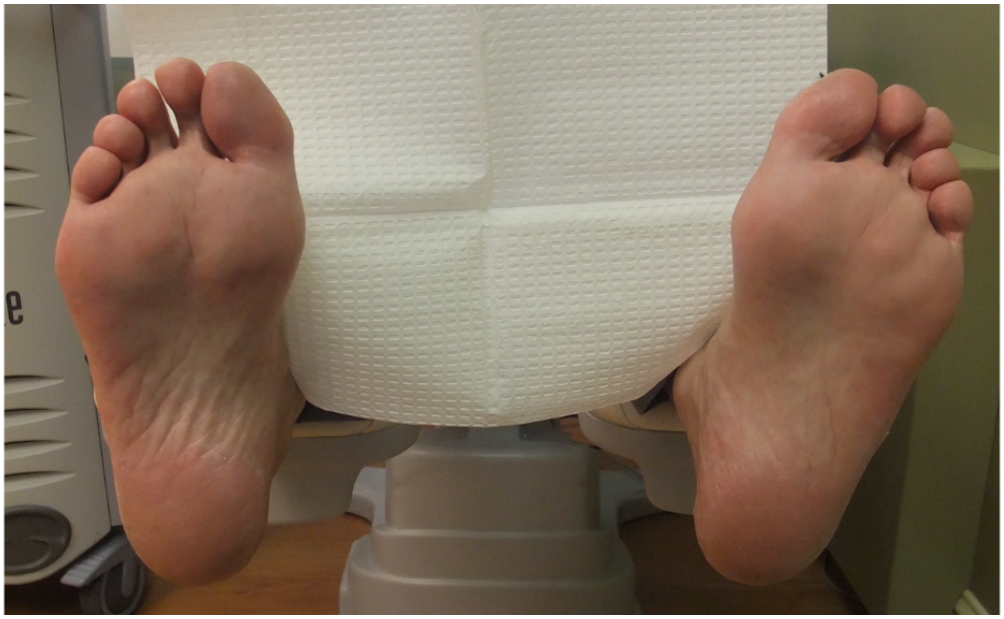


Figure 7. Each line on the graph shows a moving average of the temperature difference (ie, left foot temperature-right foot temperature) for the hallux (blue), metatarsal points 1, 3, and 5 (orange, gray, and yellow, respectively), arch (black), and heel (green). The lines span the period the socks were worn, with time shown on the x-axis. The dashed and dotted red lines show the 2.2°C temperature threshold for the left and right foot, respectively. MTP: metatarsal point.

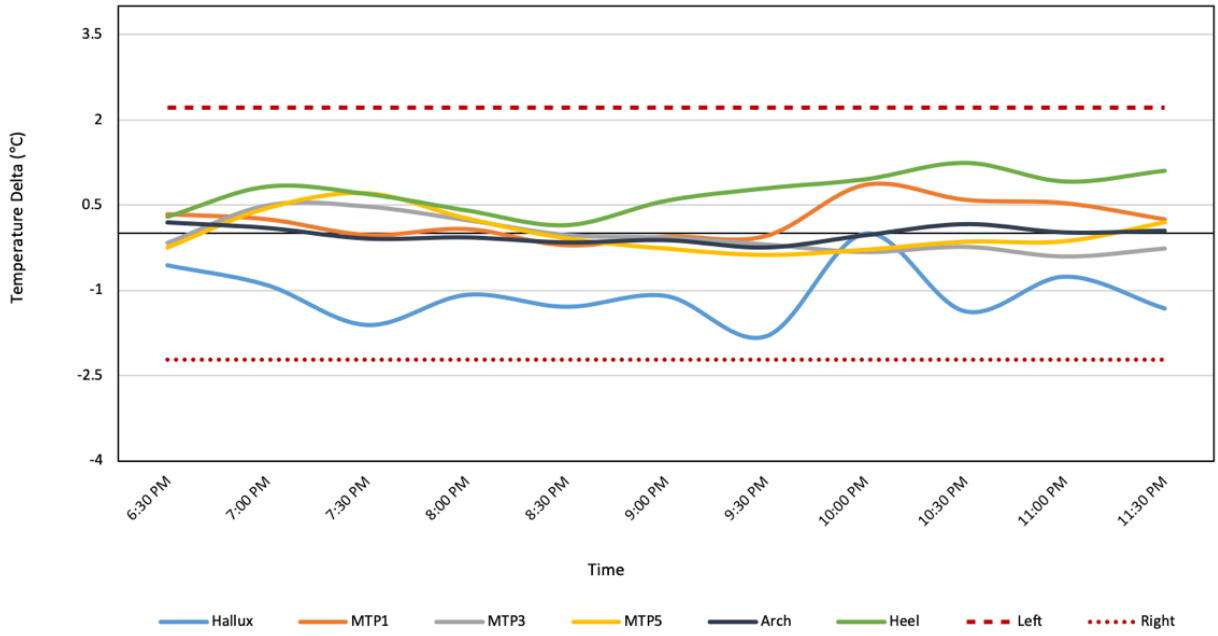


Figure 8. A digital photograph of patient 30's feet show Charcot of the right foot with collapsed midfoot (arch), designated by the red circle.

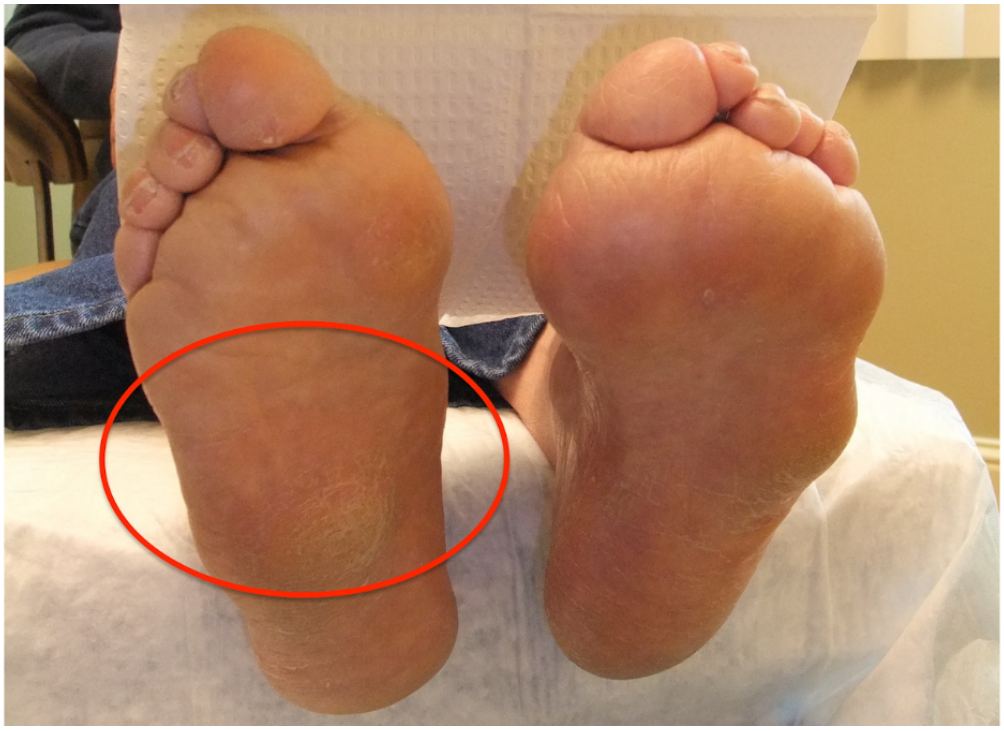


Figure 9. Each line on the graph shows a moving average of the temperature difference (ie, left foot temperature-right foot temperature) for the hallux (blue), metatarsal points 1, 3, and 5 (orange, gray, and yellow, respectively), arch (black), and heel (green). The moving average of the temperature difference shows elevated temperatures of the right foot compared to the left foot at all points except the heel. MTP: metatarsal point.

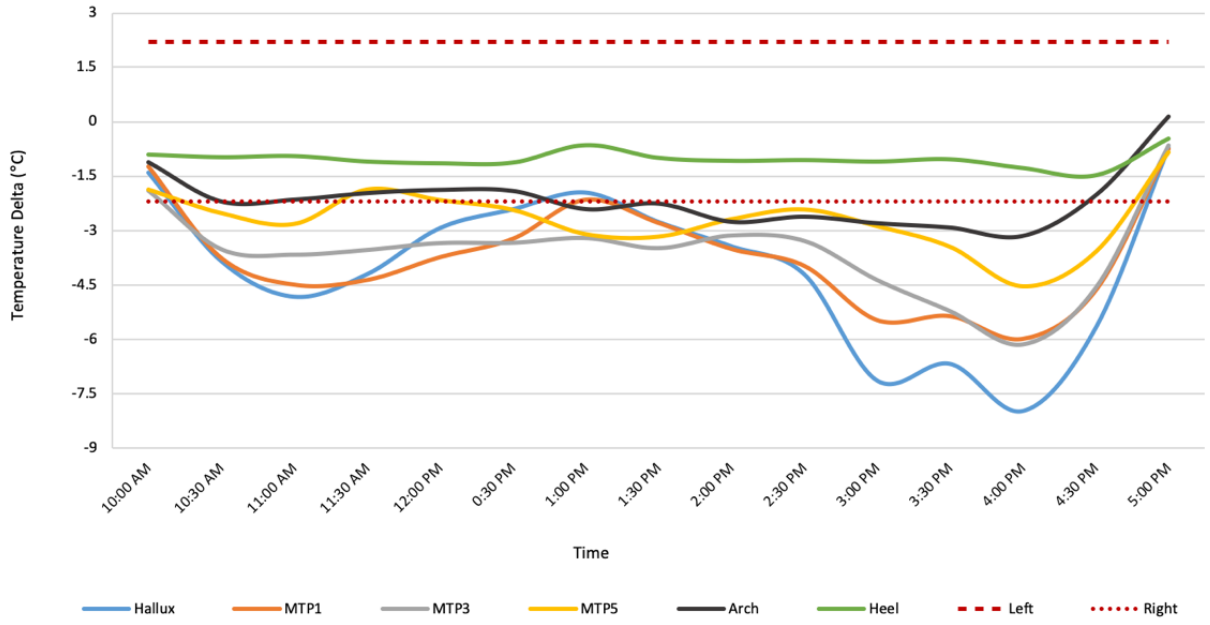


Figure 10. A digital photograph of patient 16's feet show a preulcerative lesion between the second and third metatarsal.

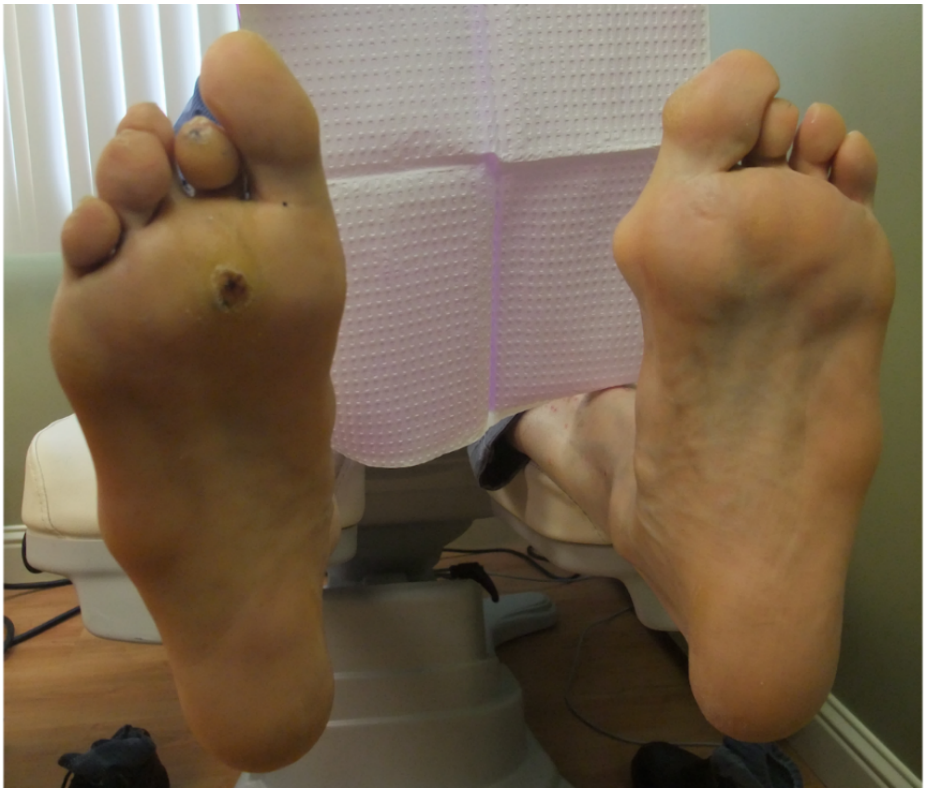
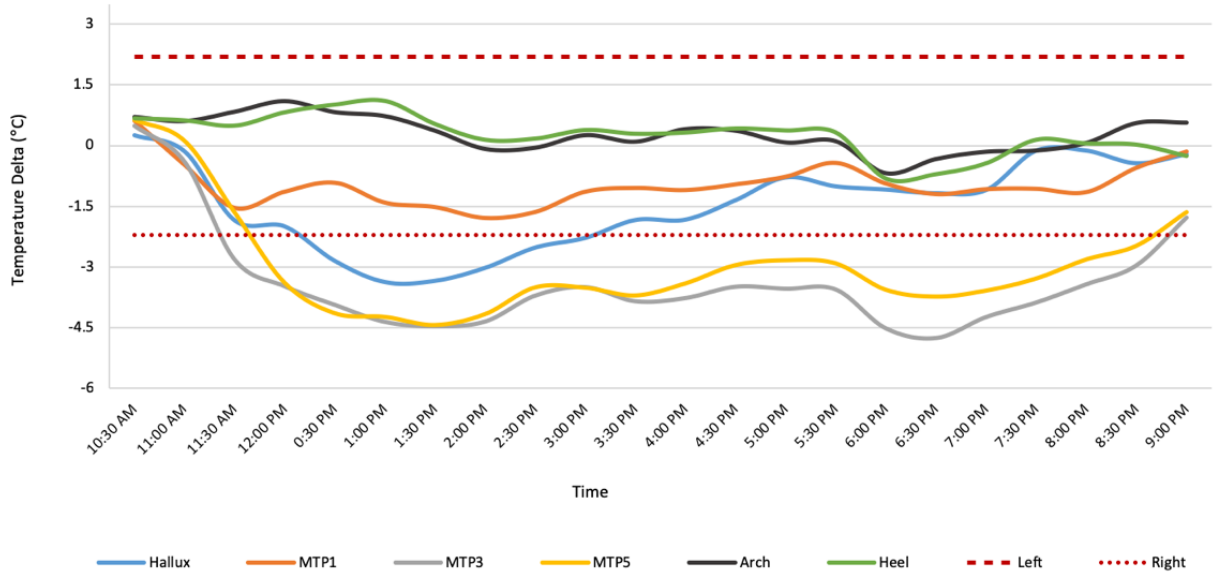


Figure 11. Each line on the graph shows a moving average of the temperature difference (ie, left foot temperature-right foot temperature) for the hallux (blue), metatarsal points 1, 3, and 5 (orange, gray, and yellow, respectively), arch (black), and heel (green). The moving average of the temperature difference shows elevated temperatures at metatarsal points 3 and 5 of the right foot. MTP: metatarsal point.



Discussion

Principal Findings

To our knowledge, this is the first study to introduce wireless continuous temperature monitoring of feet for daily and home use in patients with diabetes. The sensor-embedded socks introduced here contain microsensors embedded directly and seamlessly into the fabric and are designed to look and feel like any other garment. Particularly important for daily and home use, the socks are wireless. Wireless transmission of data is achieved via Bluetooth technology. Through the mobile app, wearers can view the current temperature as measured at six points on the user's foot. While the app was not set up to generate alerts in this study, users can receive a notification, alert, or text message when a temperature increase is detected between contralateral positions.

The aim of this study was to assess whether these sensor-embedded socks can measure temperature accurately on a continuous basis, whether the temperature findings are consistent with clinical observations, and to obtain feedback on

patient experience in using the socks. The temperature studies presented here show that the sensors used in the socks are reliable and accurate at detecting temperature.

In this pilot study of 35 patients, participants found the socks to be no different from standard socks in terms of wearability and reported feelings of comfort and safety. Patients found the app interface to be useful. Furthermore, as shown in the illustrative cases, the temperature differences between the two feet as recorded by the sensor-embedded socks were consistent with the clinical status of the patient.

An earlier report described socks made of optical fiber designed for the clinic environment [28]. A more recent design of smart socks has woven-in sensors and measures temperature at 10-minute intervals, but contains wires [29]. The sensor-embedded socks described in this study are designed for daily wear, both inside and outside the home; are wireless; are machine washable; and do not need to be recharged. Unlike garments made of optical fiber, these socks are made of smart textile, which is designed to be made on standard industrial equipment and can be used anywhere without assistance.

For patients with diabetes and neuropathy, continuous temperature monitoring for the feet now offers information that was not previously available or possible, as was the case with the introduction of continuous glucose-monitoring technology for blood glucose levels.

Static or once-a-day measurements can present a risk of reporting false positives. With continuous monitoring, algorithms can be designed to identify and filter out outliers in measurements spanning several hours and, thus, can potentially reduce false positives by taking into consideration trends over time instead of a single static threshold. As shown in Case Study 2, temperature measurements of the entire foot may be particularly beneficial to patients with Charcot arthropathy. For patients undergoing treatment for an existing injury as in Case Study 3, continuous temperature monitoring provides an objective method to identify injuries.

Patterns of temperature can be obtained via continuous temperature monitoring that are specific to individuals and, in the future, variations from a person's typical pattern may trigger alerts, rather than a single one-size-fits-all temperature threshold for all individuals. Monitoring patients over time may reveal temporal changes in individual temperature patterns. With further research, algorithms can be developed to detect temperature differences within one foot, without the need for the contralateral foot. Advanced statistical pattern recognition analysis could be used to determine patterns indicative of diabetes-related foot complications.

In future iterations, sensor-embedded socks can be coupled with built-in activity tracking to improve adherence and monitor patient compliance: data from the socks can be used to monitor patient activity and determine whether the patient is compliant with set activity guidelines.

This unique new data stream opens up questions regarding the manner in which the results are best reported, on the content and frequency of notifications, whether preulcerative lesions can be prevented from developing into ulcers, and whether amputations can be reduced. The pilot study reported here was not statistically powered to assess the performance characteristics of this novel device.

Further studies are planned to address such questions, with patient follow-up to obtain data on correlations of the temperature findings with patient outcomes.

Strengths and Limitations

The strengths of the study are as follows. The sensor-embedded socks were found to work reliably and consistently. The temperature differences reported matched clinical observations. Importantly, the study confirmed that patients can use the socks as a part of their daily lives, within or outside the home. Furthermore, the automatic collection and analysis of the data remove the element of subjectivity from the measurements as currently exists in visual inspection [30].

This study was not without limitations. As it was a single-day study, the findings could not be correlated with longer-term outcomes. More research is needed to further understand data points in continuous temperature monitoring, including as it relates to patient activity and timely intervention. Socks with built-in activity tracking and monitoring are planned to reliably and accurately measure activity concurrent with temperature measurement to further reduce subjective reporting. Future studies will be statistically powered to collect and analyze temperatures and correlate the findings to patient outcomes.

Conclusions

In this study, we explored the first use of wireless continuous temperature monitoring for daily and home use in patients with diabetes and neuropathy. This noninvasive device designed to behave as a normal sock is the first of its kind to combine wireless continuous temperature monitoring into a wearable device. The socks appear to the wearers to be no different than standard socks. When used with the mobile app, the wearer is kept informed about temperature increases in one foot relative to the other. The socks can reliably and consistently collect temperature data from the wearer's feet, which are consistent with clinical observations. Continuous temperature monitoring has emerged as a promising tool which could serve as an early warning system for the management of foot ulcers, Charcot foot, and reulceration.

Acknowledgments

We wish to thank Gayana Sarkisova for assistance with clinic logistics; Jaap J Van Netten, Edgar Peters, Klaus Kirketerp-Møller, Dan Burnett, Shawn Becker, and Yingmo Fang for helpful discussions; and the patients who participated in this study.

Authors' Contributions

AMR, HJS, and RM designed the study. AMR, KK, XS, and MM executed the study. EY designed the sensor-embedded socks used in the study. RM, MM, JF, and RP performed data analysis. RP prepared the manuscript. All authors reviewed and approved the manuscript.

Conflicts of Interest

RM, HJS, XS, EY, JF, and MM are employees and shareholders of Siren. RP is a consultant to Siren. AMR is an advisor to Siren and shareholder of Siren. The study was sponsored by Siren.

References

1. Centers for Disease Control and Prevention. National Diabetes Statistics Report 2017: Estimates of Diabetes and Its Burden in the United States. Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services; 2017. URL: https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf [accessed 2018-11-30] [WebCite Cache ID 74KJk4F3s]
2. Armstrong DG, Boulton AJM, Bus SA. Diabetic foot ulcers and their recurrence. N Engl J Med 2017 Dec 15;376(24):2367-2375. [doi: 10.1056/NEJMra1615439] [Medline: 28614678]
3. Singh N, Armstrong DG, Lipsky BA. Preventing foot ulcers in patients with diabetes. JAMA 2005 Jan 12;293(2):217-228. [doi: 10.1001/jama.293.2.217] [Medline: 15644549]
4. Armstrong DG, Lavery LA, Harkless LB, Van Houtum WH. Amputation and reamputation of the diabetic foot. J Am Podiatr Med Assoc 1997 Jun;87(6):255-259. [doi: 10.7547/87507315-87-6-255] [Medline: 9198345]
5. Lazzarini PA, Clark D, Derhy PH. What are the major causes of lower limb amputations in a major Australian teaching hospital? The Queensland Diabetic Foot Innovation Project, 2006-2007. J Foot Ankle Res 2011;4(Suppl 1):O24. [doi: 10.1186/1757-1146-4-S1-O24]
6. Barshes NR, Sigireddi M, Wrobel JS, Mahankali A, Robbins JM, Kougias P, et al. The system of care for the diabetic foot: Objectives, outcomes, and opportunities. Diabet Foot Ankle 2013 Oct 10;4:1-12 [FREE Full text] [doi: 10.3402/dfa.v4i0.21847] [Medline: 24130936]
7. Lazzarini PA, Pacella RE, Armstrong DG, van Netten JJ. Diabetes-related lower-extremity complications are a leading cause of the global burden of disability. Diabet Med 2018 May 23:1-6. [doi: 10.1111/dme.13680] [Medline: 29791033]
8. Boulton AJM, Vileikyte L, Ragnarson-Tennvall G, Apelqvist J. The global burden of diabetic foot disease. Lancet 2005 Nov 12;366(9498):1719-1724. [doi: 10.1016/S0140-6736(05)67698-2] [Medline: 16291066]
9. Lavery LA, Peters EJG, Armstrong DG. What are the most effective interventions in preventing diabetic foot ulcers? Int Wound J 2008 Jun;5(3):425-433. [doi: 10.1111/j.1742-481X.2007.00378.x] [Medline: 18593392]
10. Chantelau E. The perils of procrastination: Effects of early vs delayed detection and treatment of incipient Charcot fracture. Diabet Med 2005 Dec;22(12):1707-1712. [doi: 10.1111/j.1464-5491.2005.01677.x] [Medline: 16401316]
11. Margolis DJ, Malay DS, Hoffstad OJ, Leonard CE, MaCurdy T, Tan Y, et al. Economic burden of diabetic foot ulcers and amputations: Data Points #3. In: Data Points Publication Series. Rockville, MD: Agency for Healthcare Research and Quality; Mar 08, 2011.
12. Skrepnek GH, Mills JL, Lavery LA, Armstrong DG. Health care service and outcomes among an estimated 6.7 million ambulatory care diabetic foot cases in the US. Diabetes Care 2017 Dec;40(7):936-942. [doi: 10.2337/dc16-2189] [Medline: 28495903]
13. Skrepnek GH, Mills JL, Armstrong DG. A diabetic emergency one million feet long: Disparities and burdens of illness among diabetic foot ulcer cases within emergency departments in the United States, 2006-2010. PLoS One 2015;10(8):e0134914 [FREE Full text] [doi: 10.1371/journal.pone.0134914] [Medline: 26248037]
14. Peters EJG, Lavery LA, Armstrong DG. Diabetic lower extremity infection: Influence of physical, psychological, and social factors. J Diabetes Complications 2005;19(2):107-112. [doi: 10.1016/j.jdiacomp.2004.06.002] [Medline: 15745841]
15. Frykberg RG, Gordon IL, Reyzelman AM, Cazzell SM, Fitzgerald RH, Rothenberg GM, et al. Feasibility and efficacy of a smart mat technology to predict development of diabetic plantar ulcers. Diabetes Care 2017 Dec;40(7):973-980. [doi: 10.2337/dc16-2294] [Medline: 28465454]
16. Lavery LA, Higgins KR, Lanctot DR, Constantinides GP, Zamorano RG, Athanasiou KA, et al. Preventing diabetic foot ulcer recurrence in high-risk patients: Use of temperature monitoring as a self-assessment tool. Diabetes Care 2007 Jan;30(1):14-20. [doi: 10.2337/dc06-1600] [Medline: 17192326]
17. van Netten JJ, van Baal JG, Liu C, van der Heijden F, Bus SA. Infrared thermal imaging for automated detection of diabetic foot complications. J Diabetes Sci Technol 2013 Sep 01;7(5):1122-1129 [FREE Full text] [doi: 10.1177/193229681300700504] [Medline: 24124937]
18. Armstrong DG, Lavery LA. Monitoring healing of acute Charcot's arthropathy with infrared dermal thermometry. J Rehabil Res Dev 1997 Jul;34(3):317-321. [Medline: 9239625]

19. Najafi B, Wrobel JS, Grewal G, Menzies RA, Talal TK, Zirie M, et al. Plantar temperature response to walking in diabetes with and without acute Charcot: The Charcot Activity Response Test. J Aging Res 2012;2012:140968 [FREE Full text] [doi: 10.1155/2012/140968] [Medline: 22900177]
20. van Netten JJ, Prijs M, van Baal JG, Liu C, van der Heijden F, Bus SA. Diagnostic values for skin temperature assessment to detect diabetes-related foot complications. Diabetes Technol Ther 2014 Nov;16(11):714-721. [doi: 10.1089/dia.2014.0052] [Medline: 25098361]
21. Hazenberg CEVB, van Netten JJ, van Baal SG, Bus SA. Assessment of signs of foot infection in diabetes patients using photographic foot imaging and infrared thermography. Diabetes Technol Ther 2014 Jun;16(6):370-377. [doi: 10.1089/dia.2013.0251] [Medline: 24690146]
22. Bus SA, van Netten JJ, Lavery LA, Monteiro-Soares M, Rasmussen A, Jubiz Y, et al. IWGDF Guidance on the Prevention of Foot Ulcers in At-Risk Patients With Diabetes. The Netherlands: International Working Group on the Diabetic Foot; 2015. URL: http://www.iwgdf.org/files/2015/website_prevention.pdf [accessed 2018-11-30] [WebCite Cache ID 74KKHXUg7]
23. Liu C, van der Heijden F, Klein M, van Baal JG, Bus S, van Netten JJ. Infrared dermal thermography on diabetic feet soles to predict ulcerations: A case study. In: Proceedings of Advanced Biomedical and Clinical Diagnostic Systems XI 2013. Bellingham, WA: SPIE; 2013 Mar 22 Presented at: Advanced Biomedical and Clinical Diagnostic Systems XI 2013; February 3-7, 2013; San Francisco, CA. [doi: 10.1117/12.2001807]
24. Liu C, van Netten JJ, van Baal JG, Bus SA, van der Heijden F. Automatic detection of diabetic foot complications with infrared thermography by asymmetric analysis. J Biomed Opt 2015 Feb;20(2):26003. [doi: 10.1117/1.JBO.20.2.026003] [Medline: 25671671]
25. Aan de Stegge WB, Mejaiti N, van Netten JJ, Dijkgraaf MGW, van Baal JG, Busch-Westbroek TE, et al. The cost-effectiveness and cost-utility of at-home infrared temperature monitoring in reducing the incidence of foot ulcer recurrence in patients with diabetes (DIATEMP): Study protocol for a randomized controlled trial. Trials 2018 Sep 24;19(1):520 [FREE Full text] [doi: 10.1186/s13063-018-2890-2] [Medline: 30249296]
26. Basatneh R, Najafi B, Armstrong DG. Health sensors, smart home devices, and the Internet of Medical Things: An opportunity for dramatic improvement in care for the lower extremity complications of diabetes. J Diabetes Sci Technol 2018 May;12(3):577-586. [doi: 10.1177/1932296818768618] [Medline: 29635931]
27. Killeen AL, Walters JL. Remote temperature monitoring in diabetic foot ulcer detection. Wounds 2018 Apr;30(4):E44-E48 [FREE Full text] [Medline: 29718822]
28. Najafi B, Mohseni H, Grewal GS, Talal TK, Menzies RA, Armstrong DG. An optical-fiber-based smart textile (Smart Socks) to manage biomechanical risk factors associated with diabetic foot amputation. J Diabetes Sci Technol 2017 Dec;11(4):668-677 [FREE Full text] [doi: 10.1177/1932296817709022] [Medline: 28513212]
29. El-Nahas M, El-Shazly S, El-Gamel F, Motawea M, Kyrillos F, Idrees H. Relationship between skin temperature monitoring with Smart Socks and plantar pressure distribution: A pilot study. J Wound Care 2018 Aug 02;27(8):536-541. [doi: 10.12968/jowc.2018.27.8.536] [Medline: 30086248]
30. Bus SA. Innovations in plantar pressure and foot temperature measurements in diabetes. Diabetes Metab Res Rev 2016 Jan;32 Suppl 1:221-226. [doi: 10.1002/dmrr.2760] [Medline: 26467347]

Abbreviations

DFU: diabetic foot ulcer
DPN: diabetic peripheral neuropathy
IQR: interquartile range
MTP: metatarsal point

Edited by G Eysenbach; submitted 09.10.18; peer-reviewed by M Dini, B Arnoldussen; comments to author 26.10.18; revised version received 01.11.18; accepted 01.11.18; published 17.12.18.

Please cite as:
Reyzelman AM, Koelewyn K, Murphy M, Shen X, Yu E, Pillai R, Fu J, Scholten HJ, Ma R
Continuous Temperature-Monitoring Socks for Home Use in Patients With Diabetes: Observational Study
J Med Internet Res 2018;20(12):e12460
URL: http://www.jmir.org/2018/12/e12460/
doi:10.2196/12460
PMID:30559091

Review

Decision Support Tools for Regenerative Medicine: Systematic Review

Ching Lam[1,2], MEng; Edward Meinert[2,3], MA, MSc, MBA, MPA, CEng FBCS, FCMI; Abrar Alturkistani[3], BSc, MPH; Alison R Carter[2], BA, PhD, FRSA; Jeffrey Karp[4], BEng, PhD; Aidong Yang[1], BSc, PhD; David Brindley[2], MEng, DPhil, RAC (US), PGCert IP, FRSA, FRSB; Zhanfeng Cui[1], BSc, MSc, MA, PhD, DSc, CEng CSci

[1]Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, United Kingdom

[2]Healthcare Translation Research Group, Department of Paediatrics, University of Oxford, Oxford, United Kingdom

[3]Global Digital Health Unit, Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

[4]Division of Engineering in Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
Edward Meinert, MA, MSc, MBA, MPA, CEng FBCS, FCMI
Healthcare Translation Research Group
Department of Paediatrics
University of Oxford
John Radcliffe Hospital
Oxford,
United Kingdom
Phone: 44 7824446808
Email: edward.meinert@paediatrics.ox.ac.uk

Abstract

Background: Decisional tools have demonstrated their importance in informing manufacturing and commercial decisions in the monoclonal antibody domain. Recent approved therapies in regenerative medicine have shown great clinical benefits to patients.

Objective: The objective of this review was to investigate what decisional tools are available and what issues and gaps have been raised for their use in regenerative medicine.

Methods: We systematically searched MEDLINE to identify articles on decision support tools relevant to tissue engineering, and cell and gene therapy, with the aim of identifying gaps for future decisional tool development. We included published studies in English including a description of decisional tools in regenerative medicines. We extracted data using a predesigned Excel table and assessed the data both quantitatively and qualitatively.

Results: We identified 9 articles addressing key decisions in manufacturing and product development challenges in cell therapies. The decision objectives, parameters, assumptions, and solution methods were analyzed in detail. We found that all decisional tools focused on cell therapies, and 6 of the 9 reviews focused on allogeneic cell therapy products. We identified no available tools on tissue-engineering and gene therapy products. These studies addressed key decisions in manufacturing and product development challenges in cell therapies, such as choice of technology, through modeling.

Conclusions: Our review identified a limited number of decisional tools. While the monoclonal antibodies and biologics decisional tool domain has been well developed and has shown great importance in driving more cost-effective manufacturing processes and better investment decisions, there is a lot to be learned in the regenerative medicine domain. There is ample space for expansion, especially with regard to autologous cell therapies, tissue engineering, and gene therapies. To consider the problem more comprehensively, the full needle-to-needle process should be modeled and evaluated.

(J Med Internet Res 2018;20(12):e12448) doi:10.2196/12448

KEYWORDS

decisional tool; systematic review; regenerative medicine; cell therapy; decision support techniques; cell- and tissue-based therapy

Introduction

Rationale

Decisional tools or decision support tools are tools that can be used to support complex decision making and problem solving. Since their advent in the 1970s [1], these tools have been used to support evidence-based decision making in various industries, including health care [2], agriculture [3], and the environment [4].

In the biopharmaceutical industry, decisional tools have been applied to monoclonal antibody and vaccine manufacturing decisions for over 20 years. These tools have proved to be useful for understanding cost structures and risks in order to inform decisions in various areas, including technology evaluation, facility fit, and capacity planning [5-10].

Decision support tools such as cost-of-goods modeling have proved themselves instrumental in informing the industry about the economic drivers in switching to new technologies. One such example is the shift from stainless steel to single-use production strategies for biologics over the last 15 years across the biopharma industry, which allowed faster campaign turnover, lower initial capital costs, and manufacturing cost savings [7,9,11]. Through providing a better understanding of the cost drivers for change, decisional tools were able to help build a valid commercial case to influence decision makers in making important business and bioprocess decisions, from technology choice and process change, to supply chain and project portfolio management [6,12-15].

Regenerative medicine, as defined by Mason and Dunnill in 2008, “replaces or regenerates human cells, tissues or organs, to restore or establish normal function, with approaches such as use of soluble molecules, gene therapy, stem cell transplantation, gene therapy, tissue engineering and the reprogramming of cell and tissue types” [16]. By 2018, the field had seen major breakthroughs. With US Food and Drug Administration approvals of the genetically modified T-cell therapies tisagenlecleucel (Kymriah) and axicabtagene ciloleucel (Yescarta) for refractory or relapsed acute lymphoblastic leukemia and large B-cell lymphoma, respectively, and voretigene neparvovec (Luxturna) for retinal dystrophy [17], the industry is slowly living up to its expectations. As more regenerative medicine products are commercialized, decisions such as cost-of-goods optimization and process design become more critical.

Objectives

Rekhi et al [18] reviewed the existing decisional tools for monoclonal antibodies and cell therapy bioprocessing and identified plenty of room for expansion. With this review, we aimed to provide a systematic update of the regenerative medicine decision support tool landscape, with a focus on tissue engineering, and cell and gene therapies, to identify the gaps in the literature and inform future development of decisional tools in the area.

The key research questions this review aimed to address were as follows. First, what decisional tools are available in the regenerative medicine domain? Second, what issues have been addressed? Third, what are the gaps in decisional tools for regenerative medicine?

Methods

We conducted this systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. As we used only publicly available information, the review did not require ethics review board approval.

Eligibility Criteria

To be included in this review, each article had to meet the following criteria: (1) addressed regenerative medicines, either autologous or allogeneic, (2) described a decisional tool, and (3) was available in English.

We enforced the following exclusion criteria to allow the review to focus on the outcomes of fresh research reported with sufficient details: (1) review articles, (2) conference abstracts, and (3) book chapters.

Search Strategy and Study Selection

We searched the literature on May 23, 2018 to identify suitable studies indexed in the MEDLINE database and electronically identified the bibliographical references of the articles. We also manually searched Google Scholar.

To find relevant studies, we used the keywords in Textbox 1.

We constructed a search string by pairing a regenerative medicine term and a decisional tool search term—for example: (regenerative medicine) AND bioprocess economics; (cell therapy) AND bioprocess design.

We screened all titles and abstracts that we identified for relevance. Subsequently, we obtained full-text articles and reviewed eligible articles.

Data Collection

We analyzed the relevant articles in five aspects that are typically shared by decision support tools found in the literature (Figure 1). We extracted key data (Figure 1) from each source, by following the same structure, into a predesigned Excel spreadsheet (Microsoft Corporation).

Textbox 1. Search terms.

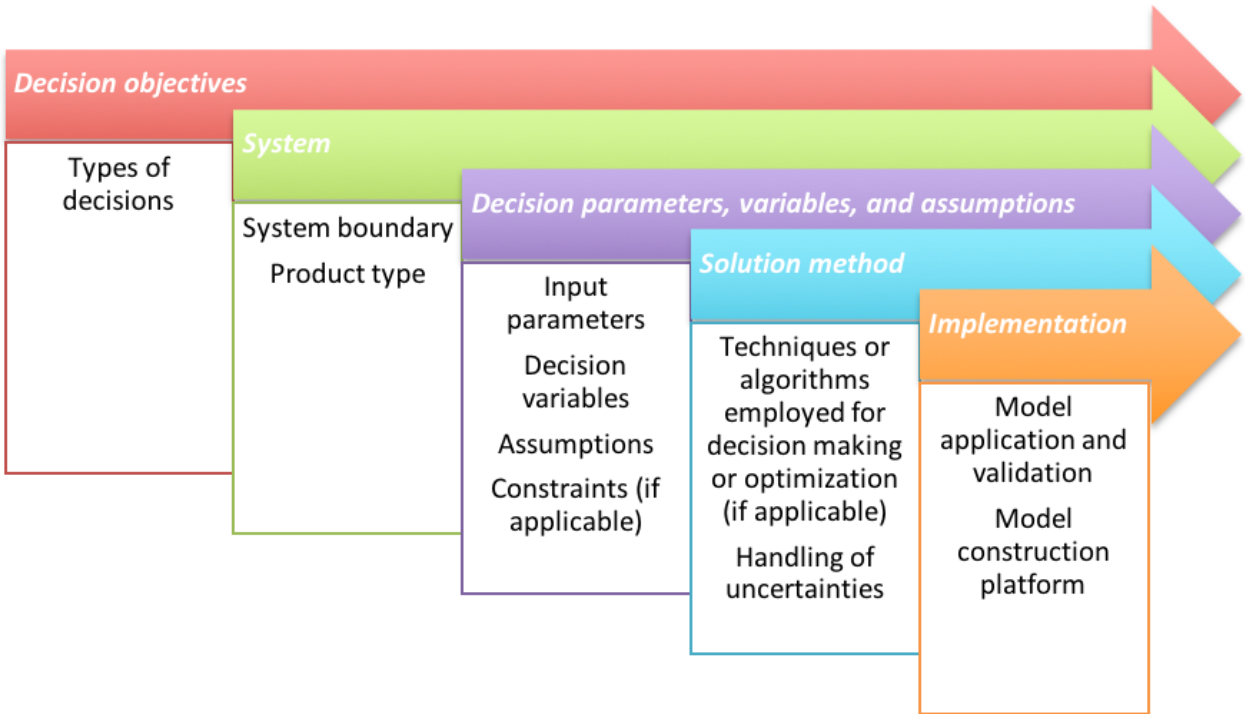
Regenerative medicine search items

- Regenerative medicine
- Cell therapy
- Tissue engineering
- Gene therapy
- Exosomes

Decisional tools search terms

- Bioprocess economics
- Bioprocess design
- Decision* tool
- Evaluation framework

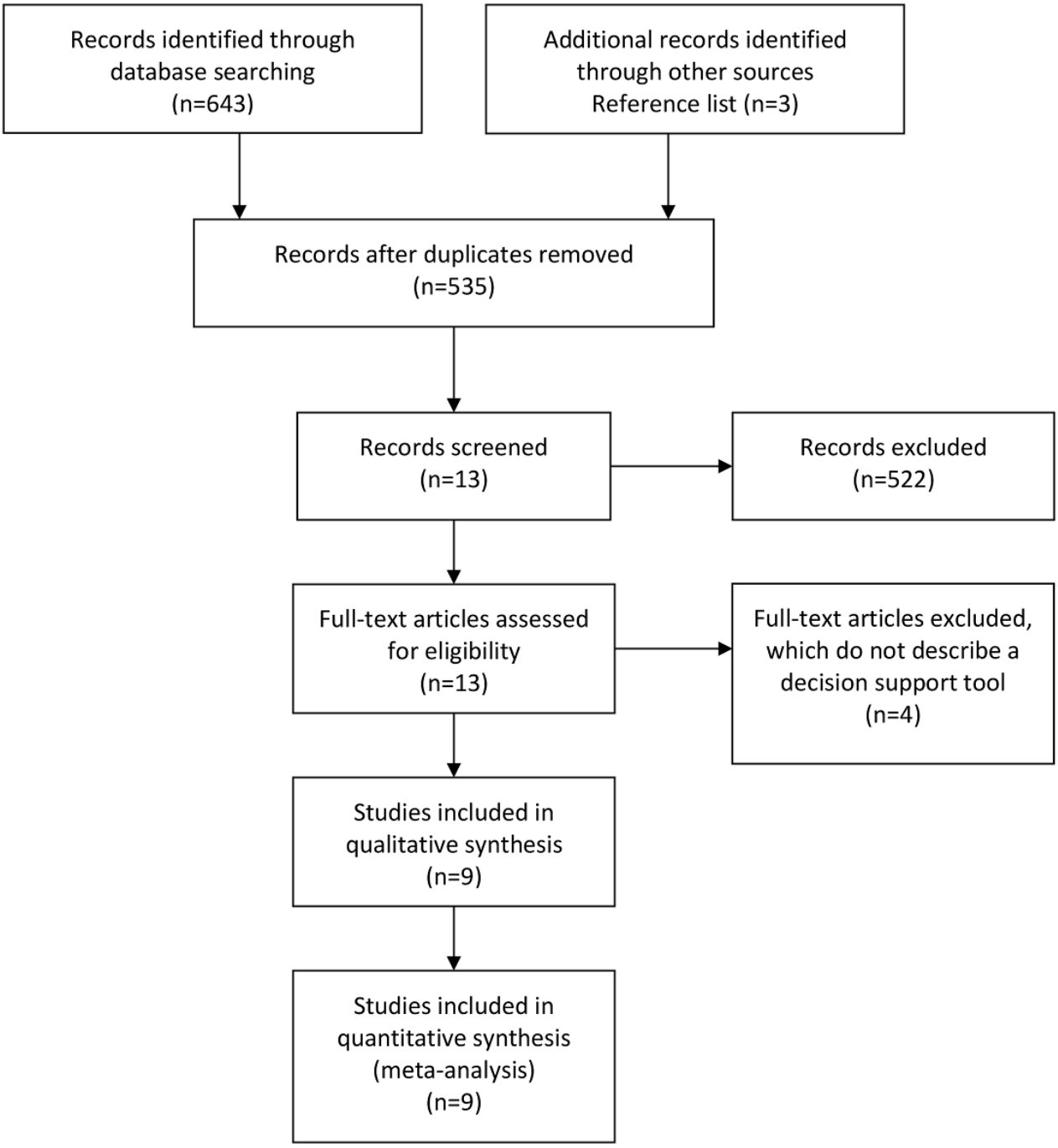
Figure 1. Key data extracted from the eligible literature.



Results

Figure 2 shows the PRISMA flowchart of the literature search process. The database searches identified 646 articles for review. At the screening stage, we deemed 13 articles to be relevant for full-text eligibility assessment. We excluded 4 full-text articles after screening, as they did not contain a description of a decisional tool and, hence, did not meet the eligibility criteria. Thus, we identified 9 articles that met the inclusion criteria and reviewed them in full detail for subsequent assessments. Multimedia Appendix 1 shows the executed data abstraction form. The small number of articles we initially identified and the small number resulting after screening indicate the novelty of this research area.

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses(PRISMA) flowchart of the literature review process.



Decision Objectives

Table 1 [19-27] shows the various decision objectives of published decisional models. Ungrin et al aimed to enhance the upstream cell expansion yield [19]. Lambrechts et al [20] described a visualization tool for upstream expansion processes cited from other literature, and hence did not have a decision objective. All other models focused on optimization of the costs for manufacturing or product development. Manufacturing cost of goods was further broken down to show subcategories such as raw material, labor, consumables, and capital equipment in various analyses. Product development costs relate to investment required to bring the product from bench to bedside, including particularly clinical trial costs. Optimizing these costs is critical in the sustainable development of companies and their operational efficiency. Project net present value (NPV) is a commonly used method in project evaluation [28]. Through evaluating the NPV as an impact of process change in the development timeline, Hassan et al [21] reflected the risks and benefits of making a process change from one technology to another.

Table 1. Decision objectives of various decisional models in the reviewed articles.

Decision objectives	Articles
Operational yield for cell expansion process	Ungrin, 2012 [19]
Cost of goods	Upstream: Simaria, 2014 [22] Downstream: Hassan, 2015 [23] Overall: Weil, 2017 [24]; Harrison, 2018 [25]; Jenkins, 2018 [26]
Investment costs	McCall, 2013 [27]; Hassan, 2016 [21]
Risk-adjusted net present value	Hassan, 2016 [21]
Not applicable	Lambrechts, 2016 [20]

System

System Boundary

Parnell et al [29] define a system boundary as a physical or conceptual boundary that contains all the essential elements, subsystems, and interactions necessary to address a systems decision problem. Different decision objectives would motivate different definitions of systems boundaries.

We generalized the systems in the eligible articles into two types: (1) product development systems, and (2) manufacturing and supply chain systems.

Product Development Systems

Two identified articles looked into the development phase of cell therapies, both describing generic processes that can be applied to any cell therapies. McCall [27] defined the systems boundary as being between preclinical trials and phase 3 clinical trials in order to look into the costs of developing a cell therapy. Hassan et al [21] defined their systems boundaries as being between phase 1 clinical trials and regulatory approval in order to study the impact of process changes along the development phases on NPV of their project.

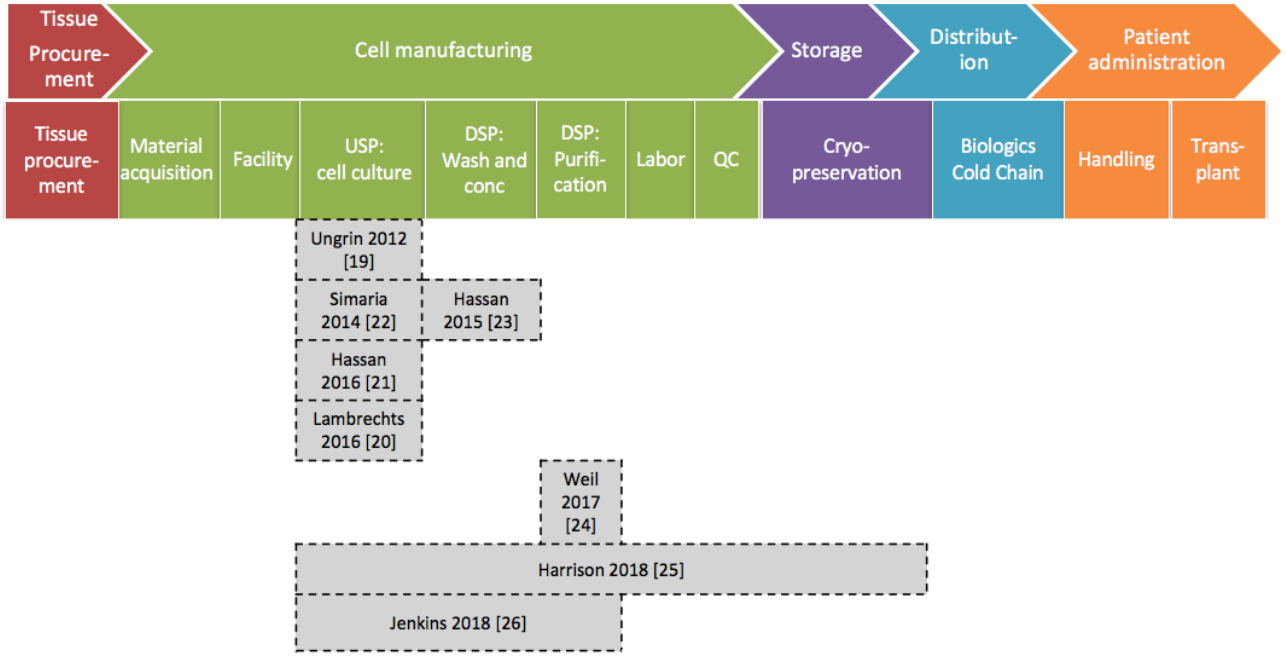
Manufacturing and Supply Chain Systems

A total of 8 articles investigated decision making for manufacturing and supply chain systems. Figure 3 shows the system boundaries in these 8 articles mapped against the needle-to-needle or patient-to-patient (ie, from patient tissue procurement to therapy administration) cost-of-goods roadmap proposed by Lipsitz et al [30].

Ungrin et al [19] and Lambrechts et al [20] addressed optimization of the cell expansion process upstream through experiments, bioprocess modeling, and visualization. Hassan et al [21] focused on process change impacts along the product development pathway using the change of upstream processing technology as an illustrative example; hence we have included their study in this section as well.

Simaria et al [22], Hassan et al [23], Weil et al [24], Harrison et al [25], and Jenkins and Farid [26] evaluated different technology options for the studied steps within their defined system boundaries to better understand the advantages, disadvantages, and bottlenecks in adopting different technology options and their implications for manufacturing cost of goods.

Figure 3. Coverage of existing decisional tools. conc: concentration; DSP: downstream processing; QC: quality control; USP: upstream processing.



Product Type

All the articles we identified focused on cell and ex vivo gene therapy; we found no decisional tools in tissue engineering or in vivo gene therapy. Table 2 [19-27] shows the product types and type of transplant considered in the articles.

There was considerably more focus on allogeneic therapies and mesenchymal stem cells (MSCs) as the cell source. The study by Weil et al [24] is the only one in this review that considered autologous processes, but it focused only on downstream affinity purification.

Decision Variables, Parameters, and Assumptions

Decision variables conceptually or mathematically represent decisions to be made in order to (best) achieve the decision objectives. For a given decision objective, the determination of a decision variable is typically affected by either internal or external characteristics, or by both of them, which are referred to as decision parameters. In other words, decision parameters, together with other assumptions taken for internal or external settings, form the input to a decision process, while the determined values for the decision variables are its output.

Decision Variables

Decision variables are variables that the decision-maker controls. Such variables are dependent on the decision objectives and the problems they seek to answer.

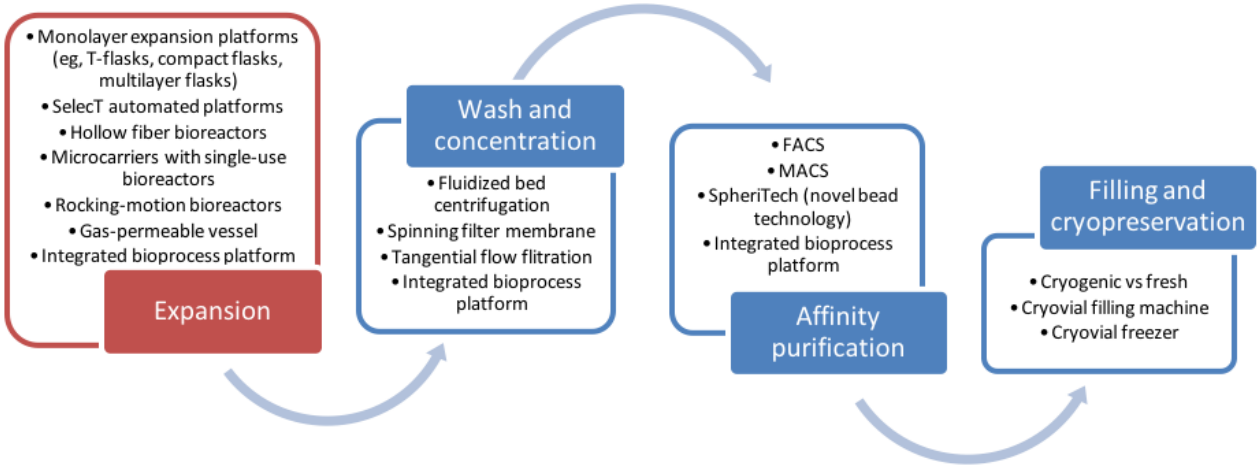
For product development systems, the models seek to answer to objectives such as minimizing development duration, risks, and investment costs. The timing for technology change was the decision variable modeled by Hassan et al to study the impact of process change in product development on the project NPV [21]. McCall [27] studied the impacts of product development risks and uncertainties, as well as rework probability on the investment costs; as there was no optimization module in this study, no decision variable was identified.

Table 2. Cell types and type of transplant.

Transplant type	Cell type				
	Mesenchymal stem cells	Chimeric antigen receptor T-cell	Human pluripotent stem cell/induced pluripotent stem cells	Not specified	Not applicable
Allogeneic	Hassan, 2015 [23], Hassan, 2016 [21], Lambrechts, 2016 [20], Harrison, 2018 [25]	Jenkins, 2018 [26]	N/A[a]	Simaria, 2014 [22]	N/A
Autologous	N/A	N/A	Weil, 2017 [24]	N/A	N/A
Not specified	N/A	N/A	Ungrin, 2012 [19]	N/A	N/A
Not applicable	N/A	N/A	N/A	N/A	McCall, 2013 [27]

[a]N/A: not applicable.

Figure 4. Upstream and downstream operations considered in the reviewed articles. FACS: fluorescence-activated cell sorting; MACS: magnetic-activated cell sorting.



For manufacturing and supply chain systems, the models seek to answer to objectives such as process yield [19] and manufacturing cost minimization [22,24-26]. In these systems, the choice of unit operations in the manufacturing process sequence, chosen equipment, and their capacities are critical decision variables to be considered. Figure 4 shows the upstream and downstream unit operations options considered in the articles. Together with process parameters, process flowsheet, and requirements, the processes were modeled in order to understand the associated cost of goods. Conclusions on the relative advantages of various technology options under different demand scenarios were drawn for bottleneck analysis and decision recommendations.

Input Parameters

Input parameters are defined as "a constant element or factor, especially serving as a limit or boundary" [31]. These are inputs defined as prerequisites of the objective function; in other words, they are constants of the objective function and not the values to be optimized. In the 9 articles we identified, these included scale, throughput, demand goal, and technical process parameters.

Scale, Throughput, and Demand Goal

Early articles in the area generally looked at demands several times more than the recent articles. As more commercial case studies of regenerative medicine products arise, the demand landscape is better understood and the estimation for cost-of-goods modeling has been lowered from the monoclonal antibodies ballpark (1000-500,000 doses of allogeneic MSCs per year [22] to around 2500 doses/y for a regional center for allogeneic MSCs [25] and 500-5000 doses/y for chimeric antigen receptor T cells (CAR-T) [26]). Demand scenarios were chosen depending on the cell type and the therapeutic target. As more and more real-world commercial cases have emerged, recent articles gave a lot more consideration to the clinical applications and their specific demands and, hence, proposed more realistic demand scenarios.

Upstream and Downstream Operations Process Parameters

For upstream operations, comparisons were drawn from a range of equipment scale, cell culture modes, and extent of automation. Table 3 [19,22,25,26] summarizes process parameters previously considered and explicitly mentioned in their respective articles.

In all the articles, planar culture flasks (eg, T-flasks and multilayer flasks) were consistently found to be the most expensive of all evaluated technologies for allogeneic therapies and infeasible for higher cell number per lot. The number of cells harvested per surface area was found to be the most important cost driver, as it dictates the number of expansion units required and, hence, the raw materials and consumables requirements [22,26].

Table 4 [23,24,26] shows the process parameters for downstream processing discussed in the articles.

Table 3. Input process parameters for upstream processing.

Input process parameters	Ungrin, 2012 [19]	Simaria, 2014 [22]	Harrison, 2018 [25]	Jenkins, 2018 [26]
Studied technologies	Microwell	T-flasks, multilayers, compact flasks, compact multilayers, multilayer bioreactors, hollow fiber bioreactors	T-175 flasks, SelecT automated platform	Planar culture flasks, rocking-motion bioreactor, gas-permeable vessel, integrated bioprocess platform
Cell culture process parameters				
Population doublings	Yes	No	No	No
Inoculation cell count	No	No	Yes	No
Seeding density	No	Yes	Yes	No
Harvest density	No	Yes	No	Yes
Technology process parameters				
Surface area	Yes	Yes	No	No
Equipment size and volume range	No	Yes	Yes	Yes
Number of expansion stages	No	Yes	No	No
Perfusion rate	No	No	No	Yes
Maximum units	No	Yes	No	No
Biosafety cabinet requirement	No	Yes	No	No
Incubator capacity requirement	No	Yes	Yes	No
Time duration assumptions				
Seed time	No	Yes	No	No
Feed time	No	Yes	No	No
Harvest time	No	Yes	No	No
Cell culture duration	No	No	Yes	No
Material use and cost assumptions				
Media requirements	No	Yes	Yes	Yes
Labor requirements	No	Yes	Yes	Yes
Consumable unit price	No	Yes	Yes	Yes
Capital charge	No	Yes	Yes	Yes

Table 4. Input process parameters for downstream processing.

Input process parameters	Hassan, 2015 [23]	Weil, 2017 [24]	Jenkins, 2018 [26]
Wash and concentration: studied technologies	Tangential flow filtration, fluidized bed centrifugation	N/Aa	Fluidized bed centrifugation, spinning filter membrane, integrated bioprocess platform
Purification: studied technologies	N/A	Fluorescence-activated cell sorting, magnetic-activated cell sorting, novel bead	Magnetic-activated cell sorting, integrated bioprocess platform
Technology process parameters			
Number of washes/cycles	Yes	Yes	No
Equipment size and volume range	Yes	Yes	Yes
Maximum cell processing capacity	Yes	Yes	Yes
Step yield	Yes	Yes	Yes
Time duration assumptions			
Maximum time	Yes	No	No
Material use and cost assumptions			
Raw material requirements	Yes	Yes	Yes
Labor requirements	Yes	Yes	Yes
Consumable unit price	Yes	Yes	Yes
Capital charge	Yes	Yes	Yes

[a]N/A: not applicable.

The downstream process starts with the wash and concentration step, and common wash concentration unit operations were discussed in detail by Hassan et al [23] and Jenkins and Farid [26]. Hassan et al reported that wash and concentration downstream steps were a bottleneck for high-cell-dose lots at high demand. As the demand estimate was lowered in the study of Jenkins and Farid, wash and concentration was no longer shown to be a bottleneck except in integrated bioprocess platforms such as CliniMACS Prodigy, which has a relatively smaller volume-reduction capacity.

Following wash and concentration, affinity purification has also been a target for modeling and optimization. Weil et al [24] and Jenkins and Farid [26] looked into affinity purification for autologous induced pluripotent stem cells and allogeneic CAR-T cells, respectively. Weil et al compared fluorescence-activated cell sorting (FACS) versus magnetic-activated cell sorting (MACS) and evaluated a novel technology that does not require cell labeling. MACS was determined to be more cost effective for dose sizes with a higher cell count (>7.0×10^7 cells/dose), as FACS is limited by its process scale. The model by Jenkins and Farid [26], for affinity purification, considered only MACS and integrated bioprocess platform.

Assumptions and Constraints

The articles made other assumptions besides technology-related assumptions and constraints.

For scheduling-related problems where one task follows another, task precedence constraint was used. McCall [27] used the task precedence constraint for dictating the start and end of a task, which is useful for setting up the scheduling problem. Iteration loops were built in the development pathway with assumption of learning. Hassan et al [21] assumed a linear project development pathway with failure probability. They constructed a database with information on clinical trial development times and failure rates of all 592 commercial cell therapy projects from 1981 to 2011 to estimate the duration and failure rate of similar products. This approach allows more industrially relevant benchmark assumptions to be made and, hence, gives rise to higher-quality results.

For resource utilization, McCall [27] assumed a fixed and steady-rate consumption of resources and a renewable resource pool throughout the project duration. Similar assumptions were made in all the other cost models to better understand the impact of resource utilization on the overall cost of goods. For instance, Simaria et al [22] showed that efficient use of equipment and facility can lower the depreciation costs shared across doses. Harrison et al [25] looked into the impact of human resource turnover in detail to understand the impact of increased operators on the relative cost of labor in overall cost of goods.

Having reasonable cost assumptions is one of the most important factors determining the validity of the model. Table 5 [22-26] shows some of the cost assumptions used in various models.

Table 5. Extracted quality control (QC), labor, and facility cost assumptions.

Cost type	Simaria, 2014 [22]	Hassan, 2015 [23]	Weil, 2017 [24]	Harrison, 2018 [25]	Jenkins, 2018 [26]
QC	US$10,000/lot	US$10,000/lot	QC costs per dose = £3250	QC costs based on Athersys Simpler assay panel: US$5934.53 Advanced assay panel: US$37512	QC quality assurance cost = 1 × operating labor cost
Labor	Operating labor = US$200/h, other labor cost multiplier = 0.2	Labor cost = hourly rate × number of operators × number of equipment × number of lots/y	Fluorescence-activated cell sorting operator wage = £57,500/y Magnetic-activated cell sorting operator wage = £46,000/y	Salary by salary band, taking into account pension, overheads, and training costs	Operator cost = US$120,000/y
Facility costs	Depreciation over 10 years	Lang factor: 23.67 Maintenance (% capital investment) = 10% Depreciation (% capital investment) = 7%	N/A[a]	Office space, business rates, service charge cleanroom space costed per square meter Depreciation over 10 years	N/A[a]

[a]N/A: not applicable.

The cost assumptions of Harrison et al [25] were considerably more detailed than those in the rest of the articles. Quality control test panels were in accordance with good manufacturing practice requirements of the specific product and are listed in detail in the supplementary material of Harrison et al and based on industry information from Athersys. Depending on product characteristics, each product needs different quality control tests and assays requirements and, hence, the costs can be quite different. For instance, genetically modified cells would require assays on transformed cell populations to demonstrate appropriate and reproducible expression of newly acquired characteristics [25].

Labor costs can be quite different depending on geographical location. In an interview with the chief executive of Nanjing Legend, a Chinese company, he estimated that the manufacturing costs for CAR-T in China can be one-sixth of those in the United States due to cheaper overheads [32]. Simaria et al [22] suggested in their sensitivity analysis that labor rate is one of the most important cost drivers for less-automated processes.

The two main methods for accounting for facility costs in the studies were equipment-factored estimates (eg, Lang factor) and estimates of cost per square meter. Facility costs are averaged over the period of depreciation and shared among all doses. The Lang factor is a commonly used method in project cost estimates in the engineering industry and is recommended by the American Association of Cost Engineers [33]. The Lang factor used by Hassan et al [23] was taken from Pollock et al [12], which took into account pipework and installation, process control, instrumentation, electrical power, building works, detail engineering, construction and site management, commissioning, and contingency factor. It was unclear what Harrison et al [25] included in their method of cost per square meter applied to cleanroom space. In addition, a different cleanroom grade would constitute a different cleanroom space running cost and, hence, it is important to understand the good manufacturing practice requirements of the manufacturing environment.

Solution Method

A solution method is required to relate the decision variables to the decision objective. Process models were built in all the identified studies. Table 6 [19-27] summarizes the approaches to solution methods used in the 9 articles. The two main approaches were process economics modeling in the form of what-if studies and multi-attribute decision making.

Table 6. Techniques and algorithms used for solution methods.

Technique or algorithm	Ungrin, 2012 [19]	McCall, 2013 [27]	Simaria, 2014 [22]	Hassan, 2015 [23]	Hassan, 2016 [21]	Lambrechts, 2016 [20]	Weil, 2017 [24]	Harrison, 2018 [25]	Jenkins, 2018 [26]
Process economics modeling	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Value systems modeling	No	Yes	No	No	No	No	No	No	No
Design structure matrix	No	Yes	No	No	No	No	No	No	No
What-if scenario analysis	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Multi-attribute decision making	No	No	No	No	No	No	No	No	Yes
Database evaluation	No	Yes	No	No	Yes	Yes	No	No	No
Latin hypercube	No	Yes	No	No	No	No	No	No	No
Monte Carlo simulation	No	No	No	No	Yes	No	No	No	Yes
Sensitivity analysis	No	No	Yes	No	No	No	Yes	No	Yes
Deterministic process evaluation	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes
Stochastic model	No	Yes	No	No	Yes	No	No	No	Yes
Data Visualization	No	No	No	No	No	Yes	No	No	No

Process Economics and Value Systems Modeling

To ensure all relevant costs are identified, typically models simulating the actual manufacturing or product development process are constructed and costs associated with each step are summed. Costs were analyzed by Simaria et al [22], Hassan et al [23], Hassan et al [21], Weil et al [24], Harrison et al [25], and Jenkins and Farid [26]. This method allows for process-centric costing, which in turn supports cost analyses based on technology options.

A value system modeling is a way of modeling the firms by sets of activities that the firms use to create value and competitive advantage [34]. McCall [27] modeled the set of activities in product development and accounted for development process characteristics such as interdependency, iteration, activity cost, and duration uncertainties. Through this model, McCall was able to highlight the critical processes, resources, and risks in product development. The report highlighted the importance of early-stage investment, clinical trials rework, and regulatory requirements.

Design Structure Matrix

Design structure matrix is a method developed by Steward and other in 1981 for planning and communicating engineering works [35]. The matrix represents the events, their sequence, and the interdependencies between events. McCall [27] used a design structure matrix to clearly represent the precedence constraints while considering iteration circuits inherent in product development.

What-If Scenario Analysis

Several articles included what-if scenario analysis, where the dose sizes, lot sizes, and demand for products were varied. These studies were used to provide guidance for technology selection under different circumstances.

Single Objective Versus Multi-Attribute Decision Making

While most articles dealt only with manufacturing or investment cost optimization, Jenkins and Farid's model employed a multicriteria decision-making methodology to assess bioprocess flowsheets [26].

The weighted sum technique provided a way to account for both quantitative and qualitative attributes of a solution, and, by assigning weightings, considered the perceived relative importance of different attributes. Weighted sum, however, is just one of many methods of multi-attribute decision making. Hester and Velasquez [36] conducted a comprehensive review and comparison of the methods commonly used. The analytic hierarchy process allows for pairwise comparisons to compare alternatives, which is less data intensive and more suitable for qualitative performance-type problems and resource management applications.

Handling of Risks and Uncertainties

Common themes incorporated into these manufacturing and development cost models are the risks and uncertainties lurking in the industry. The major methods of capturing risks and uncertainties in the studied models were stochastic modeling, Latin hypercube sampling, Monte Carlo analyses, and sensitivity analyses.

Deterministic Versus Stochastic Modeling

Deterministic models use discrete values, meaning that, for a certain input, the output will always be the same. Stochastic models have at least one quantity with random values, leading to an ensemble of different outputs [27].

Stochasticity was accounted for in 3 of the articles, where triangular distributions were applied to parameters to capture the uncertain and variable nature of the systems. McCall [27] applied a triangular probability distribution to the task duration to capture the uncertainties in development step duration. Hassan

et al [21] applied a probability distribution to the success rate of each development step. Jenkins and Farid [26] assigned probability distributions to the weighting of quality attributes in the multi-attribute decision-making module; hence, they modeled only the variability in preference of quality attribute.

We noted that probability distribution had not been assigned to variability and uncertainties in the manufacturing process in these studies.

Latin Hypercube Sampling and Monte Carlo Simulation

McCall [27] categorized the risks into product risk factors and enterprise risk factors. Product risks were defined as risks that can harm the patient, namely the choice of cell type, manufacturing processes, and delivery mechanism. Enterprise risks were defined as risks that affect the commercialization of the product and the business developing the product, namely technical risks and market risks. The Latin hypercube sampling method was used to consider the probability of failure and duration for each task along the product development pathway and the interdependencies. It is worthwhile to note that in this model, iteration caused by failures and impact of failures during each phase were considered using three matrices: design structure, rework probability, and rework impact.

Hassan et al [21] simulated the risks and uncertainties of process change along the product development pathway through Monte Carlo analyses. To adjust the project NPV according to risk, they used a discount rate based on the riskiness and expected development time.

Sensitivity Analysis

Sensitivity analysis is a common strategy to account for uncertainties and identify key cost drivers [22,24,26]. It is useful in understanding which assumptions and parameters the overall system is most sensitive to. Simaria et al [22] focused on upstream production for MSCs and found that the main cost drivers were microcarrier area, harvest density, media price, and downstream yield. Weil et al [24] and Jenkins and Farid [26], who considered processes requiring differentiation or gene modification, consistently cited the key cost drivers to be the efficiency of differentiation and gene modification. This is particularly interesting, as Weil et al modeled an autologous process [24] while Jenkins and Farid modeled an allogeneic process with a different cell type [26].

Hypothetical Case Studies

Hypothetical case studies where the demand and dose sizes varied were conducted for manufacturing systems, including by Simaria et al [22], Hassan et al [23], Weil et al [24], and Jenkins and Farid [26]. The case studies were useful in the evaluation of process bottlenecks and technology-switch sweet spot analysis.

Implementation

Model Validation

Data Mining

McCall [27] and Hassan et al [21] both established databases using real commercial case studies to inform some of their assumptions. McCall collected data from development programs surrounding orphan and non–orphan cell therapies [27], while Hassan et al collected information on clinical trial development times and failure rates of all 592 commercial cell therapy projects that entered development from 1981 to the end of 2011 [21]. These databases are useful for benchmarking purposes and increase the validity of the development time assumptions.

Laboratory Experiments

Ungrin et al [19], Weil et al [24], and Harrison et al [25] all included experimental results to inform assumptions in their studies. Using experimental results to support key assumptions is a powerful tool in validation. For instance, performance data of unit operations may not be as good as the vendor claims and the use case may be different, hence leading to varying results. Also, conducting experiments with different cell types can give valuable insights into characterization of the inherent variability of the process, lending the model more credibility.

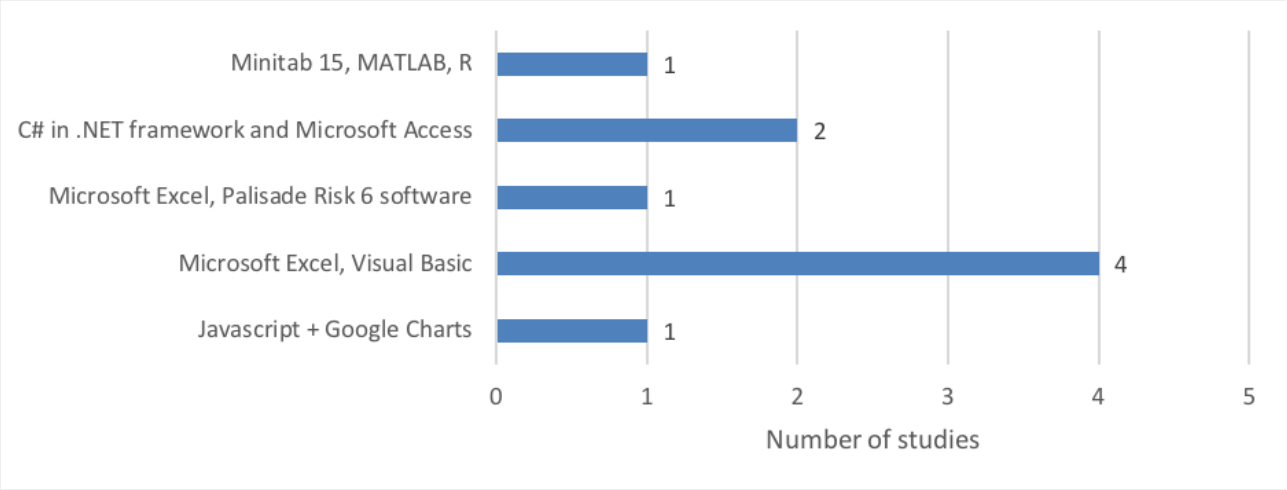
Simulation Platforms

Figure 5 shows the simulation platforms employed in the different tools. Depending on the goal of the decisional tool, different simulation tools have been used. For simpler models, using Microsoft Excel with Visual Basic has appeared to be sufficient. For instance, the Excel model constructed by Weil et al [24] consisted of a cost model with mass balance, design, sizing, resource utilization and cost-of-goods equations, database of bioprocess technology data and iterative algorithms, and scenario analysis developed using Visual Basic. However, Visual Basic codes are susceptible to Excel program upgrades, and changing formats (eg, adding a column or a row) may cause changes in the functions. C# or MATLAB allows more versatile coding experience, and for models requiring many runs, such as uncertainty or stochasticity analysis, these platforms may be more suitable.

For models with larger databases, it is worth looking into database software such as Microsoft Access. Database software provides better scalability if the volume of data is huge, and the links can be built in a more robust ways than in spreadsheets.

Visualization software such as Google Charts allows for information to be easily updated and visualized and hence is useful for presenting a lot of data in a meaningful way [20]. Dedicated add-ons, such as Palisade Risk 6, used by McCall [27], allow for Monte Carlo simulations and sensitivity analysis to be carried out a bit more easily. Without experience of using the software, however, the usability of this software has not been formally compared with implementations based on C# or MATLAB.

Figure 5. Choice of simulation platforms.



Discussion

Principal Findings

We systematically reviewed 9 articles describing decisional tools in the regenerative medicine area. The diversity of decisions that have been modeled in this area is very limited compared with work relating to monoclonal antibodies and small-molecule pharmaceuticals. Some areas that are not yet studied include scheduling [37], facility fit [38], capacity planning [39,40], supply chain optimization [41,42], and portfolio management [6].

In terms of product development systems, it is worthwhile to consider that regenerative medicines for serious life-threatening diseases can be eligible for regulatory shortcuts through various early-access initiatives, such as breakthrough therapy designation in the United States [43], Priority Medicines (PRIME) in the European Union, and Sakigake Designation in Japan [43-47]. Table 7 [17,44,48-60] shows some examples of cell and gene therapy products that have been granted regulatory pathways or designations that allow for acceleration of the commercialization process. As more and more products are granted these designations, to assist decision making, the implications of these pathways should be considered. For instance, breakthrough therapy designation allows for New Drug Application and Biologic License Application data to be submitted as they are accumulated, and orphan drug designation allows for approval of medicinal products within 6 months. The impact of these accelerated regulatory pathways can be evaluated using decisional tools to better understand the efficiency of these policies and inform future regulatory framework improvements.

Available models provide good guidance for the industry in terms of technology evaluation at various scales for large-scale allogeneic process unit operations. However, autologous products make up a significant proportion of approved regenerative medicine products approved worldwide (19 out of 36) [61]. As of 2012, more than 65% of the stem cell clinical trials contained autologous cells or tissues [62] and therefore deserve attention in the future development of decisional tools. We also noted that, despite the widespread use of simulation in the existing decisional tools, none of these used optimization algorithms that can identify and select best candidate solutions.

Additionally, as Figure 3 shows, there are no tools modeling the entire needle-to-needle process. For a more comprehensive understanding to aid decision making, the whole process from patient to patient should be considered. For instance, tissue or cell procurement can be a major constraint on the lot size and final cell count. Population doubling levels should be considered, as a higher passage number has been shown to be negatively correlated to the therapeutic potential of the cultured MSCs [63]. Harrison et al [25] conducted experiments on 3 donor samples and through their model established the challenge of donor variability on equipment sizing and of expansion potential on the final cost of goods. To provide a comprehensive account of scheduling for administration of a therapy, the availability of hospital resources should be considered. Models can be extended to cover tissue procurement and institutional requirements surrounding therapy administration in order to optimize the cost and overall patient-to-patient supply chain robustness.

As noted previously, the efficiency of differentiation and gene modification steps have shown to be a key cost driver for both autologous and allogeneic cell therapies. More in-depth evaluation of different gene editing technologies may be beneficial for driving the industry to adopt more robust and cost-effective strategies in the process step.

Similar methodologies can be applied to other novel therapeutic modalities. The approvals of alipogene tiparvovec (Glybera) in the European Union in 2012 and of voretigene neparvovec (Luxturna) in the United States in 2017 show the potential of adeno-associated virus vector-based gene therapies [54,64]; over 50 clinical candidates are using adeno-associated virus vectors [65]. Furthermore, since the approval of the first clinical trial for the clusters of regularly interspaced short palindromic repeats (CRISPR) genome editing technology in 2016, there are now over 20 active trials registered on ClinicalTrials.gov.

Table 7. Examples of cell and gene therapy products that have been granted early-access designations.

Regulatory agency and regulatory pathway	Example cell and gene therapy products
United States: Food and Drug Administration	
Priority review (1992)	Novartis: Kymriah
Accelerated approval (1992)	Pfizer: bosutinib
Fast track (1998)	Renova: RT-100 AC6 gene transfer (Ad5.hAC6); DNAtrix therapeutics: DNX-2401; AveXis: AVXS-101
Breakthrough therapy (2012)	Enzyvant: RVT-802; Juno and Celgene: JCAR017; Adaptimmune and GlaxoSmithKline: NY-ESO-1c259T; Bluebird and Celgene: bb2121
Expedited access pathway (2015)	Avita: Recell [48]
Orphan drug designation (1983)	uniQure: AMT-130 [49]
Rare pediatric disease priority review (2014)	Spark Therapeutics: Luxturna [17]
Regenerative medicine advanced therapy designation (2017)	Abeona Therapeutics: ABO-102 [50]; Mesoblast: mesenchymal precursor cell therapy [51]
European Union: European Medicines Agency	
Accelerated assessment (2004) [52]	Bluebird: LentiGlobin [53]
Orphan drug designation (2000) [52]	uniQure: AMT-130; Orchard Therapeutics: Strimvelis
Marketing authorization under exceptional circumstances (2005) [52]	uniQure: Glybera [54]
Conditional marketing authorization (2006) [52]	Chiesi Farmaceutici: Holoclar [44]; MolMed: Zalmoxis [55]
Adaptive pathway (2015) [52]	Atara Bio: ATA129
Priority Medicines (PRIME) (2016) [52]	uniQure: AMT-060, AMT-061; Juno and Celgene: JCAR017; Bluebird: LentiGlobin [53]; Adaptimmune and GlaxoSmithKline: NY-ESO-1c259T; Bluebird and Celgene: bb2121 [56]
Japan: Pharmaceuticals and Medical Devices Agency	
Priority review [57]	Glecaprevir/Pibrentasvir, AbbVie
Orphan designation (1993) [57]	Edison Pharmaceuticals: EPI-743
Conditional and time-limited approval (2014) [57]	No examples available
Sakigake forerunner review assignment (2015) [58]	Nippon-Shinyaku: NS-065/NCNP-01 [59]
China: State Administration for Market Regulation	
Accelerated and conditional approval (draft issued in 2017) [60]	Not yet in practice

As the industry moves toward delivering these novel therapies, learning from past clinical translation experiences (eg, the monoclonal antibody industry), better understanding of the risks, and making better informed decisions become all the more important.

Strengths and Limitations

There were some limitations to the review process. First, we focused our search on the MEDLINE database because preliminary scoping searches suggested that there would be more targeted literature in these databases than in those available in EMBASE and Scopus. This decision increased the risk of not identifying all relevant articles. Second, due to time limitations, we did not engage a second reviewer to review articles for eligibility, increasing the risk of excluding eligible reviews due to oversight. We consulted the Cochrane Library and PROSPERO database retrospectively, and we found no reviews to be relevant to the review question.

These limitations notwithstanding, this study is, to our knowledge, the first to systematically review the methods and logic for the design of decisional tools in aiding regenerative medicine translation and manufacturing. The small number of published studies highlights the opportunities to develop further decision support tools for different decisions and product types. The detailed design method analysis can be helpful for future development of these tools in a systematic manner in order to facilitate the translation of novel therapies into clinics more time and cost efficiently. Furthermore, the identification of the gaps in the literature can be useful for other bioprocess researchers working in the area.

Conclusions

We systematically reviewed the decisional tool landscape for regenerative medicine. Decisional tools have been shown to be instrumental in the commercialization of monoclonal antibodies through informing various decisions in manufacturing technologies, capacity planning, scheduling, and investment. As more and more regenerative medicine products receive regulatory approval, decisional tools offer a systematic way of evaluating different commercialization decisions and options.

Studies within the regenerative medicine area have largely addressed the manufacturing challenges and cost-reduction drivers for allogeneic cell therapies. Decisional tools in tissue engineering and gene therapies are lacking. To more comprehensively understand the overall costs and supply chain robustness of these lifesaving cell therapies, the entire process from tissue procurement to postadministration should be considered. To put forward industrially relevant decisional tools, costs and process assumptions should be industrially validated to ensure that any results derived from the model are useful and relevant. Future decisional tools to integrate the different facets of the available decisional tools should be developed to inform decision making in the rapidly expanding and transformative field of regenerative medicine.

Acknowledgments

This work received support from the China Regenerative Medicine International (CRMI)-Oxford Technology Centre, EIT Health, the Sir David Cooksey Fellowship in Healthcare Translation, the SENS Research Foundation, and the Oxford Academic Health Science Centre.

CL is funded by the CRMI-Oxford Technology Centre Studentship. JK is funded by the US National Institutes of Health, Grant/Award Number HL095722.

Authors' Contributions

CL conceptualized and wrote the manuscript. EM provided systematic review expertise and assisted in the development of the protocol methodology. AY provided systems engineering expertise and edited the paper for better structure, relevance, and clarity. AA, DB, ZC, AC, and JK provided comments for better relevance and clarity. All authors approved the final manuscript.

Conflicts of Interest

Funding support from the CRMI-Oxford Technology Centre is gratefully acknowledged. JK has been a paid consultant in the field of regenerative medicine for companies including Stempeutics, Sanofi, Celltex, LifeVaultBio, Takeda, and Mesoblast. JK is also an inventor on a patent that was licensed to Mesoblast. JK holds equity in Frequency Therapeutics, a company that has licensed intellectual property generated by JK that may benefit financially if it is further validated. The interests of JK were reviewed and are subject to a management plan overseen by his institutions in accordance with its conflict of interest policies. DB is a stockholder in Translation Ventures Ltd (Charlbury, Oxfordshire, UK) and IP Asset Ventures Ltd (Oxford, Oxfordshire, UK), companies that, among other services, provide cell therapy biomanufacturing, and regulatory and financial advice to pharmaceutical clients. DB is also subject to the CFA Institute’s codes, standards, and guidelines, so he must stress that this piece is provided for academic interest only and must not be construed in any way as an investment recommendation. Additionally, at the time of publication, DB and the organizations with which he is affiliated may or may not have agreed or have pending funding commitments from the organizations named here.

Multimedia Appendix 1

Executed data abstraction form.

[PDF File (Adobe PDF File), 52KB - jmir_v20i12e12448_app1.pdf]

References

1. Keen PGW. Decision support systems: the next decade. Decis Support Syst 1987;3(3):253-265. [doi: 10.1016/0167-9236(87)90180-1]
2. Lucidi S, Maurici M, Paulon L, Rinaldi F, Roma M. A simulation-based multiobjective optimization approach for health care service management. IEEE Trans Autom Sci Eng 2016;13(4):1480-1491 [FREE Full text] [doi: 10.1109/TASE.2016.2574950]
3. Rose DC, Sutherland WJ, Parker C, Lobley M, Winter M, Morris C, et al. Decision support tools for agriculture: towards effective design and delivery. Agricultural Syst 2016;149:165-174 [FREE Full text] [doi: 10.1016/j.agsy.2016.09.009]
4. Sullivan T. Evaluating environmental decision support tools. Upton, NY: Environmental Sciences Department, Environmental Research & Technology Division, Brookhaven National Laboratory; 2002 Oct. URL: https://www.bnl.gov/isd/documents/30163.pdf [accessed 2018-11-30] [WebCite Cache ID 74Jp8bUI7]
5. Mustafa MA, Washbrook J, Lim AC, Zhou Y, Titchener-Hooker NJ, Morton P, et al. A software tool to assist business-process decision-making in the biopharmaceutical industry. Biotechnol Prog 2005 Feb 04;21(1):320. [doi: 10.1021/bp049581c] [Medline: 15913287]
6. Rajapakse A, Titchener-Hooker NJ, Farid SS. Modelling of the biopharmaceutical drug development pathway and portfolio management. Comput Chem Eng 2005 May;29(6):1357-1368. [doi: 10.1016/j.compchemeng.2005.02.010]

7. Farid SS, Washbrook J, Titchener-Hooker NJ. Decision-support tool for assessing biomanufacturing strategies under uncertainty: stainless steel versus disposable equipment for clinical trial material preparation. Biotechnol Prog 2005;21(2):486-497. [doi: 10.1021/bp049692b] [Medline: 15801789]
8. Farid SS, Washbrook J, Titchener-Hooker NJ. Combining multiple quantitative and qualitative goals when assessing biomanufacturing strategies under uncertainty. Biotechnol Prog 2005;21(4):1183-1191. [doi: 10.1021/bp050070f] [Medline: 16080700]
9. Lim AC, Washbrook J, Titchener-Hooker NJ, Farid SS. A computer-aided approach to compare the production economics of fed-batch and perfusion culture under uncertainty. Biotechnol Bioeng 2006 Mar 05;93(4):687-697. [doi: 10.1002/bit.20757] [Medline: 16259001]
10. Siganporia CC, Ghosh S, Daszkowski T, Papageorgiou LG, Farid SS. Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. Biotechnol Prog 2014;30(3):594-606 [FREE Full text] [doi: 10.1002/btpr.1860] [Medline: 24376262]
11. Langer ES, Rader RA. Single-use technologies in biopharmaceutical manufacturing: a 10-year review of trends and the future. Eng Life Sci 2014 May 1;14(3):238-243. [doi: 10.1002/elsc.201300090]
12. Pollock J, Ho SV, Farid SS. Fed-batch and perfusion culture processes: economic, environmental, and operational feasibility under uncertainty. Biotechnol Bioeng 2013 Jan;110(1):206-219. [doi: 10.1002/bit.24608] [Medline: 22806692]
13. Nie W. Cost Evaluation and Portfolio Management Optimization for Biopharmaceutical Product Development [doctoral thesis]. London, UK: University College London; 2015. URL: http://discovery.ucl.ac.uk/1461733/ [accessed 2018-11-28] [WebCite Cache ID 74GNz3SX7]
14. Tsang K. Vaccine supply chain optimisation. London, UK: Imperial College London; 2006 Jan. URL: http://spiral.imperial.ac.uk/handle/10044/1/7545 [accessed 2018-05-30] [WebCite Cache ID 74GO7zSBQ]
15. Chhatre S, Francis R, O'Donovan K, Titchener-Hooker NJ, Newcombe AR, Keshavarz-Moore E. A decision-support model for evaluating changes in biopharmaceutical manufacturing processes. Bioprocess Biosyst Eng 2007 Jan;30(1):1-11. [doi: 10.1007/s00449-006-0086-8] [Medline: 17093973]
16. Mason C, Dunnill P. A brief definition of regenerative medicine. Regen Med 2008 Jan;3(1):1-5. [doi: 10.2217/17460751.3.1.1] [Medline: 18154457]
17. Morrison C. Fresh from the biotech pipeline—2017. Nat Biotechnol 2018 Jan 22:131-136. [doi: 10.1038/nbt.4068] [Medline: 29355851]
18. Rekhi R, Smith JA, Arshad Z, Roberts M, Bountra C, Bingham I, et al. Decision-support tools for monoclonal antibody and cell therapy bioprocessing: current landscape and development opportunities. BioProcess Int 2015;13(11) [FREE Full text]
19. Ungrin MD, Clarke G, Yin T, Niebrugge S, Nostro MC, Sarangi F, et al. Rational bioprocess design for human pluripotent stem cell expansion and endoderm differentiation based on cellular dynamics. Biotechnol Bioeng 2012 Apr;109(4):853-866 [FREE Full text] [doi: 10.1002/bit.24375] [Medline: 22139975]
20. Lambrechts T, Sonnaert M, Schrooten J, Luyten FP, Aerts J, Papantoniou I. Large-scale mesenchymal stem/stromal cell expansion: a visualization tool for bioprocess comparison. Tissue Eng Part B Rev 2016 Dec;22(6):485-498. [doi: 10.1089/ten.TEB.2016.0111] [Medline: 27333790]
21. Hassan S, Huang H, Warren K, Mahdavi B, Smith D, Jong S, et al. Process change evaluation framework for allogeneic cell therapies: impact on drug development and commercialization. Regen Med 2016 Apr;11(3):287-305. [doi: 10.2217/rme-2015-0034] [Medline: 26979513]
22. Simaria AS, Hassan S, Varadaraju H, Rowley J, Warren K, Vanek P, et al. Allogeneic cell therapy bioprocess economics and optimization: single-use cell expansion technologies. Biotechnol Bioeng 2014 Jan;111(1):69-83 [FREE Full text] [doi: 10.1002/bit.25008] [Medline: 23893544]
23. Hassan S, Simaria AS, Varadaraju H, Gupta S, Warren K, Farid SS. Allogeneic cell therapy bioprocess economics and optimization: downstream processing decisions. Regen Med 2015;10(5):591-609 [FREE Full text] [doi: 10.2217/rme.15.29] [Medline: 26237703]
24. Weil BD, Jenkins MJ, Uddin S, Bracewell DG, Wellings D, Farid SS, et al. An integrated experimental and economic evaluation of cell therapy affinity purification technologies. Regen Med 2017 Dec;12(4):397-417 [FREE Full text] [doi: 10.2217/rme-2016-0156] [Medline: 28621179]
25. Harrison RP, Medcalf N, Rafiq QA. Cell therapy-processing economics: small-scale microfactories as a stepping stone toward large-scale macrofactories. Regen Med 2018 Dec;13(2):159-173 [FREE Full text] [doi: 10.2217/rme-2017-0103] [Medline: 29509065]
26. Jenkins MJ, Farid SS. Cost-effective bioprocess design for the manufacture of allogeneic CAR-T cell therapies using a decisional tool with multi-attribute decision-making analysis. Biochem Eng J 2018 Sep 15;137:192-204. [doi: 10.1016/j.bej.2018.05.014]
27. McCall M, Williams DJ. Developing Cell Therapies: Enabling cost prediction by value systems modeling to manage developmental risk. JCB 2013 Apr 01;19(2):27-39 [FREE Full text] [doi: 10.5912/jcb585]
28. Remer DS, Nieto AP. A compendium and comparison of 25 project evaluation techniques. Part 1: net present value and rate of return methods. Int J Prod Econ 1995 Nov 01;42(1):79-96. [doi: 10.1016/0925-5273(95)00104-2]

29. Parnell G, Driscoll P, Henderson D. Decision Making in Systems Engineering and Management. 2nd edition. Hoboken, NJ: Wiley-Blackwell; 2010:978.
30. Lipsitz YY, Milligan WD, Fitzpatrick I, Stalmeijer E, Farid SS, Tan KY, et al. A roadmap for cost-of-goods planning to guide economic production of cell therapy products. Cytotherapy 2017 Dec;19(12):1383-1391. [doi: 10.1016/j.jcyt.2017.06.009] [Medline: 28935190]
31. Burchfield R. What's a parameter? BMJ 1998 Jun 20;316(7148):1877A [FREE Full text] [Medline: 9632408]
32. Crow D, Xueqiao W, Hancock T. Financial Times. 2018. Healthcare: cancer breakthrough leads China's biotech boom URL: http://www.ftchinese.com/story/001077339/ce?archive [accessed 2018-11-13] [WebCite Cache ID 73tfe5QgB]
33. Borowicz JJ, Bredehoeft PR, Dysert LR, Parker DE, Pickett TW, Pietlock BA, et al. Required skills and knowledge of project cost estimating. AACE International Recommended Practice No. 46R-11.: AACE; 2013. URL: http://web.aacei.org/docs/default-source/toc/toc_46r-11.pdf [accessed 2018-11-28] [WebCite Cache ID 74GQ2CgHG]
34. Hergert M, Morris D. Accounting data for value chain analysis. Strategic Manage J 1989 Mar 1;10(2):175-188. [doi: 10.1002/smj.4250100207]
35. Steward DV. The design structure system: a method for managing the design of complex systems. IEEE Trans Eng Manage 1981 Aug;EM-28(3):71-74. [doi: 10.1109/TEM.1981.6448589]
36. Hester P, Velasquez M. An analysis of multi-criteria decision making methods. Int J Oper Res 2013;10(2):56-66 [FREE Full text]
37. Jain V, Grossmann IE. Resource-constrained scheduling of tests in new product development. Ind Eng Chem Res 1999 Aug 1;38(8):3013-3026. [doi: 10.1021/ie9807809]
38. Yang Y, Farid SS, Thornhill NF. Data mining for rapid prediction of facility fit and debottlenecking of biomanufacturing facilities. J Biotechnol 2014 Jun 10;179:17-25 [FREE Full text] [doi: 10.1016/j.jbiotec.2014.03.004] [Medline: 24637375]
39. Levis AA, Papageorgiou LG. A hierarchical solution approach for multi-site capacity planning under uncertainty in the pharmaceutical industry. Comput Chem Eng 2004 May 15;28(5):707-725. [doi: 10.1016/j.compchemeng.2004.02.012]
40. Timpe CH, Kallrath J. Optimal planning in large multi-site production networks. Eur J Oper Res 2000 Oct 16;126(2):422-435. [doi: 10.1016/S0377-2217(99)00301-X]
41. Sousa RT, Liu S, Papageorgiou LG, Shah N. Global supply chain planning for pharmaceuticals. Chem Eng Res Design 2011 Nov 1;89(11):2396-2409. [doi: 10.1016/j.cherd.2011.04.005]
42. Papageorgiou LG. Supply chain optimisation for the process industries: advances and opportunities. Compu Chem Eng 2009 Dec 10;33(12):1931-1938. [doi: 10.1016/j.compchemeng.2009.06.014]
43. Corrigan-Curay J, McKee AE, Stein P. Breakthrough-therapy designation - an FDA perspective. N Engl J Med 2018 Dec 12;378(15):1457-1458. [doi: 10.1056/NEJMc1801222] [Medline: 29641963]
44. Elsanhoury A, Sanzenbacher R, Reinke P, Abou-El-Enein M. Accelerating patients' access to advanced therapies in the EU. Mol Ther Methods Clin Dev 2017 Dec 15;7:15-19 [FREE Full text] [doi: 10.1016/j.omtm.2017.08.005] [Medline: 28971109]
45. Smith JA, Brindley DA. Conditional approval pathways: the "special" case of global regenerative medicine regulation. Rejuvenation Res 2017 Feb;20(1):1-3 [FREE Full text] [doi: 10.1089/rej.2017.1932] [Medline: 28125919]
46. Kondo H, Hata T, Ito K, Koike H, Kono N. The current status of Sakigake designation in Japan, PRIME in the European Union, and Breakthrough Therapy Designation in the United States. Ther Innov Regul Sci 2017 Jan;51(1):51-54. [doi: 10.1177/2168479016662682] [Medline: 30235998]
47. Raggio M. Regulatory Education for Industry (REdl). 2015. Overview of FDA expedited programs with a focus on breakthrough therapy URL: https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/UCM466480.pdf [accessed 2018-04-02] [WebCite Cache ID 74GU6jz0J]
48. Avita Medical. FDA grants ReCell expedited access pathway designation. 2017. URL: https://www.meddeviceonline.com/doc/fda-grants-recell-expedited-access-pathway-designation-0001 [accessed 2018-11-28] [WebCite Cache ID 74GT1h3bI]
49. UniQure. uniQure announces FDA orphan designation for AMT-130 in Huntington's disease. 2017 Jan. URL: https://globenewswire.com/news-release/2017/10/06/1142262/0/en/uniQure-Announces-FDA-Orphan-Drug-Designation-for-AMT-130-in-Huntington-s-disease.html [accessed 2018-11-28] [WebCite Cache ID 74GVohIsl]
50. Abeona Therapeutics. Abeona announces FDA grants RMAT designation to ABO-102 gene therapy in MPS IIIA. 2018. URL: http://investors.abeonatherapeutics.com/phoenix.zhtml?c=63510&p=irol-newsArticle&ID=2343895 [accessed 2018-05-29] [WebCite Cache ID 73nSifcfX]
51. Mesoblast Limited. Global news wire. 2017 Dec 21. Mesoblast receives FDA regenerative medicine advanced therapy designation for its cell therapy in heart failure patients with left ventricular assist devices URL: https://bit.ly/2BezCtP [accessed 2018-11-28] [WebCite Cache ID 74GTFvPOT]
52. McBlane JW. Regulatory landscape for cell therapy--EU view. Biologicals 2015 Sep;43(5):433-436. [doi: 10.1016/j.biologicals.2015.04.004] [Medline: 25997566]
53. Stanton D. BioPharma Reporter. 2016 Sep 26. Bluebird gene therapy granted EMA accelerated assessment status URL: https://www.biopharma-reporter.com/Article/2016/09/27/Bluebird-gene-therapy-granted-EMA-accelerated-assessment-status [accessed 2018-05-29] [WebCite Cache ID 73nSeHSeT]

54. Ylä-Herttuala S. Endgame: glybera finally recommended for approval as the first gene therapy drug in the European union. Mol Ther 2012 Oct;20(10):1831-1832 [FREE Full text] [doi: 10.1038/mt.2012.194] [Medline: 23023051]
55. Mullard A. 2016 EMA drug approval recommendations. Nat Rev Drug Discov 2017 Feb 02;16(2):77. [doi: 10.1038/nrd.2017.17] [Medline: 28148940]
56. European Medicines Agency. List of products granted eligibility to PRIME. 2018. URL: https://www.ema.europa.eu/documents/report/list-products-granted-eligibility-prime_en-0.xlsx [accessed 2018-11-28] [WebCite Cache ID 74GTTPYry]
57. Jokura Y, Yano K, Yamato M. Comparison of the new Japanese legislation for expedited approval of regenerative medicine products with the existing systems in the USA and European Union. J Tissue Eng Regen Med 2018 Feb;12(2):e1056-e1062 [FREE Full text] [doi: 10.1002/term.2428] [Medline: 28211195]
58. Ministry of Health, Labour and Welfare. Strategy of SAKIGAKE. 2014. URL: https://www.mhlw.go.jp/english/policy/health-medical/pharmaceuticals/140729-01.html [accessed 2018-11-28] [WebCite Cache ID 74GTfFPC8]
59. Nakamura T. PMDA Updates. 2018 Presented at: WCBP 2018; Jan 30–Feb 1, 201; Washington, DC, USA URL: https://cdn.ymaws.com/www.casss.org/resource/resmgr/wcbp_speaker_slides/2018_WCBP_NakamuraTakahiro.pdf
60. State Administration for Market Regulation. [Encouraging Innovation and Accelerating the Review and Approval of New Drugs and Medical Devices (Draft for comment)]. 2017. URL: http://www.sfda.gov.cn/WS01/CL0087/172567.html [accessed 2018-11-09] [WebCite Cache ID 73nSO5868]
61. Alliance for Regenerative Medicine. Available products. URL: https://alliancerm.org/available-products/ [accessed 2018-11-09] [WebCite Cache ID 73nSz4haZ]
62. Li MD, Atkins H, Bubela T. The global landscape of stem cell clinical trials. Regen Med 2014 Jan;9(1):27-39. [doi: 10.2217/rme.13.80] [Medline: 24236476]
63. Bonab MM, Alimoghaddam K, Talebian F, Ghaffari SH, Ghavamzadeh A, Nikbin B. Aging of mesenchymal stem cell in vitro. BMC Cell Biol 2006 Mar 10;7:14 [FREE Full text] [doi: 10.1186/1471-2121-7-14] [Medline: 16529651]
64. Spark Therapeutics Inc. FDA approves Spark Therapeutics LUXTURNATM (voretigene neparvovec-rzyl), a one-time gene therapy for patients with confirmed biallelic RPE65 mutation-associated retinal dystrophy. URL: http://ir.sparktx.com/news-releases/news-release-details/fda-approves-spark-therapeutics-luxturnatm-voretigene-neparvovec [accessed 2018-11-09] [WebCite Cache ID 73nSoMzfQ]
65. Naso MF, Tomkowicz B, Perry WL, Strohl WR. Adeno-associated virus (AAV) as a vector for gene therapy. BioDrugs 2017 Aug;31(4):317-334 [FREE Full text] [doi: 10.1007/s40259-017-0234-5] [Medline: 28669112]

Abbreviations

CAR-T: chimeric antigen receptor T cells
FACS: fluorescence-activated cell sorting
MACS: magnetic-activated cell sorting
MSC: mesenchymal stem cell
NPV: net present value
PRIME: Priority Medicines
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 08.10.18; peer-reviewed by S Rasmussen, L Xifeng, HT Yang; comments to author 01.11.18; revised version received 19.11.18; accepted 20.11.18; published 19.12.18.

Please cite as:
Lam C, Meinert E, Alturkistani A, Carter AR, Karp J, Yang A, Brindley D, Cui Z
Decision Support Tools for Regenerative Medicine: Systematic Review
J Med Internet Res 2018;20(12):e12448
URL: http://www.jmir.org/2018/12/e12448/
doi:10.2196/12448
PMID:30567696

Original Paper

Enhancing Privacy Controls for Patients via a Selective Authentic Electronic Health Record Exchange Service: Qualitative Study of Perspectives by Medical Professionals and Patients

Ala Sarah Alaqra[1], BSc, MSc; Simone Fischer-Hübner[1], Prof Dr; Erik Framner[2], BSc

[1]Privacy and Security Research Group, Department of Computer Science, Karlstad University, Karlstad, Sweden

[2]Department of Information Systems, Karlstad University, Karlstad, Sweden

Corresponding Author:
Ala Sarah Alaqra, BSc, MSc
Privacy and Security Research Group
Department of Computer Science
Karlstad University
Universitetsgatan 2
Karlstad, 65188
Sweden
Phone: 46 547001734
Email: alaa.alaqra@kau.se

Abstract

Background: Patients' privacy is regarded as essential for the patient-doctor relationship. One example of a privacy-enhancing technology for user-controlled data minimization on content level is a redactable signature. It enables users to redact personal information from signed documents while preserving the validity of the signature, and thus the authenticity of the document. In this study, we present end users' evaluations of a Cloud-based selective authentic electronic health record (EHR) exchange service (SAE-service) in an electronic health use case. In the use case scenario, patients were given control to redact specified information fields in their EHR, which were signed by their doctors with a redactable signature and transferred to them into a Cloud platform. They can then selectively disclose the remaining information in the EHR, which still bears the valid digital signature, to third parties of their choice.

Objective: This study aimed to explore the perceptions, attitudes, and mental models concerning the SAE-service of 2 user roles: signers (medical professionals) and redactors (patients with different technical knowledge) in Germany and Sweden. Another objective was to elicit usability requirements for this service based on the analysis of our investigation.

Methods: We chose empirical qualitative methods to address our research objective. Designs of mock-ups for the service were used as part of our user-centered design approach in our studies with test participants from Germany and Sweden. A total of 13 individual walk-throughs or interviews were conducted with medical staff to investigate the EHR signers' perspectives. Moreover, 5 group walk-throughs in focus groups sessions with (N=32) prospective patients with different technical knowledge to investigate redactor's perspective of EHR data redaction control were used.

Results: We found that our study participants had correct mental models with regard to the redaction process. Users with some technical models lacked trust in the validity of the doctor's signature on the redacted documents. Main results to be considered are the requirements concerning the accountability of the patients' redactions and the design of redaction templates for guidance and control.

Conclusions: For the SAE-service to be means for enhancing patient control and privacy, the diverse usability and trust factors of different user groups should be considered.

(J Med Internet Res 2018;20(12):e10954) doi:10.2196/10954

KEYWORDS

privacy; patient data privacy; electronic health record; user control; data protection; data security; eHealth; human computer interaction

Introduction

Background

Privacy has been acknowledged by the Council of Europe's Convention on Human Rights in 1950 as a basic human right. A well-acknowledged definition of privacy was provided by the German Constitutional Court, which defined privacy as the right to informational self-determination [1], allowing individuals to determine for themselves (and thereby control) what personal information about themselves they disclose under which conditions to others.

Due to the sensitivity of medical data, the privacy of patients has been seen as essential for trust relationship between medical professionals and patients [2] over centuries, as addressed by the Hippocratic Oath [2,3].

One fundamental privacy principle that entails control over information is data minimization. It states that privacy can be best protected if personal data are not collected nor processed at all or if the amount of personal data processing is limited to the minimum necessary, at least. As the European Union (EU) General Data Protection Regulation (GDPR) requires in its Art 5 I (c), personal data shall be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed" (data minimization) [4].

Broad ranges of Privacy Enhancing Technologies (PETs) have been developed for technically enforcing data minimization that play a key role when designing systems for privacy. One example of such a PET for data minimization on content level is redactable signatures (also called malleable signatures), which enable the redaction (*blacking-out*) of personal information from signed documents while preserving the validity of the signatures [5].

In the EU H2020 project PRivacy and Security MAintaining services in the CLOUD (PRISMACLOUD), redactable signatures are used for developing a Cloud-based selective authentic electronic health record (EHR) exchange service (SAE-service) in a privacy-enhanced electronic health (eHealth) use case. In contrast to traditional digital signatures, which imply that any changes to a signed document will invalidate the signature, redactable signatures allow the controlled redaction of certain parts of the signed data without the signature losing its validity. Any unauthorized modification would, however, invalidate the signature. Hence, both authenticity and integrity of the data are protected.

An EHR is defined as "computerized record of a person's health and/or medical history..." [6-8]. In our studies, we have considered the EHR term in the hospital system for referring to medical documents. However, some might consider signed EHRs in the Cloud portal of our scenario to be personal health records (PHRs). As the concept of PHR has been noted by Wiljer et al [8] to be controversial, and was stated that no widely accepted definition exists, we, therefore, refrained from using the term PHR in our study. In addition, in our scenario, medical documents are to be used for medical purposes (second diagnosis). For simplicity reasons, we chose to use the terms EHR and medical document or record interchangeably in this paper for both the hospital and the Cloud portal.

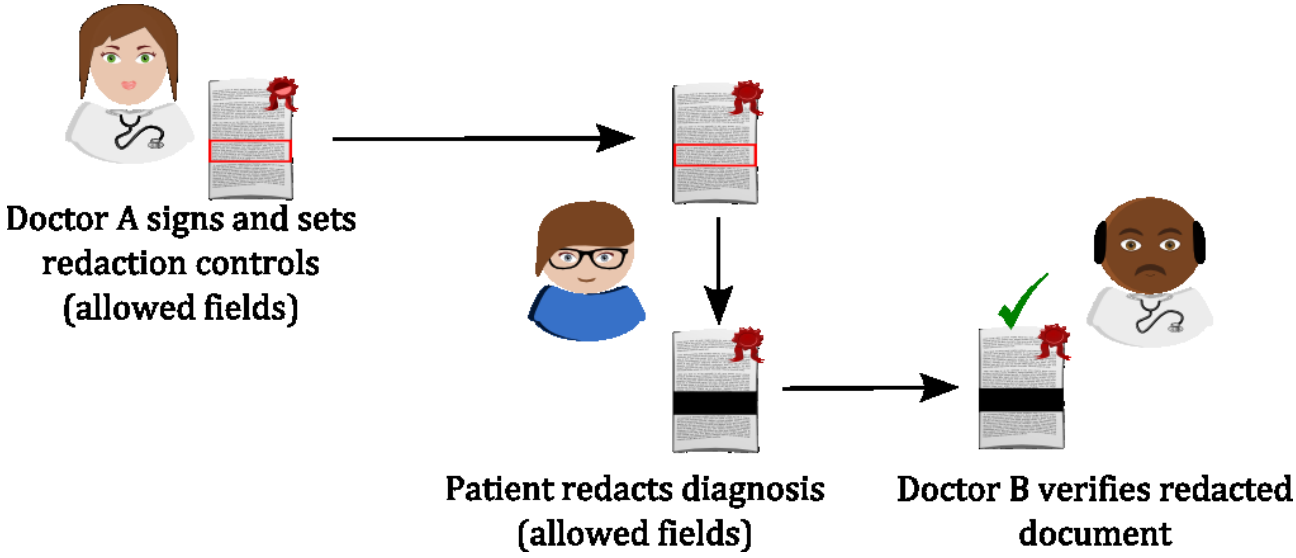
In the PRISMACLOUD eHealth use case, patients are given control and allowed to redact information in their EHR (Figure 1). In a hospital system, a medical professional (doctor A) signs the EHR with a redactable signature. The EHR is then transferred to the patient's account on a hospital Cloud platform. The patient is then able to *black-out* the predefined redactable fields of information from the signed EHR copy on the Cloud portal. Meanwhile, the signature of doctor A remains valid and the authenticity of the medical document is maintained as long as the patient is following the redaction rules. For instance, if the patient wants to get a second opinion on a diagnosis of their EHR containing blood test results, the diagnosis fields could be redacted from the EHR by the patient. The redacted EHR including only the blood values is then made available on the Cloud portal to a specialist of the patient's choice. The specialist (doctor B) can validate the signature by doctor A, and thus, verify the authenticity of the patient's blood value data (that they are indeed medical data that were collected by doctor A), which is important for protecting the patient's safety.

Hence, both user-controlled data minimization and authenticity of the selectively disclosed medical data can be provided. In an alternative use case, for producing a signed sick leave letter for the employer, the patient could redact all fields except for the fields stating the period for that the patient stayed in the hospital.

Redactions can be either implemented as an unkeyed operation that allows any party to redact the document or as a keyed operation requiring that the redactor uses a secret redaction key, which means that the redactor could later also be made accountable for the redaction.

A recent Eurobarometer survey requested by the European Commission showed that a majority of respondents would like Web-based access to their medical records, whereas the question whether they would like to grant access to their records to third parties depends on the type of recipient [9]. Moreover, earlier studies revealed that patient- (or more generally, user-) determined privacy controls and restrictions on the content and/or recipient may be a prerequisite of sharing [10,11], whereas privacy concerns and a lack of selective controls have a negative influence on the intention to share medical information even with other health care providers [12] and may reduce patient care quality [13]. As discussed in Caine et al's study [14], patients would like to have granular privacy controls over their health information in medical records allowing them to differentially share their data in medical records or only parts of it, depending on the data recipient of and/or type of medical data. The SAE-service provides a technical solution for such granular privacy control that is demanded by Caine et al [14] for maintaining the level of privacy afforded by medical records and for achieving alignment with patient preferences. At the same time, the SAE-service also protects the authenticity of the selectively disclosed data for safeguarding the patient's safety.

Figure 1. Redactable signatures in the PRISMACLOUD (PRivacy and Security MAintaining services in the CLOUD) eHealth use-case.



From a human-computer interaction (HCI) standpoint, the design of user interfaces (UIs) for such SAE-service poses several challenges. In particular, privacy crypto schemes may be counter-intuitive to users. Therefore, it is a challenge to design UIs for evoking comprehensive mental models [15]. This problem is increased by the fact that redactable signatures work differently from traditional signature schemes, which in contrast to redactable signatures, get invalid if the signed document is redacted, that is, changed. This may affect the trust that users with some familiarity with crypto technologies may have in such a PET. Moreover, different user groups among medical professionals and patients may have different expectations and requirements concerning this SAE-service, which need to be appropriately addressed.

The importance of end users' participation as stakeholders in the privacy by design (PbD) process, involving multiple disciplines, including usability design in addition to engineering, has been emphasized earlier [16]. End users should ultimately profit from PbD, where it has been pointed out that UIs need to address PbD and be "human-centered, user-centric, and user-friendly, so that informed privacy decision may be reliably exercised" [17]. Throughout the PRISMACLOUD project, we have followed the user-centered design (UCD) approach [18], which meant that the focus was on users throughout the development, design, and evaluation of UI prototypes for this SAE-service. The eHealth use case addressed 2 types of stakeholders who were involved as end users in our studies. They are the signers of medical documents who are medical professionals (eg, doctors) and redactors of medical documents who are users playing the role of patients.

Objective

This study reports about the results of our research that has been addressing the following research questions:

- What are the perceptions, attitudes, and mental models that users of both roles, signers (medical professionals) and redactors (patients with different technical knowledge), from Germany and Sweden, have with regard to patient-controlled redactions as part of the SAE-service?
- What are end user requirements for making of redactable signatures as part of the SAE-service usable?

We have included individuals with varying technical background performing the redactor's role, as we were interested in investigating whether different levels of knowledge of crypto technologies will affect their understanding and trust in redactable signatures. Moreover, as this study is conducted within the scope of an EU research project, we involved end users in 2 EU countries (corresponding to project partner's locations: Germany and Sweden), which also allowed us to investigate possible national influences. The first research question looks into the broader contribution of understanding users and serves as a prerequisite to the second research question where requirements are derived for the SAE-service.

Methods

Overview

We have followed a UCD approach for developing and evaluating UI prototypes for this SAE-service. UCD approach focuses on the needs of users and integrating that into the design processes [18].

Therefore, in our previous work, we involved end users for the elicitation of an initial set of requirements (found in Table 1). It was first done via semistructured interviews and stakeholder workshops as described and analyzed in Alaqra et al's study [19].

These initial requirements were considered for the design of low-fidelity UI prototypes (mock-ups) for the SAE-service (as shown in the subsection Mock-Ups User Interface Design). The design of the mock-ups went through several iterations of experts' reviews and walkthroughs, which were used in this study for the evaluation and facilitation of the discussion of both of our studies (shown in the following sections).

As our research is explorative with the objective to investigate and gain a deeper understanding about the users' perceptions, attitudes, and mental models, we chose qualitative empirical means in our approach.

In total, there were 2 studies conducted with the 2 categories of users, that is, the signers and the redactors, respectively. Both studies followed a semistructured format, unlike structured methods, which allows the freedom and openness of the discussions to explore one's perspectives and opinions. However, both studies had a specific set of topics for the discussions. These topics are presented and highlighted in the mock-ups UI themes, which served as the main facilitator of the discussions (see section Mock-ups User Interface Design for Electronic Health).

The 2 studies conducted were corresponding to the mock-ups parts: (1) individual walk-throughs (interviews) with medical professionals to evaluate the hospital platform mock-ups and (2) group walk-throughs (focus groups; FGs) with users, prospective patients, to evaluate the Cloud portal mock-ups.

Recruitment

Both studies were conducted with participants in Germany and Sweden (specifically Värmland County), not only because the PRISMACLOUD consortium includes partners from those countries but also because they are different in terms of the digitization of eHealth infrastructures. Sweden is regarded as one of leading EU countries in eHealth use [20,21] and national EHR system development [22]. The significant progress in moving toward eHealth has been contributed by a well-developed Information and Communication Technology (ICT) infrastructure [20], with a fully integrated EHR system on both county and national level.

In accordance with the national patient summary *Nationell patientöversikt* strategy, health care professionals can be given direct access to a patient's health records that are kept by a care provider in any of the country's 21 county councils. Via the national Web-based portal, citizens in a number of counties, including Värmland, have access to view their personal health data and request services [22]. In Germany, on the other hand, patient data are mainly documented on paper; and previous studies show that Germany is facing multiple obstacles that prevent the implementation of a national EHR system [20,22,23]. Moreover, as also the latest Eurobarometer survey on data protection from 2015 confirmed, citizens in those 2 countries have different levels of perceptions with regard to control over their personal sphere on the Web and trust that individuals have in different entities: Germany had, for instance, the highest number of respondents who think that they have no control or only partial control over their personal data, whereas people in Sweden are most likely to trust their national public authorities [24]. All these different factors motivated our choice of conducting our studies in both Germany and Sweden for also analyzing possible national influences.

For both studies, we invited participants via professional and personal contact networks and offered them lunch as a compensation. All interviews and FGs took place between mid-May and mid-June 2017.

Documentation and Analysis

For every FG and interview, there were 2 expert interviewers: 1 moderator for the discussion and 1 for note taking. Voice recording was used as a reference for the notetaking process. After the sessions, the interviewers documented their notes using the voice recording. Notes were collected and combined from the interviewers, then were iteratively categorized and evaluated into themes. Finally, the initial requirements were refined and concluded under each theme, which are summarized and presented in the Results section.

Ethical Review

Our evaluation research plan was submitted to the ethics review board at Karlstad University for approval. They decided in their meeting on May 9, 2017 that our evaluation experiment would not fall under the Swedish Ethical Review Act [25] and were, therefore, approved before we started with conducting our evaluations.

Participation in the interviews and FGs was restricted to adult volunteers, who provided their consent after being informed, both orally and in written form, about our privacy policy.

According to the Swedish Ethical Review Act, ethical review by a regional ethical review board would be required if sensitive personal data were collected or processed within the scope of the research project. We conducted FGs with users in the role of a patient; however, did not collect any of their personal medical data. We clearly advised all participants to take the role of a specified persona, that is, of made-up persons, during the FG discussions. We strictly advised them to not talk about any personal matters and confine their discussions to their persona's point of view. We informed them that in case they talked about any personal sensitive information, we would stop the recording of the session directly and delete that recorded part.

Individual Walk-Throughs: Interviews

These interviews were conducted to understand medical professionals' perspectives and opinions regarding redactable signatures from the signer's point of view. Currently, in the given eHealth scenario, doctors will have to sign the EHR with redactable signatures. We chose individual walk-throughs, that is, one-on-one interviews, as medical professionals who were recruited came from different fields and had different expertise. In addition, it was technically not plausible to gather many doctors at a specific time to conduct an evaluation.

Protocol

For addressing the signers' of redactable signatures perspectives, we used the hospital platform mock-ups. Individual walk-throughs were carried out with medical staff in the form of semistructured interviews that lasted an average of 35 to 40 min. Consent forms were explained and handed out for participating in the study and for recording the session (see Multimedia Appendix 1. Consent form for interview participants). All interviewees consented to the voice recording of the sessions. An overall introduction to EHRs redactable signatures and the eHealth use case scenario was given before the mock-up's UI testing.

Participants were given an overall task: to log in, sign the EHR of a made-up patient *Josh Brown*, and then export it to the Cloud portal. The latter task given to participants is made up of a sequence of mock-ups pages. The main mock-up pages and the

main theme of discussion that corresponds to our research questions are as follows:

1. Sign-in page: 2-factor authentication for authenticating the signer
2. Dashboard page: viewing and selecting EHRs
3. View unsigned Josh Brown's medical record page: overview and showing possible redaction templates
4. Signing the document: signature visualization

Group Walk-Throughs: Focus Groups

In our scenario, once a doctor has signed the EHR with a redactable signature, patients will be able to redact their medical documents in the Cloud portal. In this study, the Cloud portal mock-ups were used for addressing the redactors' point of view, that is, the patients. In our FGs, we had gathered participants who could be those potential patients. FGs allow us to have in-depth discussions with different sets of users and understand their standpoint regarding redactions. The nature of a group encourages discussions and generates interactivity among participants. In addition, in our study design, the first part of our FG sessions included an interactive persona discussion (see below) that required a group discussion and interaction for an in-depth elaboration of the participants' attitudes and perceptions of selective disclosures.

In addition, we have chosen to involve user groups with different levels of technical expertise and knowledge of cryptographic tools to test whether background knowledge with encryption would influence trusting the validity of the signature after the redaction of medical documents. In particular, we wanted to investigate whether the technical users would expect that a redactable signature rather works similarly as a traditional digital signature and what that would imply in terms of their trust in the system.

Protocol

The FG sessions lasted approximately 2 hours, including lunch (all but FG5, which lasted 1 hour and 30 min). Consent forms were handed out for participating in the study and for recording the session (see Multimedia Appendix 2. Consent form for focus groups participants); all participants consented to the voice recording of the sessions. Participants were reminded not to disclose any personal information about themselves but rather discuss from the perspectives of the personas, that is, made-up persons that were assigned to them.

The first part of the session included a small exercise that included redactions of personas' information on papers to understand their perspectives on information privacy and sharing meaning that they were given cards with information describing their personas. A persona consisted of first name, age, weight, marital status, address, hobbies and interests, occupation, salary, medical condition(s), religious affiliation, political affiliation, and sexual orientation.

They were given a few minutes to read their persona's cards and blackout or redact information they will not disclose to their fellow FG participants. They were instructed to disclose only their personas' name, whereas the remaining information is optional. Finally, they were asked to present the information they chose to keep and share with their fellow FG members.

A general discussion followed on why some information was not shared by participants and reasoning behind selective disclosure, on the importance of hiding some information, and in which context. We asked participants to stick to the context of a FG. The discussion focused on sharing information in FGs such as the one they are participating in using their personas' cards.

After the general discussion, an overall introduction to the eHealth use case scenario was given; however, redactable signatures were not described to the nonlay user groups (FG2, FG4, and FG5) but rather later explained after the mock-up's evaluation was done. One volunteering participant was chosen to control the walk-through of the mock-ups. The main task given to participants is to sign in on behalf of the persona Josh Brown, redact the document, and then send it to the Cloud. The following are the main mock-up UI themes of discussion that corresponds to our research questions:

1. Sign-in page: 2-factor authentication for the redactor
2. Dashboard page: viewing and selecting EHRs
3. EHR redaction: blacking-out metaphor
4. Redaction templates: support and guidance

Mock-Ups User Interface Designs

Low fidelity mock-ups have been designed using Balsamiq tool for wire framing [26] to clearly signal to the test participants that the discussions should focus on the general functionality and not on specific design issues. On the basis of the requirements and analysis of redactable signatures in Alaqra et al's study [19], Table 1 shows the list of main HCI requirements that served as a basis for the mock-ups design.

Table 1. Redactable signatures' requirements to mock-ups design.

Requirement index	Description
RQ[a]1	Unobtrusive, easy-to-use, and multifactor authentication
RQ2	Private Cloud run by authorities and branding of (trustworthy) system owner
RQ3	Support (eg, templates) or guidance on redaction considering both privacy and safety
RQ4	Clear responsibilities, that is, the redactor must be accountable
RQ5	User-friendly signature solutions
RQ6	Suitable metaphors and human-computer interaction concepts

[a]RQ refers to a code used for requirement.

Following the eHealth use case scenario mentioned in the Introduction, the mock-ups UIs make up 2 parts: the hospital platform, which is used by medical staff to sign the EHR and a Cloud portal, where patient can view and redact their signed EHR. Requirement 2 (RQ2) is addressed in the UI in the form of considering the hospital's own trusted platform and a private Cloud portal. The following subsections include the description of the mock-ups UIs designed and highlight the UIs that are considering the requirement. We highlight our investigation purposes in each part that corresponds to the UI functions and/or features.

Mock-Ups Interfaces: Hospital Platform for Medical Staff

Signing-In and Two-Factor Authentication: Requirement 1 (RQ1)

On the hospital platform's in Figure 2 (1), *Sign-in* page, the user will enter a user name and a password and then click on the *Sign-in* button, a dialogue box will appear as an extra authentication factor (Figure 2; 2). In accordance with a secure authentication solution of MOXIS's [27] developed by PRISMACLOUD partner XiTrust and requirement (RQ1), users will use the 2-factor authentication for signing in. They will receive a short messaging service (SMS) text message code on their mobile phone, which is entered into Figure 2 (3) the system before completing the sign-in process. We aimed to test user's familiarity with the 2-factor authentication process and understand their thoughts regarding its usability and concerns.

Dashboard: Viewing and Selecting Electronic Health Records: Requirement 6 (RQ6)

Once the doctor has signed in, he or she will reach Figure 3 (4) the home page that shows a list of medical documents and notes produced in conjunction with patient's encounters. Below the header (in the section's body), documents are grouped patient wise. Each of these rows includes a document icon, a document title, the time in which the document was created, and a clickable *export to Cloud* icon to the far right.

Above the list, one will find a search field and filtering elements that can be used to search for a particular patient or to filter out nonsigned, signed, and/or shared documents that one is not looking for in the list. We aim to test if the icons are recognizable (ie, if they are suitable metaphors as required by RQ6) and if the documents view matches users' mental models in real application situations.

Figure 2. Signing-in and two-factor authentication in Hospital platform using MOXIS.

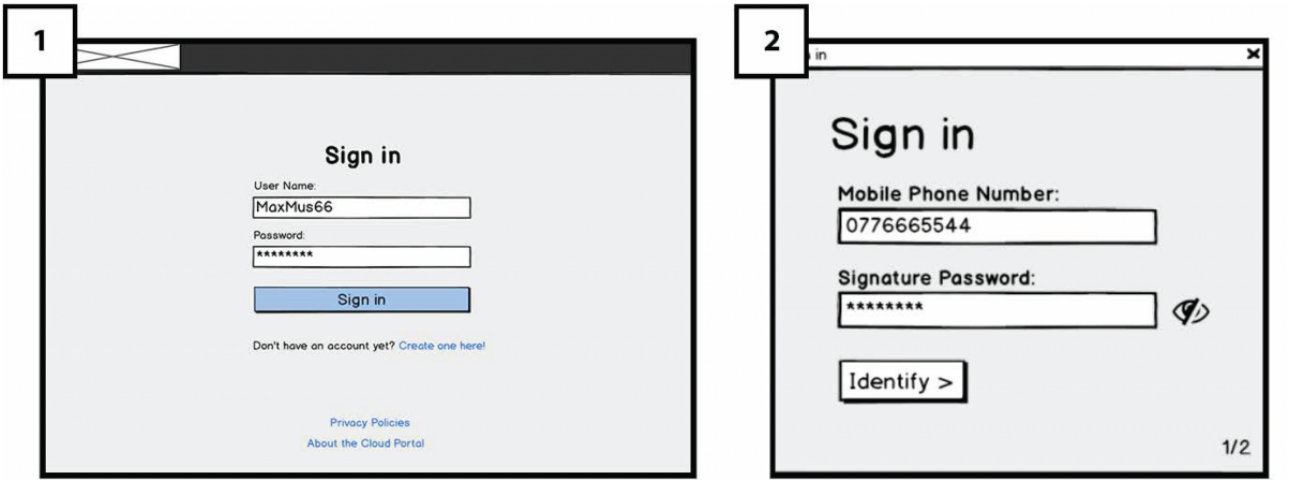
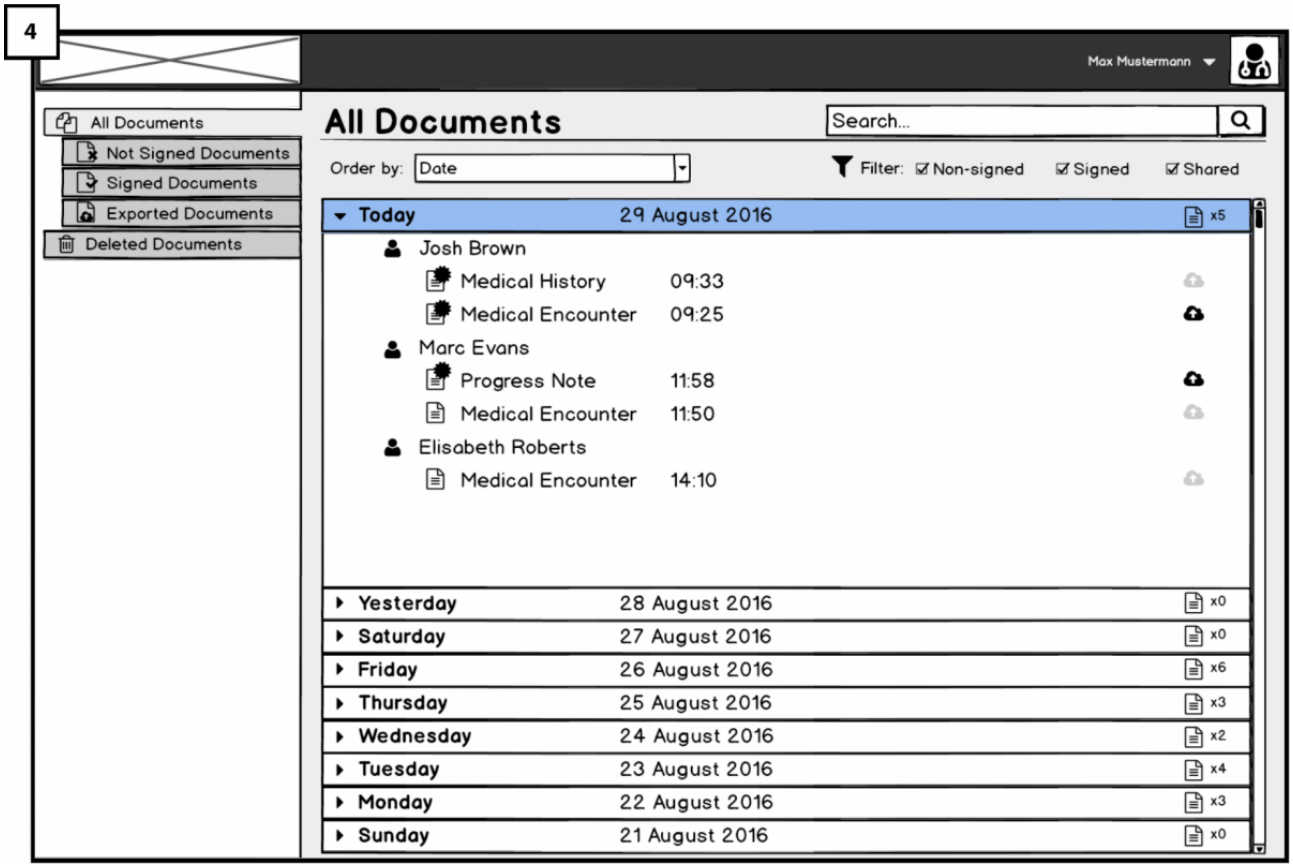


Figure 3. Dashboard of Hospital platform for viewing all electronic health records (EHRs).



Overview of Electronic Health Record and Showing Possible Redaction Templates: Requirement 3 (RQ3)

When the doctor selects a document to sign, he or she will reach an overview page shown in Figure 4 (5). On the right side of the document overview, the doctor will find a box titled *fields relevant for each type of redaction*, containing different options such as *sick leave* (allowing patients to conduct redactions for creating a sick leave letter for the employer). These options constitute templates that the patient on the Cloud portal side can use (for different purposes) to create redacted versions of the document, once it has been exported to the patient's account on the Cloud Portal. The intended use for the display of templates is to guide and show doctors what is meant by redacting documents, opening the room for discussing their opinions regarding redactions and patients redacting their document (RQ3). In addition, we intended to show doctors what might happen if they sign the documents using redactable signatures, documents will remain valid despite some fields being redacting according to redaction rules. In Figure 5 (6), different redaction templates are presented with the possible redactions; highlighted fields correspond to fields remaining after the redactions were done by the patient.

In Figure 5 (7), below the fields relevant for each type of redaction box, there is a signature placeholder that should be attached to the last page of the document before signing it. By attaching the placeholder to the document (through drag-and-drop), the doctor indicates where his or her signature should be placed on the document once it is signed. We are interested to test how the functionality of signature placeholder works and whether users understand it and find it useful.

After attaching the placeholder and clicking on the sign button, dialogue box (Figure 6; 8) will appear in which the signing is completed through a 2-factor authentication (Figure 6; 9). Thereby, the doctor allows the patient to perform redactions on the document in the future.

Signature Visualization: Requirement 5 (RQ5)

The signed EHR with the redactable signature will have a visual representation of the doctor's handwritten signature. As shown in Figure 7 (10), the handwritten presentation of the redactable signature is shown at bottom of the last page of the EHR. The use of such visual representation of the digital signature (Figure 7; 11) is thought to be more intuitive to users to have a visual confirmation that the document is signed. We aimed to test whether users understand this feature and if it is serving its purpose.

Figure 4. Overview of medical record/electronic health record (EHR) to be signed in Hospital platform.

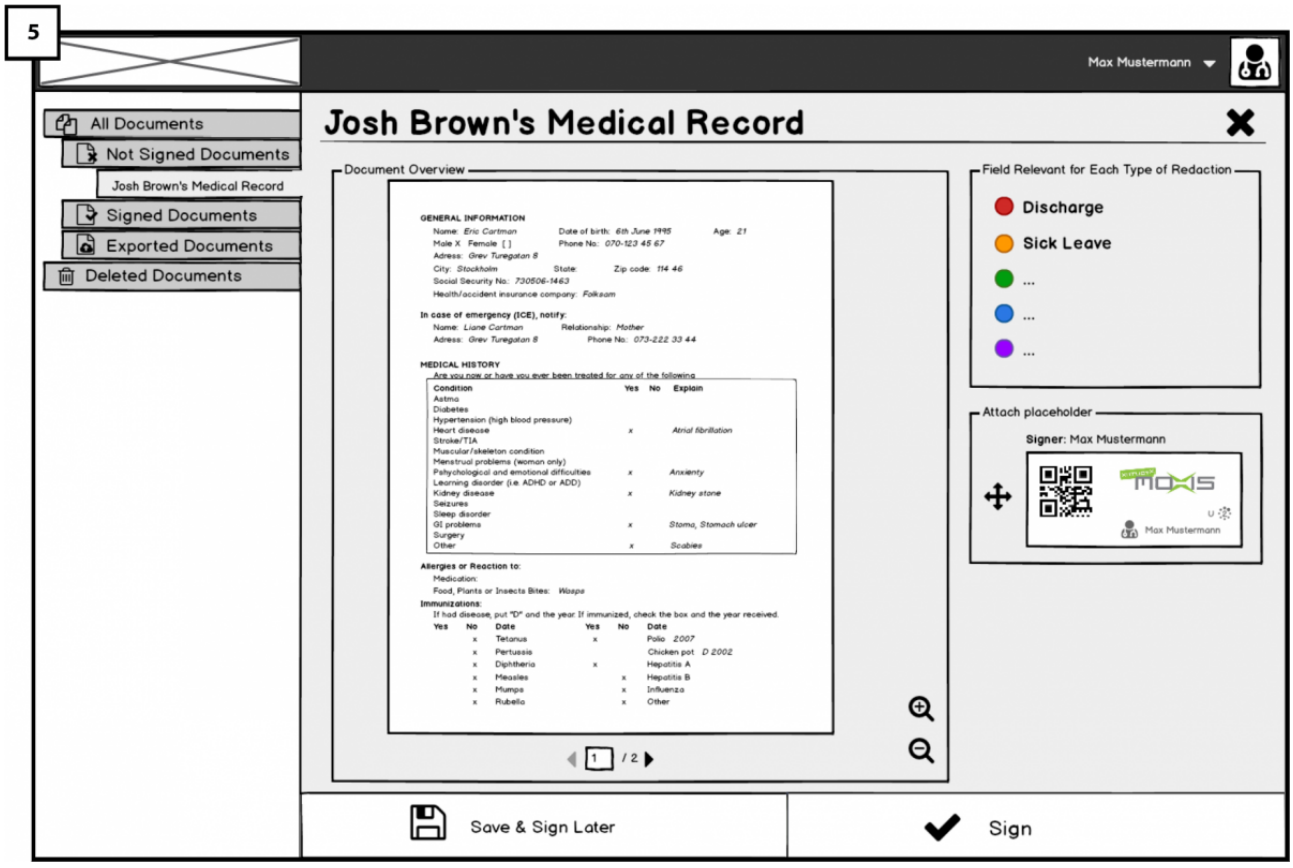


Figure 5. View of possible redaction templates.

Figure 6. Signing the electronic health record (EHR) with redactable signatures.

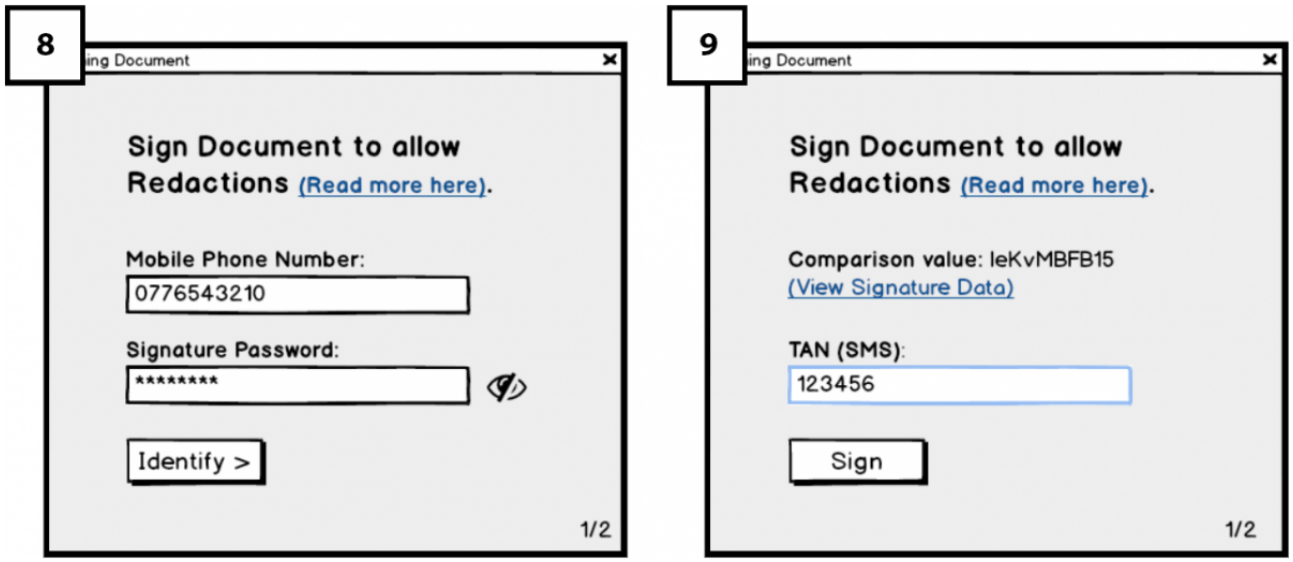
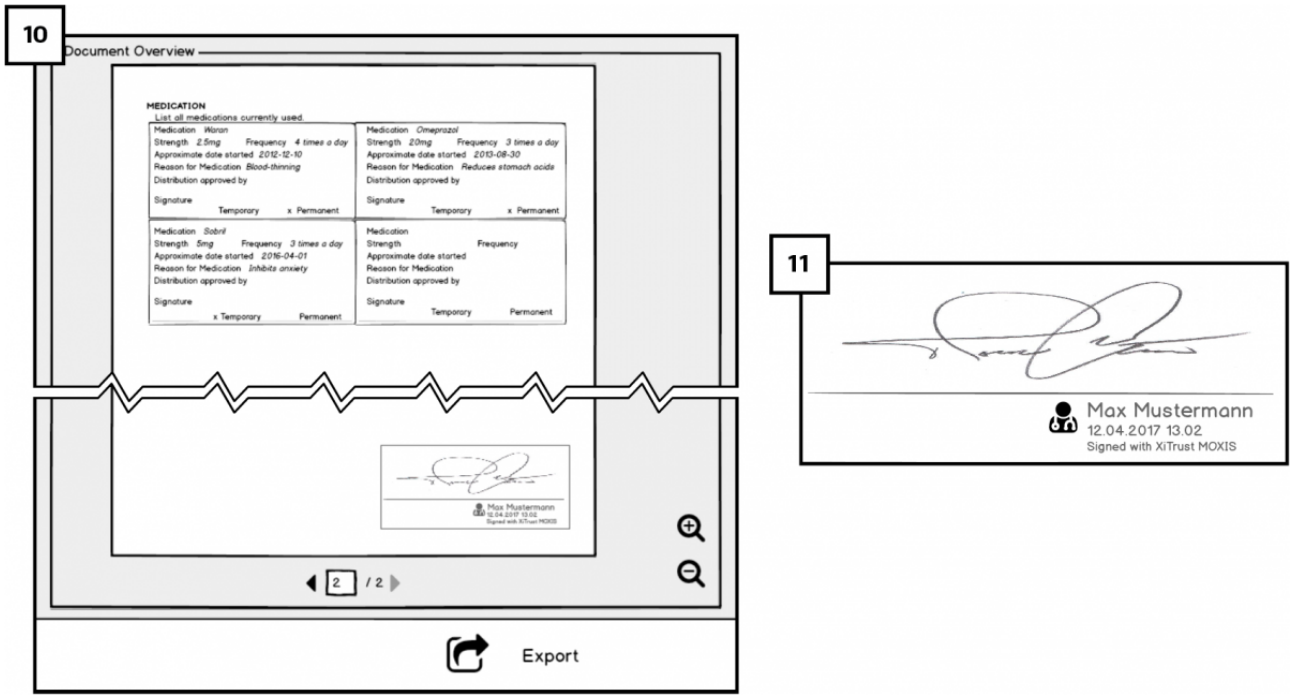


Figure 7. Visualization of the redactable signature and exporting to the Cloud.



Mock-Ups Interfaces Sequences: Cloud Portal for End Users

Signing In With Two-Factor Authentication

In similarity to the hospital platform, users will sign into the Cloud portal by using a 2-factor authentication [27].

Dashboard: Viewing and Selecting Electronic Health Records

After completing the sign-in process, users will reach Figure 8 (1) the dashboard. By clicking on + *new redaction* in the side menu, users can redact their EHRs.

Electronic Health Record Redaction Metaphor: Blacking-Out: Requirement 6 (RQ6)

Alternative views on redacted documents for showing either what text will remain and what text will be redacted is based on RQ6. When redacting the EHRs, the metaphor of blacking-out (or more precisely *graying-out*) is used in the form of a *stencil* that is placed on top of the EHR. It is intended to provide patients with guidance on the recommended amount of information to redact from the document (see Figure 9). The text to be redacted is only *grayed-out* with dark gray instead of blacked-out so that patients can still read and check what information will be redacted.

Redaction Templates: Support and Guidance Requirement 3 (RQ3)

Our mock-ups UIs address RQ3 by providing a choice of redaction templates that users can use for different contexts, and that should be created by specialists taking both privacy and safety aspects into consideration. In Figure 10 (2), after users have selected a document to redact, they select a template. They can either select a predefined template in a drop-down list (eg, *discharge* or *sick leave*) or create a new template by clicking on the + *add new template* link on the right side of the list. Below the drop-down list, the template's effect is indicated in a *document before* and document after view. In the former, users are able to choose between 2 ways of representing redaction: highlight fields that will be kept (Figure 10; 2) or fields that will be redacted (Figure 10; 3). Once the template is selected, users are allowed to redact manually more or less information if they want to, as long as the respective fields are marked as redactable. Patients are redacting information; they should receive immediate visual feedback of the graying out as well as the validity of the doctor's signature.

Signing the Redaction: Accountability of Redaction Requirement (RQ4)

Redactors are requested to perform a keyed-operation when redacting EHRs for making them accountable; thus, addressing RQ4. Therefore, after users have selected the template and/or redactable fields to be redacted, they will need to sign their redaction to complete the process.

This is done by attaching the signature placeholder to the document (similar to the hospital platform) and 2-factor authentication for signing. The final view of the medical document will have both signatures: the doctor's and the patient's. Besides, green check icons next to the word *valid* on the right side indicate the validity of both signatures. The validity of doctor's signature is dependent on the redactions performed by the patient. The patient's signature next to the doctor's (Figure 11; 4) is for showing the accountability of the patient who has redacted the document. We aimed to investigate users' opinions about accountability of redactions and whether the presented solution raises any concerns.

Figure 8. Dashboard of Cloud portal.

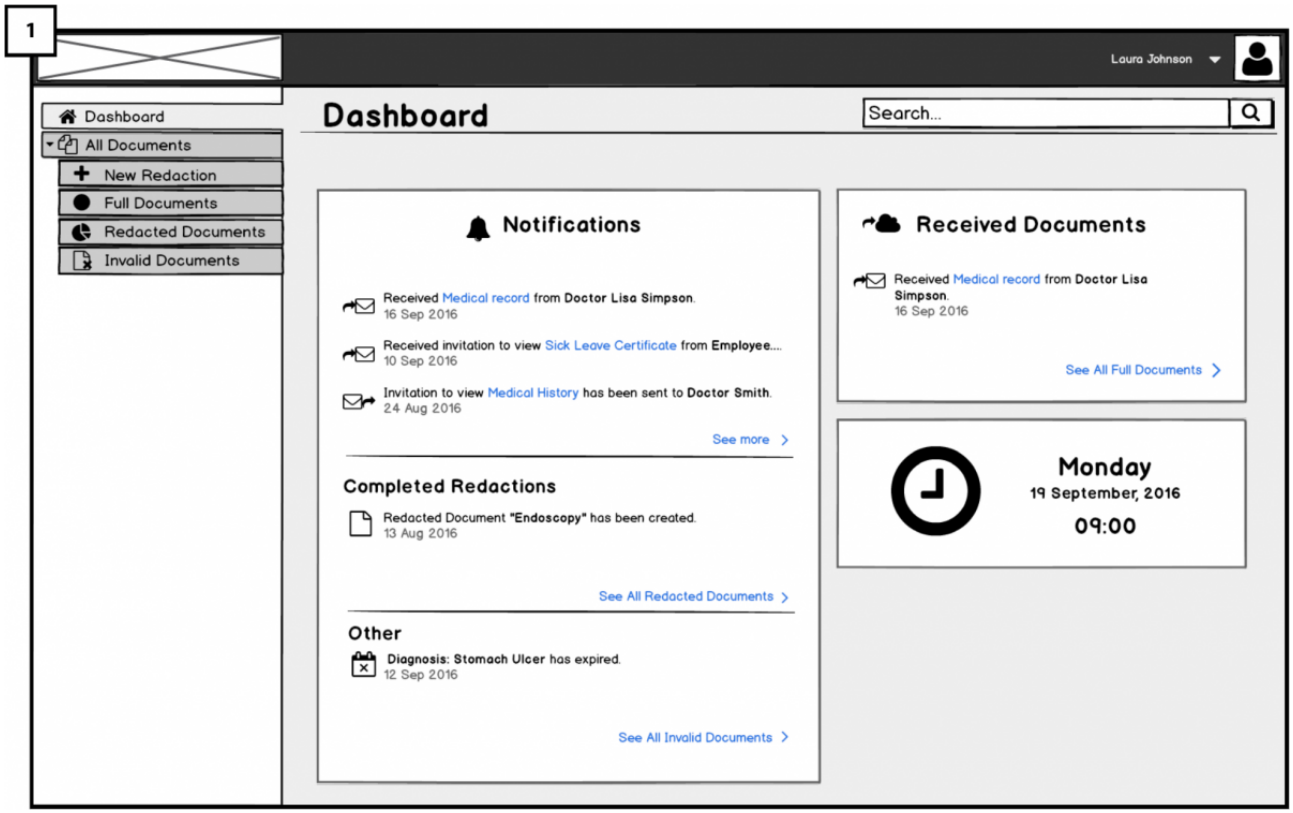


Figure 9. A redaction template metaphor shows the recommended amount of information to redact.

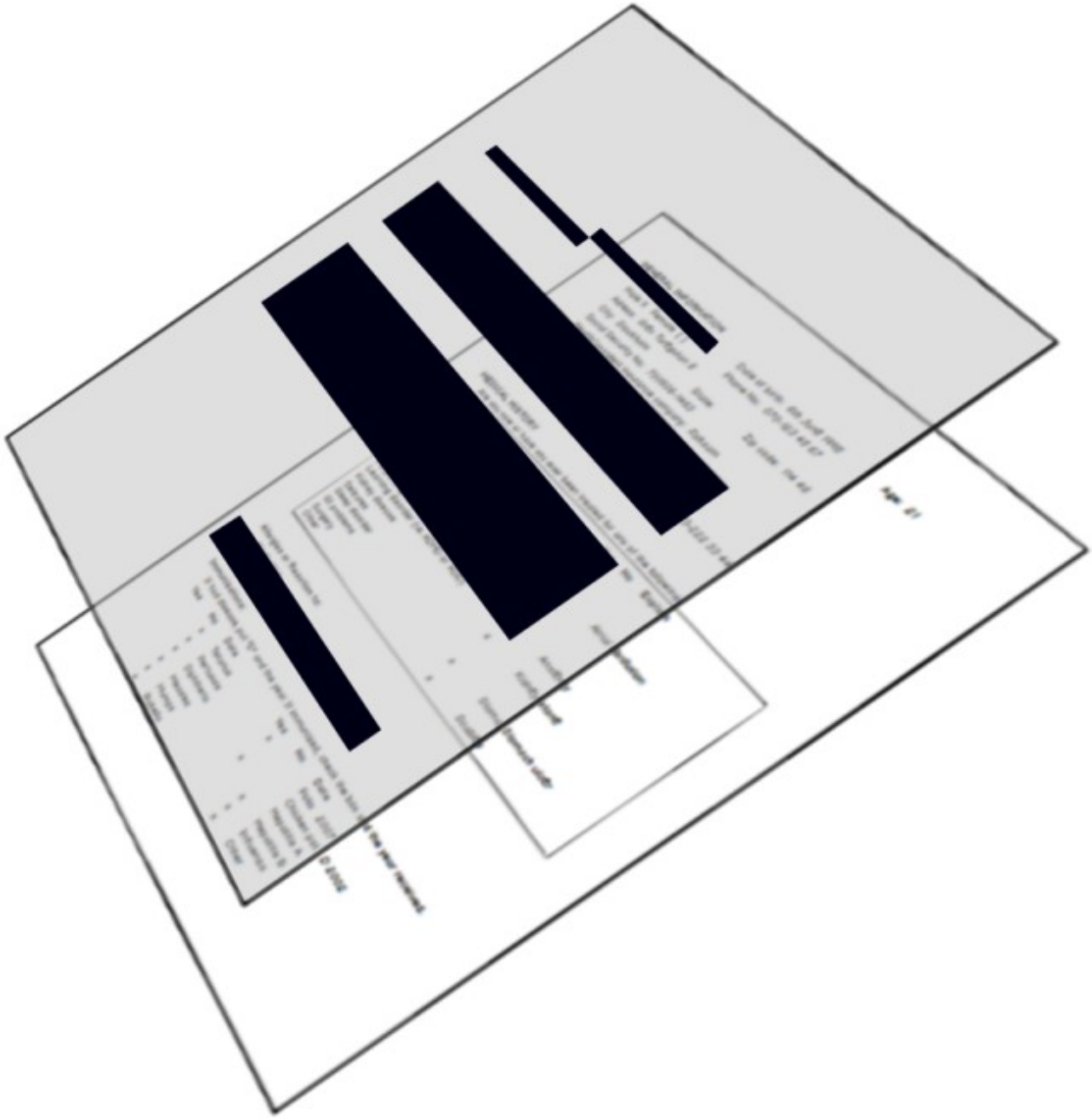


Figure 10. Templates for redaction in Cloud portal.

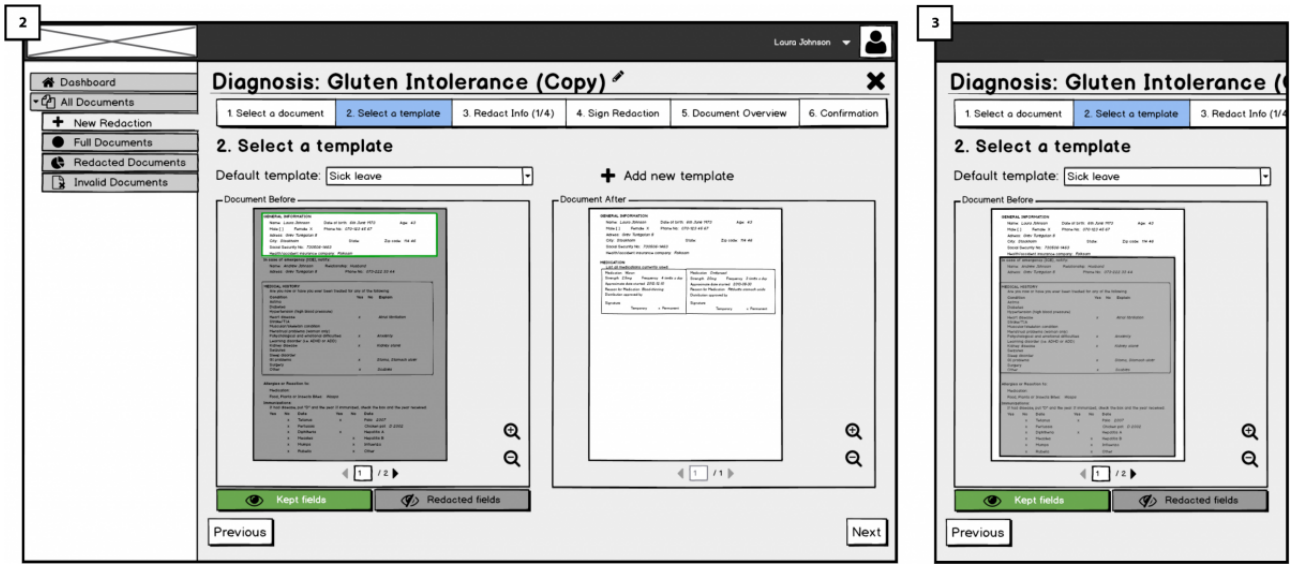
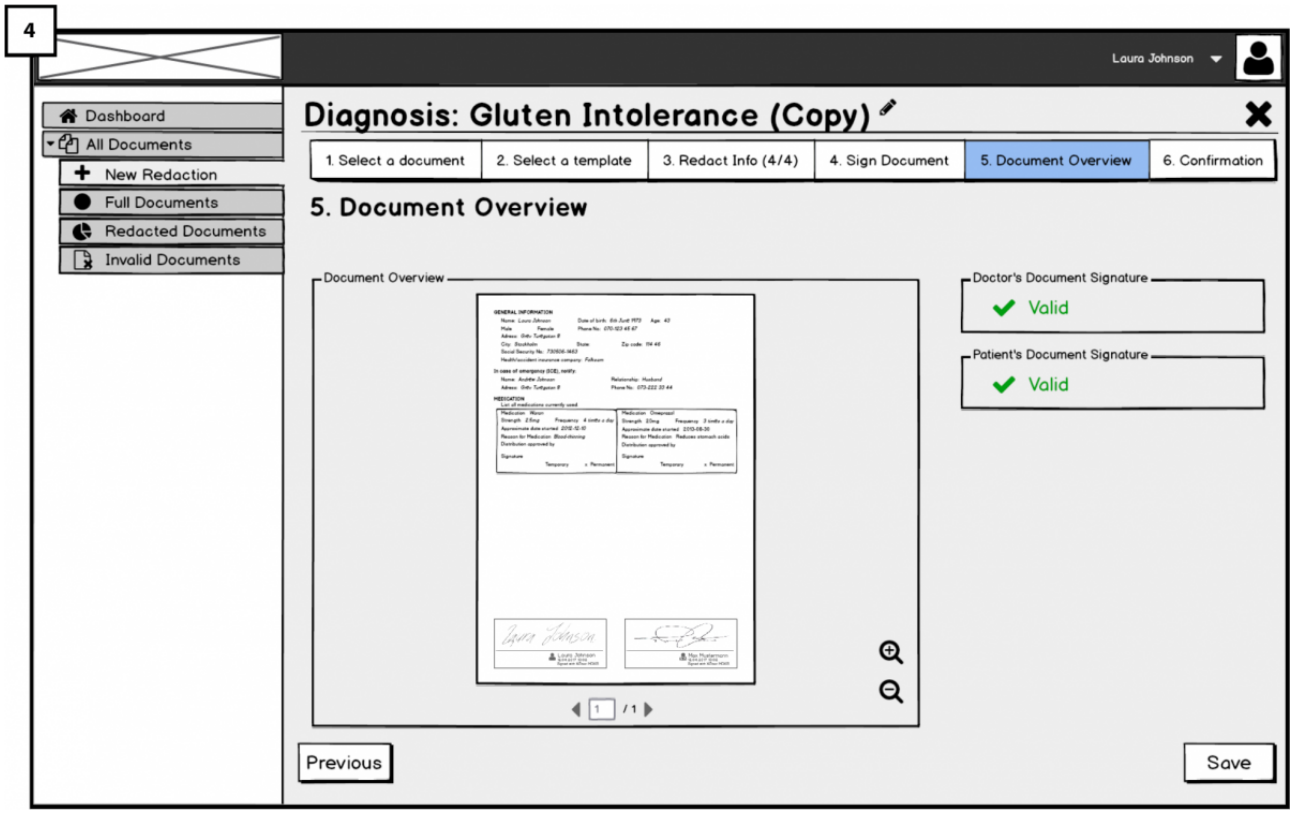


Figure 11. Validity and view of the signature.



Results

Individual Walk-Throughs or Interviews: Medical Staff Perspectives on Signing a Redactable Electronic Health Record

In total, there were 13 interviews, 5 were interviewed in Sweden, Värmland (S1-S5) and 8 in Germany, Frankfurt (G1-G3) and Hamburg (G4-G8). As shown in Table 2, eight participants are doctors in different fields, 2 foundation doctors, 1 nurse, 1 medical secretary, and 1 retired dentist (G4). Most had more than 8 years of experience in their field.

Table 2. Overview of medical staff's working titles and experience.

Index	Working title	Working experience
S1	Medical secretary or care administrator	30 years
S2	Foundation doctor at emergency section	Less than 2 years
S3	Nurse	2 years
S4	Doctor in pathology	20+ years
S5	Foundation doctor in general medicine	Less than 2 years
G1	Dermatologist	20+ years
G2	Doctor and director of cardiology	20+ years
G3	Pediatrician part time psychiatrist	22+ years
G4	Retired dentist	20+ years
G5	Doctor with quality assurance responsibility	20+ years
G6	Doctor of medicine	8 years
G7	General practitioner	25 years
G8	Rehabilitation medical doctor	30+ years

The following sections are the main results from the discussions.

Perspectives on Redactable Signatures in Electronic Health Record

Initially, the need for seeing possible redaction templates (as seen in Figure 4), which are made available to the patients as signers for future redaction, was unclear to many participants. However, after a short explanation, most acknowledged that viewing them was important and pointed out concerns regarding the redaction process and the need to consider the following aspects of redaction:

Redaction Rules' Specifications

Many participants mentioned general concerns regarding medical staff having incomplete EHRs, that is, redacted EHR. For instance, G7 emphasized the need for the full document by the recipient and would prefer a discussion with patients before any redactions; however, admitted to lack the time for doing so. Possible misuse scenarios of patients redacting medications to get more drug prescriptions from other doctors were mentioned by S3 and S5.

Many pointed out that there is a need for suitable redaction rules for the patients for conducting their redactions, which restrict the amount of redactable information for maintaining the credibility of the EHR. Suggested rule specifications included that no modifications (beyond redactions) of the EHR should be allowed (S1), the system should be *trusted* (G4 and G6), and that the patient should be the only redactor (G8). S2 and S3 mentioned that redaction rules must be strict, and that in some cases, redactions must not be allowed, for example, in the case of a pilot's medical certificate for heart diseases. The UI should communicate details of the redaction rules to both the doctors and the patients that will follow in addition to the templates.

G2 and G3 expressed strong objection toward patients redacting their medical documents and toward allowing patients to have access control of their EHR. They expressed their distrust in patients' knowledge, expertise, and ability to perform redactions and their distrust of the redacted documents. S4 showed a similar concern regarding the patient's limited knowledge. However, it was regarding patients revealing too much information; while redacting their documents, patients might keep fields that may result in indirect disclosure of sensitive information. Guidance for different types of users should be made clear and minimum data disclosure of the redaction rules should be communicated to the patients.

From these interview results, we can refine our requirements with regard to redaction rules and guidance:

- No arbitrary redactions of EHRs by patients should be allowed.
- Redaction rules are predefined by redaction templates, considering both data minimization and patient safety, in dependence of the type of recipient and purpose of selective authentic EHR exchange.
- Doctors should be able to further fine-tune and restrict the rules for redactions that are made possible to the patients via the templates, that is, the doctors keep the final control of what information is made redactable by the patients via the SAE-service, which they can also discuss and set up in cooperation with their patients.
- Redactions by the patient are restricted by clearly communicated redaction rules, which are given by the templates with possible further restrictions by the doctors.

Clear Responsibilities and Accountability

It was noteworthy that all medical staff members have concerns regarding the accountability of the redacted EHR. S4 pointed out repercussions to the doctor as the signer of the redactable EHR and mentioned that the signer might be ''sued'' in some countries for misinterpreted signed redacted document by the patient. Others (S4, S5, and G1) noted that putting trust into an SAE-service would depend on showing that an EHR was redacted and that the redactor is accountable.

Hence, these interview statements helped us confirm requirement RQ4. We derive the following as requirements that:

- Redactions should be enforced by a keyed-operation (ie, confirmed by the patient by a signing operation).
- Redacted EHRs by patients should be clearly communicated to the recipients as having been redacted by the patient, that is, the patient should be made accountable for any redaction. This should be achieved by prominently showing the electronically verifiable signature of the patient for a redacted EHR.

Usability of the Authentication and Signing Processes

The Swedish Data Protection Authority *Datainspektionen* has clearly stated, as a rule of thumb, that at least a two-factor authentication mechanism for the processing of sensitive personal data, including medical data, should be used. Although this can directly be derived from the requirement of enforcing appropriate means of security for the processing of personal data pursuant to EU Data Protection Legislation (Art 17 of the EU Data Protection Directive and Art 5 I (f) of the GDPR), our interviews revealed that many clinics in Germany still only use a simple password protection for authentication.

As the redactable signing operation requires a secure authentication of the signer, we also interviewed the medical staff with regard to their perceptions of the 2-factor authentication mechanism of XiTrust's MOXIS for system log-in and for the signing process (see Figure 2).

Efficient Authentication Process

The busy nature of the medical staffs' working environment requires efficiency for completing their tasks. S5, G2, and G6 raised concerns about the time and efforts consumed, the number of clicks needed to sign-in, and using the mobile phone in our use case, especially when timeouts occur. As S5 stated, "Every micro second counts." Moreover, 3 participants (S1, S4, and G5), who were familiar with similar 2-factor authentication mechanisms, expressed concerns using SMS for routine work where efficiency is important and thought that it was cumbersome for every instance of signing documents to go through the 2-factor authentication process.

In our studies, participants indicated signing mistakes happening when users were able to sign a document by just a mere click on a sign button. The solution, according to MOXIS developers XiTrust, is intended to eliminate redundant authentication when signing documents. Documents will still be viewed individually, and when reaching the signing process, medical staff can send the approved documents to a *tray* (for bulk signing) where later, they are able to do a group signature by following the authentication process. In this way, they can review documents once again or even have the documents sent to them by other staff (eg, secretaries) to sign. Signing mistakes that have been observed usually happen in the first stage *viewing the document*, and therefore, it is very unlikely in this way of bulk signing, with the extra step to review in the *tray*.

Hence, a multifactor authentication method to be used for secure authentication of the signer should provide efficiency in terms of minimizing the numbers of mouse-clicks required. This could, for instance, be achieved by simply saving username fields and by providing the option to sign a group of EHRs rather than requiring an electronic signature operation for each single EHR.

Practically Usable Security

When it comes to signing in, it is clear that medical staff appreciated the added layer of security (2-factor authentication with transaction authentication number(TAN)and short message service (SMS) in comparison with username and password. S1, S4, and G5 were familiar with this 2-factor authentication methods from eBanking apps and had, therefore, no problem in using it.

In addition, some suggested adding 2-factor authentication procedure before uploading medical documents to the Cloud or configuring an extra authentication step. There were, however, practical security concern regarding the use of the SMS and mobile phones for the 2-factor authentication function as expressed by 6 participants (S2, S4, G1, G4, G5, and G7). According to policies of Swedish hospitals, it is not permitted to use personal mobile phones for work purposes; instead, each medical staff is provided with a work phone that is not connected to external networks for security reasons. In addition, some doctors in hospitals in Germany have similar workplace policies for not using smartphones. G5 and G7 mentioned not using it even for personal purposes.

Therefore, requirements for authentication method for the signing operations are as follows:

- The use of commonly known secure authentication solutions that most users are familiar with should be offered (such as Bank-ID in Sweden).
- The UI should offer alternatives for different multifactor authentication methods that do not all require a mobile phone.

Human Signing-Error Support

It was reported by some participants (S1 and S4), who already use some signing functionality in their existing systems (however noncryptographic), that mistakes do occur when signing the EHR. Some examples include the hastened clicking on the sign button, especially when multiple parties are involved and discovered errors in the EHR record.

Although our use case requires more steps from the user than a hasty click to sign (authentication process), additional support for medical staff when mistakes occur during the signing of medical documents process is needed. Hence, the functionality of unsigning, that is, revoking a signature of an EHR should be added for mitigating hasty signing actions and for correcting errors.

Usability of the Signature Representation

The icon of a *seal* that corresponds to the signed EHR in the mock-ups was clear for most participants; however, S4 noted that a *tick* is more suitable and closer to the real-world analogy.

Most participants were not familiar with digital signatures; therefore, the visual representation of the digital signature was appreciated by them (Figure 7). However, S4, G3, and G8 were

familiar with digital signatures. They stated that the visual presentation was not needed and might be ‘’misleading’’ to be the actual digital signature, and therefore, not trusted.

Another concern regarding the visual representation of signatures arises from the case of having multiple parties involved in the signing process: Either multiple doctors or a combination of doctors and medical secretaries, that is, who is to sign first and whose signature is supposed to be there: secretary’s or doctor’s (S4 and G2). Concerns about privacy protection of the doctor were discussed with G5, showing the signature of the doctor on the redacted document is typically revealing the identity of the doctor to the recipient and possibly the doctor may, therefore, be mistaken to have responsibility of the redacted document.

We conclude the following:

- The responsible doctor should add the redactable signature. If the medical secretary should first sign the EHR, this signature could be implemented by noncryptographic means or could later be replaced by the doctor’s redactable signature.
- The roles of the signatures by doctors and patients (as the redactors of EHRs) and the responsibilities of these 2 parties should be made clear by the UI.

General Acceptance Criteria

While in Sweden, all EHR are stored electronically available, in Germany, they are mostly stored on paper and not digitalized. Some doctors in Germany (G7 and G8) were hesitating to store very sensitive medical attributes (eg, related to psychiatric diagnoses) electronically or even to upload them to a Cloud platform, as they think that all systems could be hacked, as G8 said, “Hackers are often ahead of things.”

Finally, when asked if they would use the SAE-service to sign a medical document with a redactable signature, S1, S2, and G4 agreed to sign a redactable medical document without further comments. Many participants said yes on the condition of accountability of the patient is clear and redaction is shown (S4, S5, G1, and G5), trusting the system (G6), stricter redaction rules to avoid abuse of drugs (S3), if it is used only for nonmedical uses (G2), and not for all kinds of patients (S5). As mentioned above, S5 and S3 were concerned about drug misuse, for example, a patient hides information about misuse or overconsumption of certain medications (eg, morphine) to get prescriptions from another doctor and suggested some patients to have blocked fields of redaction. G3 was the only participant that stated that he would not sign a redactable medical document as he does not trust the patient’s expertise to redact a document, and therefore, would not trust a redacted document either.

Our interviews showed that acceptance criteria could mostly be met by the refined requirements listed above for clear redaction rules that can be influenced by the doctors and by keeping the patients clearly responsible and accountable. Furthermore, to address any security concerns raised, doctors should have the option to exclude very sensitive fields from the EHR to be signed with their redactable signature and then uploaded to the Cloud platform.

Group Walk-Throughs or Focus Groups: Patient Perspectives on Redacting Their Electronic Health Record

For addressing patients’ perspectives, we held 5 FGs with a total of 32 participants (Table 3). Out of this, 2 took place in Sweden, 2 in Germany, and 1 in Oslo (at a seminar for information technology [IT] security PhD students in Norway and Sweden).

When recruiting participants, they were asked about their knowledge of digital signatures and redactable (malleable) signatures. Those with none were put in the lay users groups (FG1 and FG3). Those who knew of redactable (malleable) signatures were excluded from the study as our aim was to test the first-hand experience of redactable signatures and test their first thoughts, opinions, and trust they had in the validity of redactable signatures. FG2 consisted of technical users in computing science with the knowledge and experience of digital signatures, whereas FG4 had lay users in executive positions in the industry with knowledge of digital signatures. The fifth group (FG5) consisted of technical experts in the privacy and security field with knowledge and experience of digital signatures (but no knowledge about redactable signatures). The following sections are the main results from the discussions.

Users Perspectives on Information Privacy and Sharing

All FGs sessions started with an exercise of redacting personal information fields for different personas on papers (which were assigned and handed to participants). Results of paper redaction exercise are described in Multimedia Appendix 3. Overview results of blacking-out of sensitive data on paper. Almost all participants (30/32, 93%) blacked out information about medical issues. Subsequently, general information such as hobbies, demographics, and address were at the bottom of the chart, where only a few participants (5/32, 15%) redacted them across all groups. There was a clear consensus among participants of all FGs, with no visible cultural differences, regarding the sensitivity of sharing their medical information with fellow FG participants.

Table 3. Overview of focus groups participants in each group.

Index	Type of users	Number of participants	Location
FG1	Lay users	6	Sweden
FG2	Technical users	7	Sweden
FG3	Lay users	6	Germany
FG4	Lay[a] users	6	Germany
FG5	Technical[b] users	7	Norway

[a]Initially meant to be technical users; however, during the discussion, it was clear that they did not know how digital signatures technically work.

[b]These are security and crypto researchers; their technical expertise goes beyond the other technical users.

The FGs discussions revealed variety of opinions and preferences regarding which information is considered private. Only 1 participant in FG1 expressed not minding sharing all information on the persona's card, because of personal openness, ideals, and personality trait. In FG3, few participants indicated that the information they would not want to share are reflected by their cultural norms. However, the majority in every group expressed hesitations for sharing most information, especially medical information. In FG2 and FG5, many participants thought they would share more of the information than they already shared voluntarily if they were asked for it, as they do not consider the information to be confidential. In all groups, a prominent factor for sharing information is the context; a couple of participants in FG5 said that they would share different information in medical versus employment environments.

Participants mentioned different contributing factors to sharing more information such as the social environment, depending on persons asking for the information, discussion theme, social norms of the group, society and cultural influences, peer pressure, and social appeal.

As previous research results already showed for the general case [28,29], our evaluations also confirmed that people can be divided into different privacy personas: they have different preferences with regard to withholding personal information, also in dependence on the context of this study.

Hence, we can conclude:

- Different redaction templates offering default redactions should be offered in dependence of the context and type of recipient of the redacted document.
- The UI should motivate the design of the redaction templates for enforcing data minimization by default and protecting patient safety to different types of users.

Standardization for Use: Signing-In With Two-Factor Authentication

Most FGs indicated that although the 2-factor authentication is a good idea for security purposes, it is still not clear why they need to use a mobile phone to do so. However, some participants in FG1 and FG2 noted that using 2 different devices is important for a secure signing in as it provides a second secure channel. An alternative suggested by participants in FG1 is to use standard services, that is, the Swedish BANK-ID, which they already use for many other apps. The Bank-ID service is available in the form of a soft certificate that does not require an extra (mobile) device.

These results indicate the need for standardization, requiring to follow a standard format for trust and usability, preferably aligning with existing services or tools by trusted parties. The UI should tailor to secure standard authentication solutions that are commonly used in the corresponding country such as BanK-ID in Sweden.

Redaction Rules and Accountability

When discussing redactions in the FGs, the functionality of selective authentic disclosure via redactions was generally appreciated, although there was a clear concern regarding who is responsible for setting up the templates and defining rules of redactions. One participant in FG2 suggested that it should be the recipient as they need to confirm what information they need from the redactor. However, others disagreed based on not trusting the recipient enough to ask for the minimal amount of information (eg, insurance companies might be interested in receiving more information than needed).

The stencil metaphor used for graying out the parts to be redacted was, in general, well understood. However, some participants in FG4 showed concerns in trusting the redaction thinking that the redacted information will still be accessible in a hidden technical manner. It was stated that it is mainly because of their general distrust in programmers that they additionally acknowledged their lack of technical background and knowledge of the system's processes. Many participants in FG3 and FG4 showed concerns regarding redacted documents as doctors might be still able to acquire the redacted information by other unknown means without the consent of the patient. These concerns were not raised in the FGs with Swedish participants.

Hence, we conclude the following:

- Templates defining redaction rules need to be defined and/or certified by trustworthy actors that are competent to define what information is required considering the data minimization principle and patient safety.
- The fact that redacted data are actually deleted, not simply hidden and unavailable for the recipient, should be clearly communicated by the UI for establishing trust.

Redaction and Templates Guidance for Different Types of Users

The stencil metaphor for selective disclosure by blacking out (or in our case, graying out) text to be redacted corresponds to practices in the real world and was well understood by all FG participants. They also acknowledged the desire to disclose selective medical information (redacting EHRs) via redactable signatures. Participants noted that indications showing deviation from the templates were missing from the mock-ups and should be clearly shown in the UI as well as notifications when the signatures become invalid (redaction rules broken) in the process of redacting too much. For example, a warning that their current manual selection for redacting that field is invalidating the doctor's signature so that they reconsider their actions. The option to select a template for redaction and automatically get a redacted document was mentioned a few times by participants in FG2, where they do not want to further do any redactions and want to shortcut through the process. One participant in FG4 indicated that he or she do not want to choose the template, instead would rather have the system automatically assigning the template based on the recipient.

Users with different background knowledge and experience shared more or less concerns regarding the difficulty of knowing what to include and how much to redact without redacting too much for different recipients. The need for context-dependent templates for enforcing privacy by default while considering patient safety was confirmed. Therefore, the UI should offer templates based on some default recipients for more guidance. Moreover, it has to allow individual adaption of redaction templates. The UI should also show warnings and error messages for redaction rules and diversion from the templates. For users who prefer not to do manual redactions, some default steps with quick redactions following a template selection should be available and incorporated into the UI.

Hence, as concluded above:

- Redactions should be guided by the default templates that the UI should offer in dependence of the recipient.
- In the process of selecting fields to be redacted: when the patient selects fields beyond the permitted redactions, the user should be clearly informed that the doctor's signature will become invalid.

Trust of the Signature's Validity

For the UI showing the doctor's document signature as *valid* after the redaction done by the user (Figure 11), we asked all FGs whether they would trust the validity of the doctor's signature. Evoking the correct mental models for the authenticity of the redacted EHRs and particularly mediating trust in the validity of the redactable signature by the doctor after the redaction had taken place worked for lay users. FGs with lay users (FG1 and FG3) stated that they would have no issues with trusting the signature, participants from FG2 (technical users) were directly questioning the validity of the signature. Some were stating that after changing the text, the signature should be invalid, and 1 participant speculated whether the Cloud portal would create a new signature. In contrast to that, FG5 participants, consisting of experts in privacy and security, were not questioning the validity of the doctor's signature. We then directly asked them whether they would still think that the signature was valid although the text was changed. Moreover, 1 participant explained that he or she was assuming all redactable fields in the document were signed separately so that those fields with the respective signatures could easily be redacted or deleted without invalidating the validity of the other signed fields, that is, they could see some plausible technical solutions for the validity of the signature.

The FGs showed that depending on the technical knowledge, users might trust or distrust the validity of signatures of redacted documents. Typically, users with some technical knowledge may question the validity, whereas lay users may trust the validity of the signatures and security experts may find technical explanations. Hence, technical and nontechnical users may have different degrees of trust on the validity of the doctor's signature after redactions.

Therefore, the UI should offer different levels of guidance addressing redactable signatures corresponding to user expertise, that is, introductions or tutorials also have to address technical users and their potential misunderstandings of redactable signatures and the UI should offer tooltip information or a link to explanations what *validity* means for a redactable signature.

Branding and Trust

Earlier work has discussed precautions and concerns regarding storing EHR in the Cloud [30] where security, privacy, and trust requirements were stressed.

Concerning branding and trust in the system, some participants from FG2 and FG4 indicated that the UI should clearly indicate which organization is involved. Meaning, which would operate the SAE-service, including the Cloud hospital platform, for example, whether it is the hospital or municipality hosting it and operating a private Cloud. Some participants in FG3 indicated that they would not trust *new technology* that is, the Cloud portal in general; however, many indicated they would trust the governmental authorities and branding of such would be a factor for trusting the system. Some of FG1 and FG2 participants even indicated that they would only trust the authorities in Sweden, whereas others preferred to have options and alternatives. Inversely, 1 participant in FG1 stated that competent privacy or IT security companies, which would often have more skilled personnel than government agencies, are trusted rather than governmental authorities. Nonetheless, FG3 and FG4 participants indicated that there seems to be less trust in the government among the population in Germany.

Hence, based on the above mentioned statement, most participants, especially those from Sweden, seem to trust the government as an operator, some, however, would rather trust competent private IT security companies.

We can conclude overall as follows:

- A trustworthy and creditable agency is needed to brand the SAE-service and to host a private Cloud, considering culturally influenced social trust factors.

- The UI should communicate the trusted party's branding and privacy or trust certification seals that may be helpful for establishing reliable trust.

Discussion

Summary

Main findings of our user studies can be summarized as follows:

In our studies, medical staff's perspectives on redactable EHRs included concerns regarding redactions resulting in incomplete EHRs used for medical purposes and acknowledgments of patients having control of their own EHRs. Overall, they accepted the Cloud-based SAE-service under the premises of some conditions such as clear responsibilities and accountability of the redactor (patient) as well as rules defining redaction rules. The security of the 2-factor authentication was appreciated, however, required more usable and efficient means for the authentication and signing processes in a hospital environment. In addition, the hand-written signature representation of the digital redactable signature was overall appreciated.

The overall expressed opinions of the FG participants concluded that medical information is most sensitive among other types of personal information such as age, address, and income. Participants well understood the redaction process and the stencil metaphor of *blacking or graying out* of fields in the redaction process. However, they highlighted the need for more support and guidance with regard to redaction templates, for example, different default templates serving different purposes as well as the support for redaction rules. The 2-factor authentication was well received by participants who were familiar with similar apps; thus, standardization with existing solutions was highlighted. Acceptability and trust of the validity of redactable signatures depended on the familiarity and technical experiences of digital signatures. In particular, technically knowledgeable users (apart from crypto-specialists) had had more issues trusting the malleable signatures. In addition, in terms of trust or distrust in the SAE service or in agencies hosting the service, differences existed between FGs in Sweden (higher trust) and Germany (showing distrust).

Comparison With Previous Work

Addressing privacy concerns and the users' trust regarding storing their medical data in the Cloud is essential [31-33].

Related work has focused on technical means for addressing privacy and security challenges in EHR systems [34-36]. Studies have addressed patient control in terms of Web-based access to their EHRs for increasing transparency in Sweden [37,38], a dynamic, in terms of defining access control requirements for patients [36], and through a dynamic consent model [39]. Moreover, the acceptability and end-user challenges of personally controlled eHealth records [40,41] have been discussed. However, to the best of our knowledge, no previous work has addressed the end-user perspectives for a Cloud-based SAE-service based on redactable signatures as a means for enhancing patient control over authentic medical data. Consequently, this study is also the first to report about end-user perspectives and requirements for making such service usable, acknowledged, accepted, and trusted by different types of end users.

End-user perspectives for cryptographic selective disclosure technologies, especially in terms of the challenges to evoke comprehensive mental models and to establish end-user trust in the stated selective disclosure functionality have been discussed for attribute-based credentials (ABCs) [15,42] and for the German National identity card [43]. Both ABCs and the German identity card are credentials that provide related user-controlled data minimization functions, which allow the credential holder to selectively disclose attributes or characteristics of those attributes stored on the credentials (eg, they allow to reveal whether a credential holder is over 18 years instead of revealing the exact birth date or any other information stored on the credential). The studies by Wästlund et al and Benenson et al [15,42] explored different ways in which suitable mental models of the data minimization property of ABCs can be evoked on end users. The results showed that while an *adapted card metaphor* helped more than half of the test users to understand that attributes could be selectively disclosed or hidden, nevertheless better design paradigms for understanding the selective disclosure property of attribute characteristics were still needed.

However, redactable signatures used in our use case allow only traditional redactions (ie, deletion of information) as a means for selective disclosure of the remaining information, for which more adequate real-world analogies (such as the *blacking out* metaphor) exist. Blacking out text on paper (including text in letters with signatures) has been long practiced already in the offline world. This may be 1 reason why in this study, the stencil metaphor for blacking out information (illustrated by graying out text in the mockups) worked well for most of our FG participants to understand the selective disclosure property of redactable signatures.

Evoking the correct mental models with our mock-ups for the authenticity of the redacted EHRs and particularly mediating trust in the validity of the redactable signature by the doctor after the redaction had taken place seemed to work well for the lay users. However, technical users that were familiar with traditional digital signatures, which are invalidated by any modifications of the signed text, had doubts in the validity of the signature after the redaction represented by the green *check* icon. These findings are similar to research findings in Lerner et al's study [44], which report that users with technical security knowledge lacked trust in a newly designed email encryption tools, where cryptographic operations were automatic and hidden, and which, thus, seemed to behave differently to the traditional email encryption tool GNU Privacy Guard (GPG) [45] that they were familiar with. The findings of our study, however, also showed that in the case of redactable signatures, our technical users with advanced expertise in cryptography were also able to find their own technical explanations. Therefore, they were able to establish trust into the validity of the signature by the doctor after the redaction.

Comparison: Sweden and Germany

Previous studies show that Germany is facing multiple obstacles that prevent the implementation of a national EHR system.

Apart from the technical complexity and compatibility issues, there is also a conservatism [20] and strong resistance among health professional organizations against digitalization of health records [22,23]. Moreover, our study showed similar concerns by some of the medical professionals in Germany, who expressed distrust in data security in general or in the patients' knowledge, expertise, and ability to perform redactions and would, therefore, not trust the redacted documents. The interviewed medical professionals in Sweden did not voice such concerns to that extent.

A higher trust in government agencies for hosting the SAE-service by FG participants in Sweden in comparison with the ones in Germany reflected the Eurobarometer survey results [24], which revealed that people in Sweden are most likely to trust their national public authorities.

Measures providing transparency for the processing of EHRs and accountability of health personnel that are in place in Värmland, Sweden, may also be 1 reason why our FG participants in Sweden, in contrast to the participants in FGs 3 and 4 in Germany, did not voice any doubts that doctors may still be able to obtain the redacted information by other means without the patients' consent.

In an interview with medical professionals in Sweden, S2 confirmed that accountability measures are taken very seriously in Sweden with the example that doctors require to document the patients' consent witnessed by a colleague when accessing the patients' EHRs.

Legal Rules and Compliance

The proposed SAE-service is compliant with European privacy rules and regulations. It is, in particular, meeting the GDPR's privacy requirements for data minimization (Art 5 c), and for data protection by design and default (Art 25). It also supports patients to exercise their right to access their medical data by obtaining an electronic copy of their medical records pursuant to Art 15 (3) GDPR and pursuant to national patient data protection legislation (eg, Section 630g (2) Bürgerliches Gesetzbuch in Germany [46] and Section 8 Patient Data Act in Sweden [47]) complementing the GDPR.

Already today, patients would be empowered to redact data from these electronic copies of their medical records (if they are not digitally signed) or from printouts of these copies and pass it on to other parties. The proposed SAE will, in addition, protect the authenticity and integrity of the medial information, and thus, also patient safety, through the (redactable) signature by the doctor. In addition, we require that the patient is digitally signing all redactions (ie, the redaction is implemented as a keyed operation) for making the patient accountable. The secure authentication and signing solution MOXIS by XiTrust that is used for the SAE-service allows to implement the patient's signature as a qualified signature according to the European Electronic Identification, Authentication, and Trust Services (eIDAS) Regulation with the help of the Austrian trust service provider A-Trust [48]. This means that the patient's signature would, in this case, fulfill the highest security standards of eIDAS and have the same legal status as a handwritten signature.

Limitations

As our study's focus is on the user, our limitations are related to participants of our studies. One might argue that because of our recruitment process (our own network), participants might be inclined to be bias in their feedback. However, because of our research objectives and study design, we refrained from inquiring about their acceptability criteria (if they value our UIs), as that was not an interest for our research. Our focus was on mental perception of the SAE-service functions and sequences, that is, what works and why. Another point is our limited demographic data from our participants. We have had participants varying in background, gender, and age; however, we did not collect that data in our results. In addition, we intended to follow the data minimization principle in practice with our studies: collecting the minimum amount of data necessary for the study.

Conclusions

In our study, we have addressed medical professionals and patient's perspectives of our SAE-service. Allowing data minimization of the EHR through redactable signatures, supports users' control of their medical data. The need for diverse considerations for both roles of users, with different technical backgrounds as well as country they are based in, has been highlighted in our results. One important influence is the effect of users' experiences on their acceptance and perception of our proposed service that include their mental models and familiarity with existing solutions, experiences with EHRs, and/or their technical background. Therefore, it was challenging in our study to compare user's acceptance and trust based on countries they are in (Sweden and Germany) as there was a clear distinction in the familiarity and experience of EHRs in the countries addressed. The complexity of different users' experiences calls for customized designs targeting different sets of users for future usable eHealth solutions.

Acknowledgments

This study is supported by the EU Horizon 2020 research project No. 644962 PRISMACLOUD "Privacy and security maintaining services in the Cloud" [49]. The authors extend their thanks to John Sören Pettersson and Erik Wästlund for their feedback and discussions. In addition, to our project partners who aided our studies and to the volunteers who participated and contributed to this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Consent form for interviews.

[PNG File, 139KB - jmir_v20i12e10954_app1.png]

Multimedia Appendix 2

Consent form for focus groups.

[PNG File, 143KB - jmir_v20i12e10954_app2.png]

Multimedia Appendix 3

Overview results: blacking-out of data on paper.

[PNG File, 16KB - jmir_v20i12e10954_app3.png]

References

1. Bundesverfassungsgericht (German Constitutional Court). 1983. BVerfGE 65, 1 - Census [in German] URL: http://www.servat.unibe.ch/dfr/bv065001.html [accessed 2018-09-20] [WebCite Cache ID 72Zg8j5gg]
2. Rothstein M. The Hippocratic bargain and health information technology. J Law Med Ethics 2010;38(1):7-13 [FREE Full text] [doi: 10.1111/j.1748-720X.2010.00460.x] [Medline: 20446978]
3. Miles SH. The Hippocratic Oath and the Ethics of Medicine. USA: Oxford University Press; 2005.
4. European Union Regulation. EUR-Lex. 2016 May 04. 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) URL: http://data.europa.eu/eli/reg/2016/679/oj/eng [accessed 2018-11-09] [WebCite Cache ID 72ZhV094w]
5. Derler D, Pöhls HC, Samelin K, Slamanig D. A general framework for redactable signatures and new constructions. In: Information Security and Cryptology - ICISC 2015. Cham: Springer; 2015:3-19.
6. Urowitz S, Wiljer D, Apatu E, Eysenbach G, Delenardo C, Harth T, et al. Is Canada ready for patient accessible electronic health records? A national scan. BMC Med Inform Decis Mak 2008 Jul 24;8:33 [FREE Full text] [doi: 10.1186/1472-6947-8-33] [Medline: 18652695]
7. 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] [Medline: 16357345]
8. Wiljer D, Urowitz S, Apatu E, DeLenardo C, Eysenbach G, Harth T, Canadian Committee for Patient Accessible Health Records. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. J Med Internet Res 2008 Oct 31;10(4):e34 [FREE Full text] [doi: 10.2196/jmir.1061] [Medline: 18974036]
9. EU Commission. Directorate-General for Communication. Special Eurobarometer 460, Attitudes towards the impact of digitisation and automation on daily life URL: http://data.europa.eu/euodp/en/data/dataset/S2160_87_1_460_ENG [WebCite Cache ID 72Y3FE3kL]
10. Weitzman E, Kaci L, Mandl K. Sharing medical data for health research: the early personal health record experience. J Med Internet Res 2010 May 25;12(2):e14 [FREE Full text] [doi: 10.2196/jmir.1356] [Medline: 20501431]
11. Olson J, Grudin J, Horvitz E. A Study of Preferences for Sharing and Privacy. In: CHI '05 Extended Abstracts on Human Factors in Computing Systems. USA: ACM; 2005 Presented at: CHI EA '05; 2005,04, 02-07; Portland, USA p. 1985-1988. [doi: 10.1145/1056808.1057073]
12. Abdelhamid M, Gaia J, Sanders G. Putting the focus back on the patient: how privacy concerns affect personal health information sharing intentions. J Med Internet Res 2017 Sep 13;19(9):e169 [FREE Full text] [doi: 10.2196/jmir.6877] [Medline: 28903895]
13. Kisekka V, Giboney J. The effectiveness of health care information technologies: evaluation of trust, security beliefs, and privacy as determinants of health care outcomes. J Med Internet Res 2018 Apr 11;20(4):e107 [FREE Full text] [doi: 10.2196/jmir.9014] [Medline: 29643052]
14. Caine K, Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013 Jan 01;20(1):7-15 [FREE Full text] [doi: 10.1136/amiajnl-2012-001023] [Medline: 23184192]
15. Wästlund E, Angulo J, Fischer-Hübner S. Evoking comprehensive mental models of anonymous credentials. In: Camenisch J, Kesdogan D, editors. Open Problems in Network Security, Lecture Notes in Computer Science. Berlin, Heidelberg: Springer; 2012:1-14.
16. Tsormpatzoudi P, Berendt B, Coudert F. Privacy by Design: From Research and Policy to Practice – the Challenge of Multi-disciplinarity. In: Berendt B, Engel T, Ikonomou D, Le Métayer D, Schiffner S, editors. Privacy Technologies and Policy. Switzerland: Springer International Publishing; 2015:199-212.
17. Cavoukian A. Privacy by Design: Take the Challenge. Canada: Information and Privacy Commissioner of Ontario; 2009.

18. Norman DA, Draper SW. User centered system design: New perspectives on human-computer interaction. USA: CRC Press; 1986:1-4822.
19. Alaqra A, Fischer-Hübner S, Pettersson JS, Wästlund E. Stakeholders' perspectives on malleable signatures in a cloud-based eHealth scenario. In: Proceedings of the Tenth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2016). UK: Plymouth university; 2016 Presented at: International Symposium on Human Aspects of Information Security & Assurance (HAISA 2016); 2016,07,19-21; Frankfurt p. 220-230.
20. Currie WL, Seddon JJ. A cross-national analysis of eHealth in the European Union: some policy and research directions. Information & Management 2014 Sep;51(6):783-797. [doi: 10.1016/j.im.2014.04.004]
21. Ortega Egea JM, González MV, Menéndez MR. eHealth usage patterns of European general practitioners: a five-year (2002-2007) comparative study. Int J Med Inform 2010 Aug;79(8):539-553. [doi: 10.1016/j.ijmedinf.2010.05.003] [Medline: 20538514]
22. Fragidis L, Chatzoglou P. Development of Nationwide Electronic Health Record (NEHR): an international survey. Health Policy Technol 2017 Jun;6(2):124-133. [doi: 10.1016/j.hlpt.2017.04.004]
23. Hoerbst A, Kohl C, Knaup P, Ammenwerth E. Attitudes and behaviors related to the introduction of electronic health records among Austrian and German citizens. Int J Med Inform 2010 Feb;79(2):81-89. [doi: 10.1016/j.ijmedinf.2009.11.002] [Medline: 20031482]
24. EU Commission. Directorate General for Communication Networks Internet. Special Eurobarometer 431, Data protection URL: http://data.europa.eu/euodp/en/data/dataset/S2075_83_1_431_ENG [accessed 2018-03-26] [WebCite Cache ID 72Y3RQWVH]
25. Riksdagsförvaltningen. Riksdagen. Law (2003: 460) on ethics testing of human-related research [in German] URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sfs-2003-460 [accessed 2018-09-20] [WebCite Cache ID 72ZkgGpzV]
26. Balsamiq. Mockups by Balsamiq Studios LLC URL: https://balsamiq.com/ [accessed 2018-09-20] [WebCite Cache ID 72ZllQLhf]
27. XiTrust Secure Technologies. URL: https://www.xitrust.com/produkte/xitrust-moxis/ [accessed 2018-09-20] [WebCite Cache ID 72Zn3PMEU]
28. Westin AF. Social and political dimensions of privacy. J Social Issues 2003 Jun;59(2):431-453. [doi: 10.1111/1540-4560.00072]
29. Morton A, Sasse MA. Privacy is a process, not a PET: a theory for effective privacy practice. In: Proceedings of the 2012 New Security Paradigms Workshop. 2012 Presented at: NSPW '12; September 18-21, 2012; Bertinoro, Italy p. 87-104.
30. Rodrigues JP, de la Torre I, Fernández G, López-Coronado M. Analysis of the security and privacy requirements of cloud-based electronic health records systems. J Med Internet Res 2013 Aug 21;15(8):e186 [FREE Full text] [doi: 10.2196/jmir.2494] [Medline: 23965254]
31. Vodicka E, Mejilla R, Leveille S, Ralston J, Darer J, Delbanco T, et al. Online access to doctors' notes: patient concerns about privacy. J Med Internet Res 2013 Sep 26;15(9):e208 [FREE Full text] [doi: 10.2196/jmir.2670] [Medline: 24072335]
32. Carrión Señor I, Fernández-Alemán JL, Toval A. Are personal health records safe? A review of free web-accessible personal health record privacy policies. J Med Internet Res 2012 Aug 23;14(4):e114 [FREE Full text] [doi: 10.2196/jmir.1904] [Medline: 22917868]
33. Brumen B, Heričko M, Sevčnikar A, Završnik J, Hölbl M. Outsourcing medical data analyses: can technology overcome legal, privacy, and confidentiality issues? J Med Internet Res 2013;15(12):e283 [FREE Full text] [doi: 10.2196/jmir.2471]
34. Fernández-Alemán JL, Señor IC, Lozoya PA, Toval A. Security and privacy in electronic health records: a systematic literature review. J Biomed Inform 2013 Jun;46(3):541-562 [FREE Full text] [doi: 10.1016/j.jbi.2012.12.003] [Medline: 23305810]
35. Li M, Yu S, Zheng Y, Ren K, Lou W. Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. IEEE Trans Parallel Distrib Syst 2013 Jan;24(1):131-143. [doi: 10.1109/TPDS.2012.97]
36. Sahi MA, Abbas H, Saleem K, Yang X, Derhab A, Orgun MA, et al. Privacy preservation in e-Healthcare environments: state of the art and future directions. IEEE Access 2018;6:464-478. [doi: 10.1109/ACCESS.2017.2767561]
37. Grünloh C, Rexhepi H, Cajander Å, Åhlfeldt R, Myreteg G, Huvila I. Patient empowerment meets concerns for patients: a study of patient accessible electronic health records in Sweden. In: Studies in Health Technology and Informatics. Amsterdam: IOS Press; 2016.
38. Rexhepi H, Åhlfeldt RM, Cajander Å, Huvila I. Cancer patients' attitudes and experiences of online access to their electronic medical records: a qualitative study. Health Informatics J 2018 Jun;24(2):115-124. [doi: 10.1177/1460458216658778] [Medline: 27440056]
39. Williams H, Spencer K, Sanders C, Lund D, Whitley EA, Kaye J, et al. Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. JMIR Med Inform 2015 Jan 13;3(1):e3 [FREE Full text] [doi: 10.2196/medinform.3525] [Medline: 25586934]
40. Weitzman E, Kaci L, Mandl K. Acceptability of a personally controlled health record in a community-based setting: implications for policy and design. J Med Internet Res 2009 Apr 29;11(2):e14 [FREE Full text] [doi: 10.2196/jmir.1187] [Medline: 19403467]

41. Pearce C, Bainbridge M. A personally controlled electronic health record for Australia. J Am Med Inform Assoc 2014;21(4):707-713 [FREE Full text] [doi: 10.1136/amiajnl-2013-002068] [Medline: 24650635]
42. Benenson Z, Girard A, Krontiris I, Liagkou V, Rannenberg K, Stamatiou Y. User Acceptance of Privacy-ABCs: An Exploratory Study. In: Tryfonas T, Askoxylakis I, editors. Human Aspects of Information Security, Privacy, and Trust. Cham: Springer; 2014:375-386.
43. Harbach M, Fahl S, Rieger M, Smith M. On the acceptance of privacy-preserving authentication technology: the curious case of national identity cards. In: De Cristofaro E, Wright M, editors. Privacy Enhancing Technologies. Berlin, Heidelberg: Springer; 2013:245-264.
44. Lerner A, Zeng E, Roesner F. Confidante: Usable Encrypted Email: A Case Study with Lawyers and Journalists. 2017 Presented at: 2017 IEEE European Symposium on Security and Privacy (EuroS P); 2017,07,03; Paris, France p. 385-400. [doi: 10.1109/EuroSP.2017.41]
45. The GNU Privacy Guard Project. URL: https://www.gnupg.org/ [accessed 2018-09-18] [WebCite Cache ID 72WdY4575]
46. Bürgerliches Gesetzbuch: German Civil Code BGB. Civil Code in the version promulgated on 2 January 2002 (Federal Law Gazette [Bundesgesetzblatt] I page 42, 2909; 2003 I page 738), last amended by Article 4 para. 5 of the Act of 1 October 2013 (Federal Law Gazette I page 3719) URL: https://www.gesetze-im-internet.de/englisch_bgb/englisch_bgb.html [accessed 2018-09-19] [WebCite Cache ID 72YKD1veX]
47. Riksdagsförvaltningen. Patientdatalag (2008:355) Svensk författningssamling 2008:2008:355 t.o.m. SFS 2018:1214 - Riksdagen URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/patientdatalag-2008355_sfs-2008-355 [accessed 2018-09-20] [WebCite Cache ID 72Zo03kcJ]
48. EU Publications. 2014. Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC URL: https://publications.europa.eu/en/publication-detail/-/publication/23b61856-2e82-11e4-8c3c-01aa75ed71a1/language-en [accessed 2018-09-19] [WebCite Cache ID 72YHVfCrp]
49. Prismacloud. PRIvacy and Security MAintaining services in the CLOUD URL: https://prismacloud.eu/ [accessed 2018-09-20] [WebCite Cache ID 72ZoZtqVC]

Abbreviations

ABCs: Attribute-based credentials
eHealth: electronic health
EHR: electronic health record
eIDAS: electronic IDentification, Authentication and trust Services
EU: European Union
FG: focus group
GDPR: General Data Protection Regulation
GPG: GNU Privacy Guard
HCI: human–computer interaction
IT: information technology
ICT: Information and Communication Technology
PbD: privacy by design
PETs: Privacy Enhancing Technologies
PHR: personal health records
PRISMACLOUD: PRivacy and Security MAintaining services in the CLOUD
RQ: requirement
SAE-service: Selective Authentic EHR Exchange service
SMS: short messaging service
TAN: transaction authentication number
UCD: user-centered design
UI: user interface

Edited by G Eysenbach; submitted 03.05.18; peer-reviewed by B Brumen, M Abdelhamid; comments to author 07.06.18; revised version received 24.09.18; accepted 29.09.18; published 21.12.18.

Please cite as:
Alaqra AS, Fischer-Hübner S, Framner E
Enhancing Privacy Controls for Patients via a Selective Authentic Electronic Health Record Exchange Service: Qualitative Study of Perspectives by Medical Professionals and Patients
J Med Internet Res 2018;20(12):e10954
URL: http://www.jmir.org/2018/12/e10954/
doi:10.2196/10954
PMID:30578189

Review

Elements of Trust in Digital Health Systems: Scoping Review

Afua Adjekum[1], MSc; Alessandro Blasimme[1], MA, PhD; Effy Vayena[1], MSc, PhD

Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

Corresponding Author:
Effy Vayena, MSc, PhD
Department of Health Sciences and Technology
ETH Zurich
Auf der Mauer 17
Zurich,
Switzerland
Phone: 41 44 632 8301
Email: effy.vayena@hest.ethz.ch

Abstract

Background: Information and communication technologies have long become prominent components of health systems. Rapid advances in digital technologies and data science over the last few years are predicted to have a vast impact on health care services, configuring a paradigm shift into what is now commonly referred to as digital health. Forecasted to curb rising health costs as well as to improve health system efficiency and safety, digital health success heavily relies on trust from professional end users, administrators, and patients. Yet, what counts as the building blocks of trust in digital health systems has so far remained underexplored.

Objective: The objective of this study was to analyze what relevant stakeholders consider as enablers and impediments of trust in digital health.

Methods: We performed a scoping review to map out trust in digital health. To identify relevant digital health studies, we searched 5 electronic databases. Using keywords and Medical Subject Headings, we targeted all relevant studies and set no boundaries for publication year to allow a broad range of studies to be identified. The studies were screened by 2 reviewers after which a predefined data extraction strategy was employed and relevant themes documented.

Results: Overall, 278 qualitative, quantitative, mixed-methods, and intervention studies in English, published between 1998 and 2017 and conducted in 40 countries were included in this review. Patients and health care professionals were the two most prominent stakeholders of trust in digital health; a third—health administrators—was substantially less prominent. Our analysis identified cross-cutting personal, institutional, and technological elements of trust that broadly cluster into 16 enablers (altruism, fair data access, ease of use, self-efficacy, sociodemographic factors, recommendation by other users, usefulness, customizable design features, interoperability, privacy, initial face-to-face contact, guidelines for standardized use, stakeholder engagement, improved communication, decreased workloads, and service provider reputation) and 10 impediments (excessive costs, limited accessibility, sociodemographic factors, fear of data exploitation, insufficient training, defective technology, poor information quality, inadequate publicity, time-consuming, and service provider reputation) to trust in digital health.

Conclusions: Trust in digital health technologies and services depends on the interplay of a complex set of enablers and impediments. This study is a contribution to ongoing efforts to understand what determines trust in digital health according to different stakeholders. Therefore, it offers valuable points of reference for the implementation of innovative digital health services. Building on insights from this study, actionable metrics can be developed to assess the trustworthiness of digital technologies in health care.

(J Med Internet Res 2018;20(12):e11254) doi:10.2196/11254

KEYWORDS

digital health; digital health technologies; health care; health systems; trust

Introduction

Background

Digital health broadly refers to the use of information and communication technologies to improve human health, health care services, and wellness for both individuals and populations [1,2]. It has been argued that the capacity to collect, store, and analyze extensive amounts of health data is the chief driving force of digital health [3]. The accessibility of such data is rejuvenating the process involved in diagnosing, managing, and treating disease, thus exceeding the conventional boundaries of how health care institutions and providers operate. A case in point is the myriad number of smartphone apps that allow patients to seamlessly monitor various aspects of their health care beyond the confines of a health care institution [1].

There is currently no consensus on a definition for digital health. The term "digital medicine" for instance, resembles digital health, as it also refers to the use of digital technologies such as biosensors and smartphones to refine and individualize medicine [4]. Given how they are often described, electronic health, mobile health (mHealth), telecare, and telehealth could also be used interchangeably with digital health [5]. This ambiguity calls for a need to generate an inclusive definition that captures the different terms that may be used to portray digital health.

The US Food and Drug Administration (FDA) depicts digital health as comprising of mHealth, wearable devices, telehealth, telemedicine, personalized medicine, electronic health records (EHRs), and health information technology (IT) [6]. In this review, we adopt this as our working definition of digital health. Throughout this paper, the term "digital health" refers to all of the aforementioned categories. So far, there has been a prolific development of digital health technologies, and the value of such ventures continues to rise at a steady pace. In 2017 alone, the global net worth of the digital health industry was estimated at US $25 billion (£19 billion; €21 billion). Some estimates even project that digital health could cut back up to US $7 billion of US health care expenditure annually [7].

Beyond economic gains, improved safety and efficacy are among the anticipated benefits of digital health [7-10]. Current evidence supports the notion that digital health does indeed bolster safety within health systems [11]. In the domain of health care delivery, digital health promises to abate mortality, shorten hospital admissions, and decrease medication errors [11]. Despite these advances, there are privacy and data protection concerns associated with the pace of development of digital health products [7,12]. Moreover, as data from digital health tools such as mHealth apps increasingly inform medical decision making, the issue of medical liability comes to the fore [13,14]. The considerations about privacy and data protection highlight the ethical challenges that bear directly on the trustworthiness of digital health. While numerous studies have analyzed such ethical issues [15-19], the determinants of trust in digital health are yet to receive comparable levels of attention [1,3,20-22].

What is Trust?

Trust is an elusive concept that is difficult to pin down in operational terms. Relationships of trust can exist between individuals, between individuals and the organizations they come into contact with, or between 2 organizations of any given nature [23]. Trust is oftentimes illustrated as a relationship between one party (a trustor) and another (a trustee) with optimistic anticipation that the trustee will fulfill the trustor's expectations [23,24]. Trust relationships often lack enforceable obligations and are thus vulnerable to deception [25]. Consequently, different sets of reasons encourage trust relationships. Chief among them are the trustee's reliability (possessing a good reputation), competence (having the technical skills to perform the task at hand), and integrity (generally acting in an honest way) [26].

Within health systems, trust is a prominent component of doctor-patient relationships [27-29]. It improves not only health care access but also treatment outcomes and patient satisfaction [30,31]. However, whether or not it is appropriate to talk about trust between people and inanimate objects—such as technological products—remains an open question in the literature [21,32]. Indeed, the inclination of individuals to purchase or use products that are derived from "expert systems"—those structures that rely on either technical know—how or professional expertise and whose outcomes are consequently pervasive, opaque, or easily taken for granted—has been described as a tangible component of trust [33].

Some experts suggest that trust is propelled by contingency rather than risk, while others maintain that the ability to weigh risks and to choose between different actions drives trust [34]. Despite the risk of deception within any trust relationship, it is disputable whether one chooses to trust solely by weighing risks or actively by evaluating alternative options. Be that as it may, in the case of medical technologies, institutional trust and technical reliability are deeply intertwined [35]. In terms of digital health technologies, we hypothesize that trust is likely to develop if the risks and uncertainties associated with their use can be minimized.

As health care becomes increasingly dependent on digital technologies, exploring what determines and what foregoes trust in digital health is of paramount importance. Identifying the factors pertinent to trust can inform the development of novel health care services as well as meet the needs and expectations of users and patients. In addition, such factors can be taken into account for the assessment of both new and existing digital health services. Thus, this study seeks to contribute to this discourse by analyzing what the relevant stakeholders in digital health consider as the enablers and impediments of trust in digital health.

Methods

Overview

This review aimed to summarize the enabling and impeding factors of trust in digital health. To this end, we conducted a scoping review using Arksey and O'Malley's proposed framework on scoping reviews [36]. A scoping review

methodology was chosen, as it appropriately captures broad and ambiguous topics, like digital health, that may involve a myriad of study designs. We searched for studies that reported on the perspectives of different digital health stakeholders. From these perspectives, we discerned views on what was reported to facilitate trust and what hindered it. Often, some of these same factors were recognized as relevant for the acceptance of a particular technology. By acceptance, we mean adoption and use grounded in or at least co-occurring with trust on the part of users. This understanding of trust as a potential determinant of acceptance reflects some credited models of technology acceptance in the health care sector [37].

Information Sources

We searched 5 databases: MEDLINE, EMBASE, the Cumulative Index to Nursing and Allied Health Literature, PsycINFO, and Web of Science for peer-reviewed studies as well as gray literature. We worked with a research librarian at the University of Zurich, Switzerland, to identify relevant bibliographic databases and to construct a search strategy that would ensure comprehensive results.

Search Strategy

The search strategy involved formulating keywords and Medical Subject Headings around the 2 main themes of this study, namely, trust and digital health. Since the concept of trust can be ill-defined within the literature [35], we set out to include synonyms such as expectation, mistrust, confidence, and experience to capture the heterogeneity of trust descriptions within the literature (Multimedia Appendix 1). Digital health, on the other hand, was disaggregated into its distinctive components as described by the FDA: mHealth, wearable devices, telehealth, telemedicine, personalized medicine, and health IT. The searches were restricted to publications available in English, French, German, Italian, and Spanish with no publication date restrictions, to allow the search results to encompass a broad range of relevant studies. The searches commenced on July 20, 2017, and concluded on August 18, 2017. The recovered studies were then exported into the Endnote X8.2 reference software.

Eligibility Criteria of Included Studies

To capture the wide array of studies that may be relevant to this topic, we did not predefine the study designs of included studies. This allowed for the inclusion of qualitative, quantitative, intervention, and mixed-methods studies. We assessed the relevance of the retrieved studies to ensure that they related to either of the abovementioned digital health technologies. Moreover, each study was required to meet at least 1 of the following criteria: (1) investigate stakeholder perceptions, attitudes, expectations, and perspectives toward digital health or (2) highlight some potential enablers and impediments to trust in digital health technologies and services.

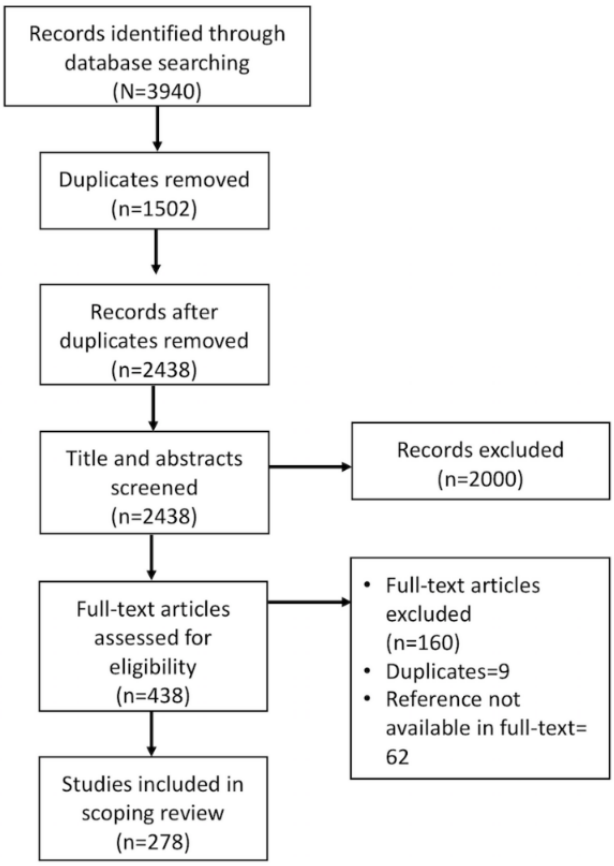
Study Selection, Categorization, and Data Extraction

As is customary in scoping reviews, we employed an iterative approach to select, categorize, and extract data from the recovered studies [36]. We used a 2-step process to select relevant articles. At first, 1 author (AA) reviewed all of the titles and abstracts derived from the search. In order to reduce sampling bias [38], a second author (AB) reviewed a random sample of 243 titles along with their associated abstracts (constituting 10% of the total sample after duplicates had been removed). To assess the level of agreement between the 2 reviewers, an interrater reliability score using Cohen kappa was computed along with its corresponding CI and *P* value. The Cohen kappa score for the 2 coders (AA and AB) was .661 (95% CI 0.465-0.857; $P<.001$). According to McHugh (2012), a kappa of.661 signifies a moderate agreement between the coders [39].

Overall, we retrieved a total of 3940 search results from the 5 databases. Of these, 1474 were identified as duplicates and discarded. However, during the screening process, we discovered an extra 28 duplicates, increasing the total number discarded to 1502. This led to screening the titles and abstracts of 2438 articles of which 438 were eligible for full-text screening. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram below (Figure 1) lays out these procedures in more detail [40]. The final number of articles included in the review was 278.

From each article, we documented the author's name, year of publication, country of origin, sample size, study design (eg, qualitative or quantitative), digital health type as well as the relevant stakeholders. A descriptive, analytical approach was used to summarize the outcomes of the studies. We identified the trust elements (enablers and impediments) by charting the key themes and issues identified from each study [36]. To develop these themes, the results section of each study was scrutinized to identify various stakeholder priorities, perspectives, expectations, perceptions, and attitudes toward a particular digital health technology or service. Multimedia Appendix 2 shows the studies from which each element was derived. Since either an enabler or impediment could be derived from the same study, we reported the overall number of studies that support each element rather than percentages. Simultaneously, we compiled a list of recurring terminologies that were used to represent or describe the various digital health technologies, which we termed "health technology types."

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



Results

Characteristics of Articles

Of the 278 articles included in this review, 51 (51/278, 18.3%) related to telemedicine and telehealth, 24 (24/278, 8.6%) to personalized medicine, 47 (47/278, 16.9%) to mHealth, 73 (73/278, 26.3%) to health IT, 73 (73/278, 26.3%) to EHRs, and 4 (4/278, 1.4%) to wearable devices, while 6 (6/278, 2.2%) concerned 2 or more digital health technologies. Most of the studies were conducted in 2015 (50/278, 18.0%), and the median year was 2014. The oldest study was conducted in 1998, while the most recent study was from 2017. There were 98 qualitative studies, 133 quantitative studies, 45 mixed method studies, and 2 intervention studies. Data from Web-based sources were collected in 7 studies. Overall, the studies were conducted in 40 countries; the United States was the most represented (101/278, 36.3%). The United Kingdom had the second highest number of studies (47/278, 16.9%) followed by Australia (16/278, 5.8%) and Canada (15/278, 5.4%; see Multimedia Appendix 3).

Digital Health Technologies and Services

For each digital health technology, we uncovered several health technology types employed to provide digital health services. Within each digital health category, there appear to be multiple terminologies to describe identical or variable technologies or services. In many instances, there were only slight variations differentiating one service from the other. For example, electronic patient records, electronic medical records, and electronic health care records were variable forms of EHRs, while Web-based consultations, online support groups, and Web-based health information were some examples of health IT. Multimedia Appendix 4 provides a list of the variable terminologies identified from the included studies.

Stakeholders

In our analysis, we identified 2 major stakeholders: *patients* or the *public* (187 studies) and *health care professionals* (HCPs; 101 studies). A third less predominant group—*health administrators* (HAs; 20 studies)—was also identified. For the sake of clarity, HCPs refer to a broad range of health care specializations that include pharmacists, occupational therapists, physical therapists, physicians, and nurses. Other stakeholders that were considerably less represented in the analyzed studies included medical and nursing students, consumer groups, health policy makers, data controllers, academic researchers, social workers, counselors, and IT technicians.

Trust Enablers and Impediments

Our findings indicate that trust in digital health technologies and services is affected by a variety of elements. In this study, trust enablers refer to those factors that encourage stakeholders' trust in digital health, while trust impediments denote the factors that can potentially hinder trust. These trust enablers and impediments, therefore, underscore the elements that influence stakeholder decisions on whether or not to place their trust in digital health technologies.

Personal Elements

By personal elements, we designate factors that influence trust in digital health at the individual level. The higher the likelihood of a digital health technology or service to enhance job performance, the more likely stakeholders are to trust it due to convenience and *usefulness* (110 studies). Moreover, *sociodemographic factors* (84 studies) such as ethnicity, income, and educational status affected an individual's trust in digital health either positively or negatively, thereby acting simultaneously as enablers and impediments. *Ease of use* (53 studies)—the propensity for systems to require minimal effort for use—also influenced trust positively. Other personal elements include *fair data access* (21 studies), *recommendations* (17 studies) from family members, acquaintances and colleagues as well as *self-efficacy* (15 studies). The latter denotes a refined acumen to manage one's own health [41]. *Altruism* (9 studies) also contributed to stakeholder involvement in digital health enterprises and was driven by the prospect of contributing to novel and beneficial therapies that would benefit society.

A number of studies reported *excessive costs* (34 studies) and *limited accessibility* (55 studies) as potential barriers to trust and, therefore, acceptance. *Fear of data exploitation* (25 studies) from third parties such as insurance and pharmaceutical companies was another palpable impediment to trusting digital health systems.

Technological Elements

The technological elements refer to the technical components of digital health technologies that make them appealing to accept and use. In terms of sensitive personal data such as genetic data, robust systems that delivered on safety and *privacy* (73 studies) were crucial to trust. There was a high affinity for *customizable design features* (28 studies) that allowed stakeholders to tailor devices to their specific needs. Since HCPs were often required to utilize disparate software programs, they requested *interoperable* (10 studies) systems that ensured that newer systems are compatible with currently existing ones. Relating to trust impediments, *defective technology* (32 studies) was a culprit for the minimal use of digital health technologies or services.

Institutional Elements

The institutional elements denote the strategies that are implemented within establishments that influence stakeholder trust in digital health. Several studies highlighted that various stakeholders had suggestions, expectations, or feedback to provide on how best to improve digital health services. Consequently, *stakeholder engagement* (71 studies), which involves taking stakeholders' opinions into account, emerged as a relevant condition to increase trust in digital health. *Improved communication* (46 studies) was a cross-cutting expectation from digital health technologies. Both patients and HCPs valued the many communication avenues that digital health provided. In 40 studies, it appeared that there was a need for *initial face-to-face* interactions prior to the introduction of digital health services. Generally, stakeholders expected digital health technologies to build upon and improve on existing systems. Hence, they preferred technologies that *decreased workloads* (82 studies).

The *reputation of service providers* (71 studies), however, served as either an enabler or impediment to trust in digital health. A good reputation encouraged trust and vice versa. *Time-consuming* (42 studies) technologies as well as those that provided *information of poor quality* (51 studies) impeded trust. Other impediments identified included *insufficient training* (54 studies) and uncertainties originating from *inadequate publicity* (44 studies) about the capabilities, existence, and risks involved in using digital health. Finally, trust was also hindered by the absence of *guidelines for standardized use* (22 studies).

In Table 1, we provide a summary of these findings and highlight the stakeholders for whom these elements appeared pertinent. In the table, found in parenthesis next to each element are the total number of studies (n). A checkmark is also used to illustrate the respective trust elements that each stakeholder is associated with.

Table 1. Trust enablers and impediments alongside their corresponding stakeholders.

Element classification	Enablers of trust	Impediments to trust	Stakeholders		
			Patients	HCPs[a]	HAs[b]
Personal elements	Altruism (n=9)	N/A[c]	✓[d]	N/A	N/A
	Ease of use (n=30)	N/A	✓	✓	✓
	N/A	Excessive costs (n=34)	✓	✓	✓
	Fair data access (n=21)	N/A	✓	✓	N/A
	N/A	Fear of data exploitation (n=25)	✓	N/A	N/A
	Recommendation by others (n=17)	N/A	✓	✓	N/A
	Self-efficacy (n=15)	N/A	✓	✓	N/A
	N/A	Limited accessibility (n=55)	✓	✓	N/A
	Sociodemographic factors (n=84)[e]	Sociodemographic factors (n=84)[e]	✓	✓	N/A
	Usefulness (n=110)	N/A	✓	✓	N/A
Technological elements	Customizable design features (n=28)	N/A	✓	✓	N/A
	N/A	Defective technology (n=32)	✓	✓	✓
	Interoperability (n=10)	N/A	N/A	✓	N/A
	Privacy (n=73)	N/A	✓	✓	N/A
Institutional elements	Decreased workloads (n=83)	N/A	N/A	✓	✓
	Guidelines for standardized use (n 22)	N/A	N/A	✓	✓
	Improved communication (n=46)	N/A	✓	✓	✓
	N/A	Inadequate publicity (n=44)	✓	✓	✓
	Initial face-to-face contact (n=40)	N/A	✓	✓	N/A
	N/A	Insufficient training (n=54)	✓	✓	✓
	N/A	Poor information quality (n=51)	✓	✓	✓
	Service provider reputation (n=71)[e]	Service provider reputation (n=71)[e]	✓	✓	N/A
	Stakeholder engagement (n=71)	N/A	✓	✓	N/A
	N/A	Time-consuming (n=42)	N/A	✓	✓

[a]HCP: health care professional.

[b]HA: health administrator.

[c]N/A: not applicable.

[d]Check mark indicates respective trust elements that each stakeholder is associated with.

[e]These elements (sociodemographic factors and service provider reputation) are simultaneously trust enablers and impediments.

Discussion

Principal Findings

This study highlights the enablers of and impediments to trust in digital health technologies and services. Our results show that digital health encompasses a wide variety of health technology types and their respective services. Altogether, we identified 3 primary stakeholders: *patients*, *HCPs*, and *HAs*. Moreover, our findings map out cross-cutting *personal*, *technological*, and *institutional* trust elements in the form of enablers and impediments to trust in digital health technologies. Of these elements, sociodemographic factors and service provider reputation acted simultaneously as enablers and impediments.

A possible interpretation of the ambivalent nature of sociodemographic factors may lie in the fact that a lack of resources, be them material or educational, render people in a vulnerable state. Within health care settings, individuals often compensate for their vulnerability by perceiving health workers as potential threats [42]. The level of risk involved in instances of unfulfilled or broken trust impacts the willingness of vulnerable people to entrust individuals, institutions, or technologies with various tasks. In a similar fashion, those sitting at the high end of the socioeconomic spectrum may be prone to trust new technologies because of their perceived ability to control them. Alternatively, they may have higher expectations with regards to health care services and, thus, set the bar of trustworthiness much higher than the more disadvantaged strata of the population.

The ambiguity that we uncovered in this study reflects what other studies on trust vis-à-vis sociodemographic status have highlighted. Available evidence on the role of sociodemographic factors (eg, ethnicity, gender, and educational status) within the health care context is mixed. For instance, 1 study, has shown that patient characteristics (with the exception of age) rarely predict trust in patient-doctor relationships [43]. Conversely, others have identified patient characteristics such as age, ethnicity, income status, educational level, and literacy levels as crucial factors affecting the use of electronic health [20,44]. In light of these discrepant findings, further research is needed to clarify the underlying effects of sociodemographic factors in digital health.

A prevalent theme throughout this review was that stakeholders appear to trust profit-making entities such as insurance and pharmaceutical companies much less than they do public institutions like universities. This is a widespread phenomenon that reflects greater public assumptions about the private sector's interests and profits [45]. Our findings support the importance of reputation to trust even though *service provider reputation* was identified as both a trust impediment and enabler. On the one hand, when a service provider embodies high ethical standards and is proficient at providing required services, they attain the advantage of shaping the expectations of stakeholders positively. In contrast, negative performance statistics of a service provider stand to give rise to negative expectations about their proficiency.

Despite stakeholder optimism about digital health tools, there are notable concerns about the accuracy of digital information exacerbated by the absence of uniform quality controls and standards [23]. Onora O'Neill has underscored the importance of enacting policies that address these challenges [26]. Based on the studies concerning Web-based health information included in this review, it was observed that patients and HCPs struggled to establish the quality of digital information. Consequently, in order to gauge the authenticity, veracity, and usefulness of digital health technologies or services, they relied quite significantly on *recommendations* from family members, colleagues, or acquaintances.

The FDA definition that we adopted for this review features personalized medicine as one of the components of digital health. Domains such as personalized medicine rely on the creation of large cohorts of deeply characterized individuals, as is the case with the 1 million participant research cohort being built for the Precision Medicine Initiative in the United States [3,46,47]. Success in this area will crucially depend on trust [48,49]. How to gain the degree of public support and personal commitment that is needed to build such infrastructures is far from obvious. In such cases, the ability to measure trustworthiness against a validated set of criteria will greatly increase the odds of success for such initiatives. Our study can be considered as a vital step in this direction, laying the conceptual groundwork for the development of such tools.

As we have shown, trust in digital health technologies and services depends on the interplay of a complex set of enablers and impediments. This study sheds light on what determines trust in digital health according to different stakeholders. More specifically, our findings can be of help in the implementation of innovative digital health technologies and services as well as in the management of existing digital health infrastructures. Building on insights from this study, actionable metrics such as the patient trust in telemedicine services tool can be developed to assess the trustworthiness of digital technologies in health care [50]. Each metric would need to undergo a validation process before being deployed in practice by HAs charged with monitoring or developing digital health services.

Overall, engaging with efforts to investigate the different dimensions of trust is particularly urgent given the growing attention from entities such as governments. This heightened level of attention is warranted due to the potential impacts of ever more innovative forms of digital health. Some approaches to digital health, in particular, those relying on big data, predictive analytics, and artificial intelligence [51-53] will require dedicated governance models in order to deliver on their promises while meeting the expectations of their users [54]. Reliable ways of measuring trustworthiness will, thus, be a key tool in such a rapidly evolving scenario.

Limitations

A drawback to this study is the unequal number of studies in each digital health category. Although this was unlikely to have skewed our findings, there were relatively fewer studies on the newer forms of digital health such as wearable devices. Despite suggestions for reviews to be screened by 2 individuals, the volume and the complicated 2-step process involved in gleaning relevant information meant that only 1 author (AA) could fully screen all of the publications. Nevertheless, a second author (AB) screened 10% of the total publications for which a kappa statistic was calculated to ensure a minimal level of bias. Even though there was a moderate interrater agreement score (kappa=.661; 95% CI 0.465-0.857; P<.001), our kappa statistic is well above the .60 value that represents an inadequate agreement threshold [39]. Lastly, we acknowledge that scoping reviews can have several shortcomings [55]. However, the poorly-defined nature of both digital health and trust within the literature required a method that could map out the discourse and, thus, pave the way for a systematic review.

Conclusion

Rapid advances in digital technologies and data science over the last few years are predicted to have a tangible impact on health care services, configuring a paradigm shift into what is now commonly referred to as digital health. Digital health, however, relies heavily on trust to succeed. What counts as the building blocks of trust in digital health systems has so far remained underexplored. In this study via a scoping review approach, we seek to fill this gap by analyzing what relevant stakeholders consider as the constitutive elements of trust in digital health. Overall, 278 qualitative, quantitative, mixed-methods, and intervention studies in English were included in this review. *Patients* and *HCPs* were the 2 most prominent stakeholders to trust, while *HAs* were a third and substantially less prominent stakeholder. Altogether, the trust elements that either enabled or hindered trust in digital health clustered into *personal*, *technological*, and *institutional* factors. This study paves the way for the implementation of the criteria

necessary to measure and anticipate trust in emerging health care technologies.

Acknowledgments

This project was funded by grants from the Swiss National Science Foundation to EV (PP000P3-157556; 4075210-167223). The authors wish to thank Martina Gosteli for her help with the search strategy and Felix Gille for valuable comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search queries.
[PDF File (Adobe PDF File), 27 KB - jmir_v20i12e11254_app1.pdf]

Multimedia Appendix 2
Studies illustrating trust enablers and impediments.
[PDF File (Adobe PDF File), 136 KB - jmir_v20i12e11254_app2.pdf]

Multimedia Appendix 3
List of study countries.
[PDF File (Adobe PDF File), 24 KB - jmir_v20i12e11254_app3.pdf]

Multimedia Appendix 4
Health technology types.
[PDF File (Adobe PDF File), 34 KB - jmir_v20i12e11254_app4.pdf]

References

1. Kostkova P. Grand challenges in digital health. Front Public Health 2015;3:134 [FREE Full text] [doi: 10.3389/fpubh.2015.00134] [Medline: 26000272]
2. World Health Organization. 2018. eHealth URL: http://www.who.int/ehealth/en/ [accessed 2018-06-07] [WebCite Cache ID 6zzwR7Y69]
3. Vayena E, Haeusermann T, Adjekum A, Blasimme A. Digital health: meeting the ethical and policy challenges. Swiss Med Wkly 2018 Dec 29;148:w14571 [FREE Full text] [doi: 10.4414/smw.2018.14571] [Medline: 29376547]
4. Steinhubl SR, Topol EJ. Digital medicine, on its way to being just plain medicine. npj Digital Med 2018 Jan 15;1(1):1-1. [doi: 10.1038/s41746-017-0005-1]
5. Boogerd EA, Arts T, Engelen LJ, van de Belt TH. "What Is eHealth": Time for An Update? JMIR Res Protoc 2015 Mar 12;4(1):e29 [FREE Full text] [doi: 10.2196/resprot.4065] [Medline: 25768939]
6. United States Food and Drug Administration. 2018. Digital Health URL: https://www.fda.gov/medicaldevices/digitalhealth/ [accessed 2018-06-07] [WebCite Cache ID 6zzwKIrhs]
7. Duggal R, Brindle I, Bagenal J. Digital healthcare: regulating the revolution. BMJ 2018 Dec 15;360:k6. [Medline: 29335296]
8. Krenn L, Schlossman D. Have Electronic Health Records Improved the Quality of Patient Care? PM R 2017 May;9(5S):S41-S50. [doi: 10.1016/j.pmrj.2017.04.001] [Medline: 28527503]
9. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood) 2005;24(5):1103-1117 [FREE Full text] [doi: 10.1377/hlthaff.24.5.1103] [Medline: 16162551]
10. McKenna RM, Dwyer D, Rizzo JA. Is HIT a hit? The impact of health information technology on inpatient hospital outcomes. Applied Economics 2017 Dec 22;50(27):3016-3028. [doi: 10.1080/00036846.2017.1414934]
11. Agboola SO, Bates DW, Kvedar JC. Digital Health and Patient Safety. JAMA 2016 Apr 26;315(16):1697-1698. [doi: 10.1001/jama.2016.2402] [Medline: 27115372]
12. Horvitz E, Mulligan D. Policy forum. Data, privacy, and the greater good. Science 2015 Jul 17;349(6245):253-255. [doi: 10.1126/science.aac4520] [Medline: 26185242]
13. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. J Med Syst 2012 Oct;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7] [Medline: 22052129]
14. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3. [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]

15. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: A systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
16. Mittelstadt B. Ethical Issues of Personal Health Monitoring: A Literature Review. 2011 Presented at: ETHICOMP; 2011; Sheffield.
17. Chan M, Estève D, Fourniols J, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012 Nov;56(3):137-156. [doi: 10.1016/j.artmed.2012.09.003] [Medline: 23122689]
18. 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] [Medline: 22343270]
19. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]
20. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
21. Torous J, Roberts LW. Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
22. 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] [Medline: 21444847]
23. Kelton K, Fleischmann KR, Wallace WA. Trust in digital information. J. Am. Soc. Inf. Sci 2008 Feb 01;59(3):363-374. [doi: 10.1002/asi.20722]
24. Adjekum A, Ienca M, Vayena E. What Is Trust? Ethics and Risk Governance in Precision Medicine and Predictive Analytics. OMICS 2017 Dec;21(12):704-710 [FREE Full text] [doi: 10.1089/omi.2017.0156] [Medline: 29257733]
25. O'Neill O. A Question of Trust : The BBC Reith Lectures 2002. Cambridge: Cambridge University Press; 2002:4-8.
26. O'Neill O. Intelligent Trust in a Digital World. New Perspectives Quarterly 2017;34(4):6-31 [FREE Full text] [doi: 10.1111/npqu.12105]
27. Dorr GS, Lipkin M. The doctor-patient relationship: challenges, opportunities, and strategies. J Gen Intern Med 1999 Jan;14 Suppl 1:S26-S33 [FREE Full text] [Medline: 9933492]
28. Bending ZJ. Reconceptualising the doctor-patient relationship: recognising the role of trust in contemporary health care. J Bioeth Inq 2015 Jun;12(2):189-202. [doi: 10.1007/s11673-014-9570-z] [Medline: 25124984]
29. O'Neill O. Autonomy and Trust in Bioethics. Cambridge: Cambridge University Press; 2002:16-21.
30. Gilson L. Trust and the development of health care as a social institution. Soc Sci Med 2003 Apr;56(7):1453-1468. [Medline: 12614697]
31. Rowe R, Calnan M. Trust relations in health care: developing a theoretical framework for the “new” NHS. J Health Organ Manag 2006;20(5):376-396. [doi: 10.1108/14777260610701777] [Medline: 17087401]
32. Hawley K. Trust, Distrust and Commitment. Noûs 2012 Oct 25;48(1):1-20. [doi: 10.1111/nous.12000]
33. Sztompka P. Trust: a Sociological Theory. Cambridge: Cambridge University Press; 1999:44-46.
34. Giddens A, Anthony. The consequences of modernity. Stanford: Cambridge, UK : Polity Press in association with Basil Blackwell, Oxford, UK, 1990; 1990:25-36.
35. Montague EN, Kleiner BM, Winchester WW. Empirically understanding trust in medical technology. International Journal of Industrial Ergonomics 2009 Jul;39(4):628-634. [doi: 10.1016/j.ergon.2009.01.004]
36. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
37. Tung F, Chang S, Chou C. An extension of trust and TAM model with IDT in the adoption of the electronic logistics information system in HIS in the medical industry. Int J Med Inform 2008 May;77(5):324-335. [doi: 10.1016/j.ijmedinf.2007.06.006] [Medline: 17644029]
38. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. J Clin Epidemiol 2006 Jul;59(7):697-703. [doi: 10.1016/j.jclinepi.2005.11.010] [Medline: 16765272]
39. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
40. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535 [FREE Full text] [Medline: 19622551]
41. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
42. Plomp HN, Ballast N. Trust and vulnerability in doctor-patient relations in occupational health. Occup Med (Lond) 2010 Jun;60(4):261-269. [doi: 10.1093/occmed/kqq067] [Medline: 20511267]
43. Hall MA, Dugan E, Zheng B, Mishra AK. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? Milbank Q 2001;79(4):613-39, v [FREE Full text] [Medline: 11789119]

44. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: A literature review. Int J Med Inform 2011 Jan;80(1):1-12. [doi: 10.1016/j.ijmedinf.2010.10.017] [Medline: 21112244]
45. Abelson J, Miller FA, Giacomini M. What does it mean to trust a health system? A qualitative study of Canadian health care values. Health Policy 2009 Jun;91(1):63-70. [doi: 10.1016/j.healthpol.2008.11.006] [Medline: 19117635]
46. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795. [doi: 10.1056/NEJMp1500523] [Medline: 25635347]
47. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. Digital Med 2018 Mar 14;1(1):1-4. [doi: 10.1038/s41746-017-0012-2]
48. Blasimme A, Vayena E. Becoming partners, retaining autonomy: ethical considerations on the development of precision medicine. BMC Med Ethics 2016 Dec 04;17(1):67 [FREE Full text] [doi: 10.1186/s12910-016-0149-6] [Medline: 27809825]
49. Blasimme A, Vayena E. "Tailored-to-You": Public Engagement and the Political Legitimation of Precision Medicine. Perspectives in Biology and Medicine 2016;59(2):172-188. [doi: 10.1353/pbm.2017.0002]
50. Velsen LV, Tabak M, Hermens H. Measuring patient trust in telemedicine services: Development of a survey instrument and its validation for an anticoagulation web-service. Int J Med Inform 2017 Dec;97:52-58. [doi: 10.1016/j.ijmedinf.2016.09.009] [Medline: 27919395]
51. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. N Engl J Med 2017 Jun 29;376(26):2507-2509. [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
52. Shapshay SM. Artificial intelligence: the future of medicine? JAMA Otolaryngol Head Neck Surg 2014 Mar;140(3):191. [doi: 10.1001/jamaoto.2013.6465] [Medline: 24457970]
53. Beam AL, Kohane IS. Translating Artificial Intelligence Into Clinical Care. JAMA 2016 Dec 13;316(22):2368-2369. [doi: 10.1001/jama.2016.17217] [Medline: 27898974]
54. Vayena E, Blasimme A. Health Research with Big Data: Time for Systemic Oversight. J Law Med Ethics 2018 Mar 27;46(1):119-129. [doi: 10.1177/1073110518766026] [Medline: 30034208]
55. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009 Jun;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]

Abbreviations

EHR: electronic health record
FDA: Food and Drug Administration
HA: health administrator
HCP: health care professional
IT: information technology
mHealth: mobile health

Edited by G Eysenbach; submitted 09.06.18; peer-reviewed by L van Velsen, C Fry; comments to author 17.08.18; revised version received 28.09.18; accepted 29.09.18; published 13.12.18.

Please cite as:
Adjekum A, Blasimme A, Vayena E
Elements of Trust in Digital Health Systems: Scoping Review
J Med Internet Res 2018;20(12):e11254
URL: https://www.jmir.org/2018/12/e11254/
doi:10.2196/11254
PMID:30545807

Original Paper

Nonpublication Rates and Characteristics of Registered Randomized Clinical Trials in Digital Health: Cross-Sectional Analysis

Mustafa Al-Durra[1,2], BSc, MSc; Robert P Nolan[3,4], PhD, CPsych; Emily Seto[1,2], PEng, PhD; Joseph A Cafazzo[1,2,5], PEng, PhD; Gunther Eysenbach[1,2], MD, MPH, FACMI

[1]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[3]Cardiac eHealth and Behavioral Cardiology Research Unit, Peter Munk Cardiac Centre, University Health Network, Toronto, ON, Canada

[4]Department of Psychology, University of York, Toronto, ON, Canada

[5]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Mustafa Al-Durra, BSc, MSc
Centre for Global eHealth Innovation
Techna Institute
University Health Network
Toronto General Hospital, R Fraser Elliott Building, 4th Floor
190 Elizabeth Street
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4800 ext 4765
Fax: 1 416 340 3595
Email: aldurra_mustafa@yahoo.ca

Abstract

Background: Clinical trials are key to advancing evidence-based medical research. The medical research literature has identified the impact of publication bias in clinical trials. Selective publication for positive outcomes or nonpublication of negative results could misdirect subsequent research and result in literature reviews leaning toward positive outcomes. Digital health trials face specific challenges, including a high attrition rate, usability issues, and insufficient formative research. These challenges may contribute to nonpublication of the trial results. To our knowledge, no study has thus far reported the nonpublication rates of digital health trials.

Objective: The primary research objective was to evaluate the nonpublication rate of digital health randomized clinical trials registered in ClinicalTrials.gov. Our secondary research objective was to determine whether industry funding contributes to nonpublication of digital health trials.

Methods: To identify digital health trials, a list of 47 search terms was developed through an iterative process and applied to the "Title," "Interventions," and "Outcome Measures" fields of registered trials with completion dates between April 1, 2010, and April 1, 2013. The search was based on the full dataset exported from the ClinlicalTrials.gov database, with 265,657 trials entries downloaded on February 10, 2018, to allow publication of studies within 5 years of trial completion. We identified publications related to the results of the trials through a comprehensive approach that included an automated and manual publication-identification process.

Results: In total, 6717 articles matched the *a priori* search terms, of which 803 trials matched our latest completion date criteria. After screening, 556 trials were included in this study. We found that 150 (27%) of all included trials remained unpublished 5 years after their completion date. In bivariate analyses, we observed statistically significant differences in trial characteristics between published and unpublished trials in terms of the intervention target condition, country, trial size, trial phases, recruitment, and prospective trial registration. In multivariate analyses, differences in trial characteristics between published and unpublished trials remained statistically significant for the intervention target condition, country, trial size, trial phases, and recruitment; the odds of publication for non-US–based trials were significant, and these trials were 3.3 (95% CI 1.845-5.964) times more likely

to be published than US–based trials. We observed a trend of 1.5 times higher nonpublication rates for industry-funded trials. However, the trend was not statistically significant.

Conclusions: In the domain of digital health, 27% of registered clinical trials results are unpublished, which is lower than nonpublication rates in other fields. There are substantial differences in nonpublication rates between trials funded by industry and nonindustry sponsors. Further research is required to define the determinants and reasons for nonpublication and, more importantly, to articulate the impact and risk of publication bias in the field of digital health trials.

(J Med Internet Res 2018;20(12):e11924) doi:10.2196/11924

KEYWORDS

clinical protocols; clinical trial; eHealth; mHealth; mobile health; publications; publication bias; randomized controlled trial; registries; telehealth; telemedicine

Introduction

Background

Empirical observations demonstrate that not all clinical studies successfully publish their results in peer-reviewed journals. Perhaps, the earliest indication of publication bias in the area of scientific research was in 1979 by Robert Rosenthal with the term "file drawer problem," acknowledging the existence of selective publication bias for studies with positive and significant results [1]. A decade later, Kay Dickersin defined publication bias as "the tendency on the part of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or strength of the study findings." [2]. The phenomenon of publication bias in clinical trials was attributed to the tendency of primary investigators and editors to submit or publish findings that are strong or statistically significant [3-5].

In 2008, a study of publication rates of clinical trials supporting successful new Food and Drug Administration drug applications found that over half of all the included trials were unpublished 5 years after obtaining approval from the Food and Drug Administration [6]. Similar findings were reported by other studies, indicating that half of all clinical trials remain unpublished in any peer-reviewed journal [7-9]. In 2014, two studies on discontinued randomized clinical trials reported discontinuation rates of 21% and 24.9%. This presents an ethical concern when considering the scarce research resources invested in the respective trials without the dissemination of any findings [10,11].

The registration of clinical trials, first proposed by Simes in 1986 [5], provides a means to mitigate publication bias by allowing researchers, scholars, and healthcare professionals to explore another source of trial results and information that may not be published [3-5]. It also helps identify discrepancies in primary outcome reporting by comparing primary outcome measures, as indicated in the trial protocols and published primary outcomes, which poses a key risk to the validity of trials [12-17]. During the past two decades, this proposal triggered numerous calls demanding mandatory registration of clinical trials [18-23]. In September 2004, the International Committee of Medical Journal Editors (ICMJE) mandated trial registration in a public registry at or before study enrollment as a prerequisite for publication in any of the ICMJE member journals and that the public trial registry should be publicly accessible at no charge and managed by a not-for-profit organization [24,25]. Soon thereafter, major medical journals announced the adoption of this new policy, including the British Medical Journal, the Lancet, and the Journal of Medical Internet Research [18,21,26]. In October 2008, the 7th revision of the Declaration of Helsinki was adopted by the World Medical Association's General Assembly, with increasing emphasis on prospective registration of trials and the ethical obligation on researchers to publish their study results [27].

Since its establishment in the year 2000, the ClinicalTrials.gov website, which is maintained by the United States National Library of Medicine at the National Institutes of Health, has become the world's largest clinical trial registry, with 286,717 registered trials, 60% of which are non-US–based as of October 11, 2018 [24,28-30].

A number of studies have analyzed and reported the characteristics of publication rates of clinical trials registered in ClinicalTrials.gov [8,9,11,31] and other data sources [6,10]. However, to our knowledge, no study has thus far analyzed and reported the characteristics of publication rates within the domain of digital health. Digital health randomized clinical trials face specific challenges, including a high attrition rate, usability issues, and insufficient prior formative research [18,32-37]. These challenges may contribute to nonpublication of trial results. This study aimed to examine the prevalence and characteristics of the nonpublication rate of digital health randomized controlled trials registered in ClinicalTrials.gov.

Research Objectives

The primary research objective was to examine the prevalence and characteristics of the nonpublication rate among digital health randomized clinical trials registered in the ClinicalTrials.gov database. The secondary research objective was to determine whether industry funding contributes to nonpublication of trials. Considering that the ClinicalTrials.gov registry is a US–based registry including 60% of non-US–based trials, we intended to explore differences in the nonpublication rate and trial size between US- and non-US–based trials [38]. We also aimed to report outcome discrepancy between prospective and published primary outcomes of the included trials.

Methods

Data Source

The ClinicalTrials.gov website provides free, global open access to the online registry database through a comprehensive website search page as well as download capabilities; for example, all registration information for a given trial can be downloaded in XML format via a Web service interface. For our study, we downloaded the entire ClinicalTrials.gov online database, with 265,657 registered clinical trials entries, on February 10, 2018.

Inclusion Criteria

The research included all eHealth-, mHealth-, telehealth-, and digital health-related randomized clinical trials that are registered in the ClinicalTrials.gov website and include any information and communication technology component, such as cellular phones, mobile phones, smart phones; devices and computer-assisted interventions; internet, online websites, and mobile applications; blogs and social media components; and emails, messages, and texts.

We also included interventional and behavioral trials with or without the results. We limited our inclusion criteria to trials with latest completion dates between April 1, 2010, and April 1, 2013. The latest date between trials' primary completion date and completion date fields was considered the latest completion date. Details regarding the evaluation of the latest completion date of trials are described in Multimedia Appendix 1 [39,40].

Justification of the Completion Date

Our search allowed for almost 5 years of a "publication lag period" between the stated trial completion date (up to April 1, 2013) and the search date for published reports (February 10, 2018). This strategy allowed us to account for longer publication cycles that may take up to several years, as indicated in prior studies [28]. For example, a study from the Netherlands that investigated the effects of a mobile phone app on the quality of life in patients with type 1 diabetes was published on May 11, 2015 [41], but the underlying clinical trial (NCT01444534) was first received by ClinicalTrials.gov on September 26, 2011, and the last update in ClinicalTrials.gov was made on October 23, 2012. To keep our data sample relevant, representative, and manageable, we chose to focus our study on a 3-year cross-sectional analysis for trials completed between April 2010 and April 2013.

Exclusion Criteria

Our search excluded registered clinical trials that were not randomized or only focused on electronic record-management systems such as electronic medical records, electronic health records, and hospital information systems as well as back-end integration systems, middleware applications, and Web services. Registered clinical trials that only reported on internet, Web-based, online, and computer-based surveys as well as television or online advertisement were also excluded. In addition, the search excluded registered clinical trials that focused only on biotechnology, bioinformatics analysis, and sequencing techniques. Finally, trials on medical devices and those only related to diagnostic imaging device, computerized neuropsychological, cognition, and oxygen assessment tools were excluded.

Search Terms

The search terms and phrases were conceptually derived from the inclusion criteria. A complete list of included search terms and phrases was developed through an iterative process (Multimedia Appendix 2 [42-52]). The following list presents the final list of the 47 search terms and phrases that were included in the search process: "smartphone," "smart-phone," "cellphone," "cell-phone," "cellular phone," "cellular-phone," "mobile phone," "cell phone," "messaging," "sms," "texting," "text reminder," "short message," "email," "e-mail," "iphone," "android," "ipad," "fitbit," "on-line," "online," "e-Health," "eHealth," "mhealth," "m-health," "internet," "e-therapies," "social-media," "social media," "facebook," "twitter," "whatsapp," "information technology," "communication technology," "app," "information application," "health application," "mobile application," "electronic application," "phone application," "touch application," "well-being application," "informatic," "computer," "digital," "web," and "wearable."

Data Extraction

Conditions

The "condition" field in ClinicalTrials.gov was defined as "the disease, disorder, syndrome, illness, or injury that is being studied" [53]. We analyzed and consolidated a total of 487 unique conditions of the 556 included registered randomized clinical trials into eight different groups, as reported in Table 2. Details of the condition classifications are provided in Multimedia Appendix 3 [54].

Discontinuation Reasons

The data exported from the ClinicalTrials.gov database includes a field "Why_Stopped" that indicates the reasons for trial discontinuation. This field is populated for trials with a withdrawn, suspended, and terminated recruitment status. We extracted and evaluated the textual content of this field as part of our recruitment analysis. Details of classification of the reasons for trial discontinuations are indicated in Multimedia Appendix 4.

Major Technology

We analyzed the descriptions of the 556 included randomized clinical trials to identify the major type of technology that was utilized within the respective interventions. Details of major technology classifications of the trials are indicated in Multimedia Appendix 5.

Table 1. Analysis of randomized clinical trials by their lead sponsor information.

Lead sponsor category (N=556)	Trials, n (%)
Foundations, Institutes, and Research Centers	72 (12.9%)
Hospitals and Medical Centers	102 (18.3%)
United States Federal Government	25 (4.5%)
University	301 (54.1%)
Other	18 (3.2%)
Industry	38 (6.8%)
Insurance	6 (15.8%)
Pharmaceuticals	2 (5.3%)
Technology and Services	29 (76.3%)
Telecommunication	1 (3.1%)

Prospective Trial Registrations

The XML data exported from the ClinicalTrials.gov database did not include an explicit field to indicate whether the trial was registered prospectively. We compared each trial's "study_first_submitted" field to the "start_date" field in order to determine if the trial was registered prospectively or retrospectively. The "study_first_submitted" field indicates the dates when the trial's primary investigator first submitted the trial record to ClinicalTrials.gov, whereas the "start_date" field indicates the date when the first participant was enrolled in the study [53]. We considered the registration to be prospective if the "study_first_submitted" date was before the "start_date."

Reporting of Study Results

The data exported from the ClinicalTrials.gov database includes a field "Has Results" to indicate whether results have been submitted for the underlying study. The XML export of the trial metadata also includes the field "FirstReceived_Results_Date," which is the date on which the study's first results were received. These fields are maintained by the primary investigators of the respective trials and, in many cases, as explained in the "Limitations" section, this field is updated voluntarily by the primary investigator and seems to be inconsistent. Our analysis showed that only 61 (11%) of all included 556 randomized clinical trials reported results in the ClinicalTrials.Gov database.

Lead Sponsor of Trials

We defined a comprehensive and specific categorization of the funding sources of trials. We analyzed the content of the "Lead_Sponsor" field, available in trials' XML files exported from ClinicalTrials.gov, which comprises information regarding the entity or individual that sponsors the clinical study [55]. We were able to categorize the "Lead_Sponsor" field into six different groups, with a more specific breakdown for industry sponsors (Table 1).

Identification of Publication

We exported all the contents of the 556 included registered randomized clinical trials from the ClinicalTrials.gov website in XML format and then identified existing publications by two processes: automated and manual identification processes. The automated identification process considered all publications referenced in the trial's registry record as well as a PubMed search according to each trial's National Clinical Trial registration number. The manual identification process was a multistep process aimed to search trial publications by key trial attributes and author details in two major bibliographic databases (PubMed and Medline) as well as the Google search engine. We only considered the results of a clinical trial to be "published" if at least one of the primary outcome measures was reported. Complete details of the publication-identification processes are described in Multimedia Appendix 6 [56-59].

Results

Screening Process

We exported the entire ClinlicalTrials.gov database, with 265,657 registered clinical trials entries as of February 10, 2018, into a local Structured Query Language server database. The 47 indicated search terms and phrases were then applied in the Structured Query Language server database as follows:

1. For every search term and phrase, identify matching records by the [Title] OR [Interventions] OR [Outcome Measures] fields. We identified 6717 matching trials.
2. Apply the latest completion date criteria between April 1, 2010, and April 1, 2013. We obtained 803 matching trials.
3. After screening against all inclusion and exclusion criteria, 247 registered clinical trials were excluded as per the following breakdown:
   - 149 trials were not randomized.
   - 52 trials had false-positive matching terms. For example, the registered clinical trial NCT01287377 examined the association between nicotine patch messaging and smoking cessation. The trial term "messaging" was a false-positive match to one of our search terms.
   - 17 trials were only related to computerized neuropsychological, cognition, and oxygen assessment tools.
   - 11 trials focused only on internet, Web-based, online, and computer-based surveys.
   - 9 trials were limited to the phone call intervention component.

- 5 trials were related to scanners and diagnostic imaging devices.
- 3 trials were related to television or online advertisement.
- 1 trial was related to electronic medical record systems.

4. Finally, 556 studies were included after screening.

A summary of the search results is presented in Figure 1.

Publication Rates

In summary, 406 of 556 (73%) trials were associated with identified outcome publications and 150 of 556 (27%) trials did not have any identified publications or their identified publications did not report any of their primary outcomes. Only 6 of the 556 (1.1%) published trials did not report any of the primary outcome measures indicated in the trial's registration protocols (Figure 2).

Analysis of Trial Characteristics

We conducted a statistical descriptive analysis, describing and summarizing the characteristics of all the 556 included registered randomized clinical trials by the following standard data elements exported from and defined by the ClinicalTrials.gov database: age group, condition, country, gender, intervention model, lead sponsor, masking, recruitment status, start date, study arms, study results, trial phase, and trial size [55]. To further our analysis, we added additional data fields that were extracted from the trial descriptions: follow-ups, latest completion date, major technology, primary outcome measure, and prospective trial registration.

We examined the relationship between trial characteristics and the nonpublication rate using bivariate and multivariate analyses. For bivariate analysis, we used the Pearson Chi-square statistical test, and for multivariate analyses, we used binary logistics regression in SPSS (IBM Corporation, Armonk, NY). The results of this analysis are depicted in Table 2.

The Pearson Chi-square test and binary logistic regression test results reported significant relationships ($P<.05$) between the nonpublication rate of trials and trial characteristics including trial condition, country, prospective registration, recruitment, trial size, and trial phases. Both tests reported no significant relationships between the nonpublication rate of trials and the age group, follow-up period, gender, intervention model, latest completion date, lead sponsor, primary outcome measures, major technology, masking, start date, study arms, and updates of trials in ClinicalTrials.gov results database.

Figure 1. Trials included from the search results.

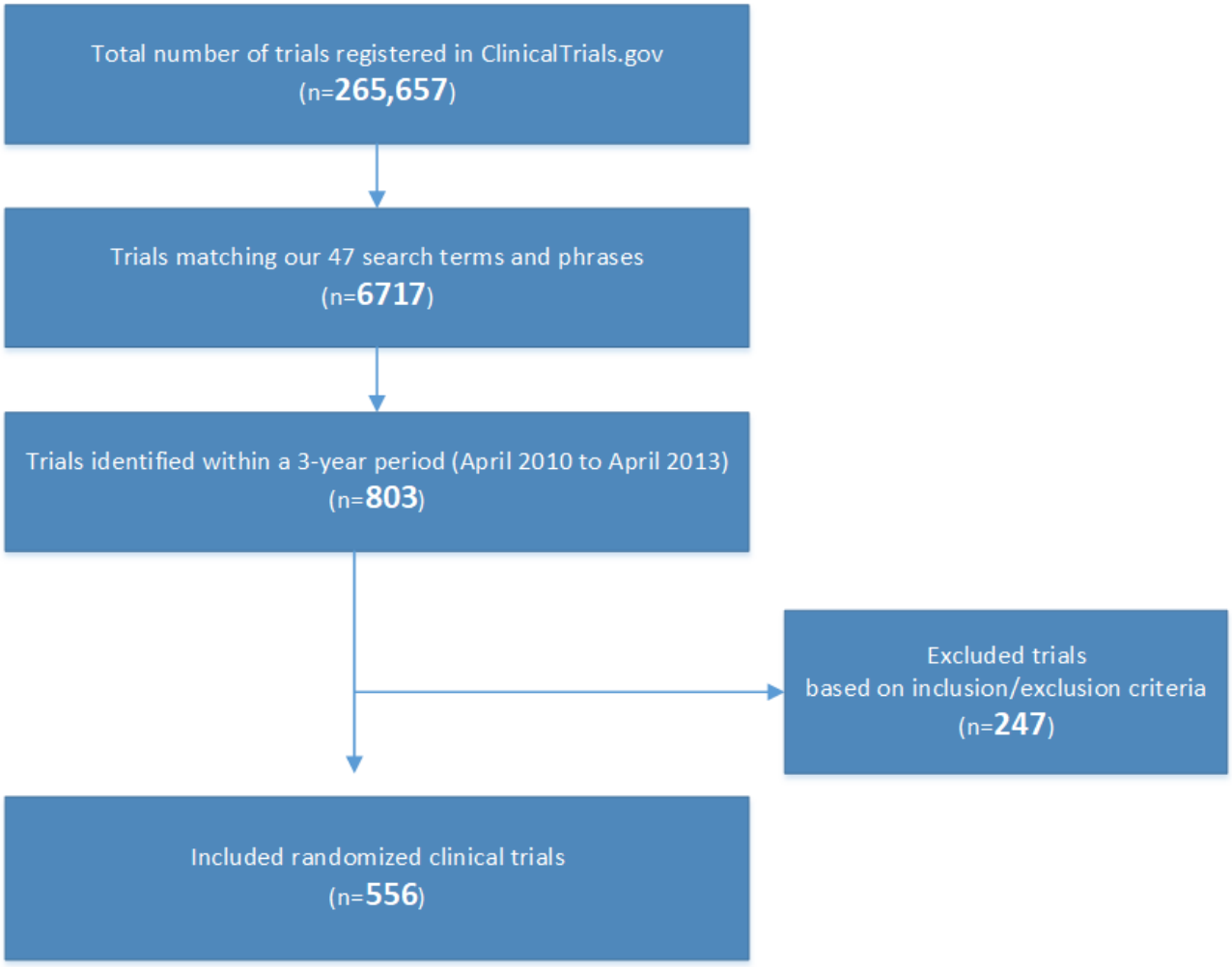
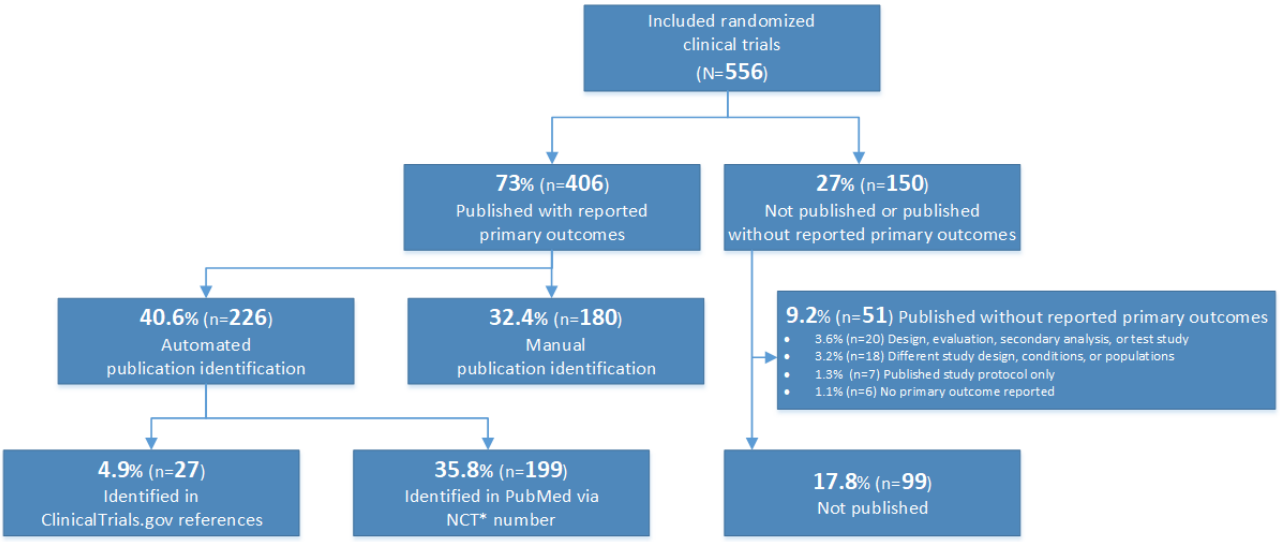


Figure 2. Results of the publication-identification process. *NCT: National Clinical Trial.



Conditions

The Pearson Chi-square test results showed a significant association (*P*=.005) between the nonpublication rate and the eight different condition groups. The highest nonpublication rate was 45.2% for randomized clinical trials focusing on the "Cancer" condition. In contrast, the lowest nonpublication rate was 15.8% for randomized clinical trials focusing on "Smoking, Alcohol Consumption, Substance Abuse and Addiction" conditions. The binary logistic regression test results showed a significant association (*P*=.01) between the nonpublication rate and intervention condition groups; however, trials on cancer or addiction/smoking conditions were not a significant predictor for nonpublication (*P*=.10, odds ratio [OR]=0.414, 95% CI: 0.740-16.173 and *P*=.12, OR=3.458, 95% CI: 0.740-16.173, respectively).

Country

The Pearson Chi-square test results showed significant differences (*P*<.001) in the nonpublication rates between the United States and other countries; the highest nonpublication rate was observed for trials in the United States (32.8%) as compared to non-US trials. The binary logistic regression test results showed a significant association between the nonpublication rate between the US and non-US trials. The odds of publication for non-US trials were significant, and these trials were 3.3 times more likely to be published than the reference group of the US–based trials (*P*<.001, OR=3.317, 95% CI: 1.845-5.964). The global distribution of all 556 randomized clinical trials included is depicted in Multimedia Appendix 7.

Lead Sponsors

Only 38 (6.8%) of the 556 included registered randomized clinical trials were funded by industry sponsors. We observed a trend of 1.5 times higher nonpublication rate for industry-funded trials than non-industry-funded trials. However, this trend was not statistically significant (*P*=.07), which may be explained by the small sample size. We also found that the percentage of industry-funded trials in the US (12%) was five times higher than that in international non-US trials (2%).

Phases

Our Pearson Chi-square test results showed significant differences (*P*=.01) between the nonpublication rate of trials and their respective study phases. Of the 556 randomized clinical trials, 427 (76.8%) had no information reported on trial phases. For 129 (23.2%) of the randomized clinical trials that reported a study phase, phase II trials (including trials registered for both phase I and II) were most commonly reported (56 trials) and had the lowest nonpublication rate (14.3%). There were 42 phase III/IV trials (including trials registered for both phase II and III), with the highest nonpublication rate of 40.5%. The binary logistic regression test results showed a significant relationship (*P*=.004) between the nonpublication rate and trial size, and phase II trials (including trials registered for both phase I and II) were 3.9 times more likely to be published (*P*=.01, OR=3.882, 95% CI: 1.460-10.318) than other phase trials. The odds of nonpublication showed a trend towards significance for phase III/IV trials (including trials registered for both phase II and III), and these trials were 3.1 times more likely to be published (*P*=.08, OR=3.112, 95% CI: 0.876-11.054); however, the trend did not reach statistical significance.

Registration of Prospective Trials

We examined the relationship between prospective trial registrations and trial nonpublication rates. Results of the Pearson Chi-square test showed a statistically significant relationship (*P*=.006) between prospective trial registrations and the nonpublication rates, with higher nonpublication rates for prospectively registered trials (11.3%) than retrospectively registered trials. Our analysis also showed that only 163 (29.3%) of all our included trials were registered prospectively. We advanced our analysis to explore the impact of the 2004 ICMJE mandate and the 2008 Declaration of Helsinki on prospective trial registrations in ClinicalTrials.gov [25,27]. Results of the Pearson Chi-square test showed a statistically significant relationship (*P*<.001) between prospective trial registration and the start date of trials, with a lower number of prospective registrations reported for trials that started after 2008 (29.7%; Table 3).

Table 2. Relationship between the characteristics of randomized clinical trials and nonpublication rate.

Trial characteristics	Unpublished RCTs[a]/Total RCTs[a], n (%)	P value[b]	Binary logistic regression	
			P value	Odds ratio (95% CI)
Overall	150/556 (27%)	—	—	—
Age Group		0.52	0.36	
Adult	27/97 (27.8%)	—	0.47	0.689 (0.254 to 1.871)
Adult/Senior	90/312 (28.8%)	—	0.73	0.864 (0.337 to 1.987)
Child	0/2 (0%)	—	0.99	>999.999 (0 to >999.999)[c]
Child/Adult	20/79 (25.3%)	—	0.29	1.738 (0.627 to 4.821)
Child/Adult/Senior	13/66 (19.7%)	—	—	Reference
Condition		0.005	0.01	
Cancer	14/31 (45.2%)	—	0.1	0.414 (0.740 to 16.173)
Chronic pain and chronic conditions (including diabetes, asthma, and COPD[d])	24/81 (29.6%)	—	0.52	0.752 (0.317 to 1.784)
Heart disease, hypertension, and stroke	15/53 (28.3%)	—	0.8	1.130 (0.436 to 2.931)
Mental health, neurodevelopmental disorders, Alzheimer, dementia, and epilepsy	14/78 (17.9%)	—	0.31	1.585 (0.648 to 3.877)
Multiconditions	23/53 (43.4%)	—	0.11	0.480 (0.197 to 1.165)
Obesity, weight management, nutrition, and physical activity	17/60 (28.3%)	—	0.11	2.455 (0.810 to 7.438)
Smoking, alcohol consumption, substance abuse, and addiction	9/57 (15.8%)	—	0.12	3.458 (0.740 to 16.173)
Others	34/143 (23.8%)	—	—	Reference
Country		<.001	<.001	
Outside the United States	39/218 (17.9%)	—	<.001	3.317 (1.845 to 5.964)
United States	111/338 (32.8%)	—	—	Reference
Enrollment		<.001	0.02	
≤5th percentile (up to 26 participants)	15/29 (51.7%)	—	0.99	>999.999 (0 to >999.999)[c]
Between the 5th and 50th percentile (between 27 and 148 participants)	58/244 (23.8%)	—	0.99	>999.999 (0 to >999.999)[c]
Between the 50th and 95th percentile (between 149-1962 participants)	59/246 (24%)	—	0.99	>999.999 (0 to >999.999)[c]
>95th percentile (more than 1962 participants)	8/27 (29.6%)	—	0.99	>999.999 (0 to >999.999)[c]
Undefined	10/10 (100%)	—	—	Reference
Follow-up period		0.14	0.21	
<1 month	13/56 (23.2%)	—	—	Reference
1-3 months	34/138 (24.6%)	—	0.44	1.436 (0.574 to 3.595)
4-6 months	32/171 (18.7%)	—	0.62	0.792 (0.314 to 1.997)
6-12 months	45/128 (35.2%)	—	0.39	0.670 (0.272 to 1.653)
12-24 months	12/40 (30%)	—	0.89	1.085 (0.330 to 3.570)
>24 months	5/17 (29.4%)	—	0.9	0.908 (0.200 to 4.124)
Undefined	9/60 (15%)	—	0.19	2.199 (0.673 to 7.185)
Gender		0.98	0.64	
Both	132/491 (26.9%)	—	0.88	0.877 (0.168 to 4.567)
Female	15/55 (27.3%)	—	0.76	1.318 (0.225 to 7.738)

Trial characteristics	Unpublished RCTs[a]/Total RCTs[a], n (%)	P value[b]	Binary logistic regression	
			P value	Odds ratio (95% CI)
Male	3/10 (30%)	—	—	Reference
Intervention model		0.09	0.29	
Single assignment	14/33 (42.4%)	—	0.99	1.475 (0.929 to 2.343)
Crossover assignment	4/21 (19%)	—	0.99	<.001 (<.001 to >999.999)[c]
Parallel assignment	121/464 (26.1%)	—	0.99	<.001 (<.001 to >999.999)[c]
Factorial assignment	11/32 (34.4%)	—	0.99	<.001 (<.001 to >999.999)[c]
Undefined	0/6 (0%)	—	—	Reference
Latest completion date by year[d]		0.07	0.06	
Before 2012	63/269 (23.4%)	—	0.06	1.636 (0.987 to 2.714)
On or after 2012	87/287 (30.3%)	—	—	Reference
Lead sponsor – industry		0.07	0.3	
No	135/518 (26.1%)	—	0.3	1.609 (0.650 to 3.986)
Yes	15/38 (39.5%)	—	—	Reference
Major technology		0.67	0.58	
Computer-based intervention (offline)	27/97 (27.8%)	—	0.99	0.995 (0.119 to 8.299)
Email notifications	7/24 (29.2%)	—	0.88	0.834 (0.082 to 8.444)
Mobile phone application	5/14 (35.7%)	—	0.84	0.771 (0.058 to 10.204)
Telemonitoring devices	16/64 (25%)	—	0.54	1.950 (0.226 to 16.842)
Text messaging	9/53 (17%)	—	0.61	1.799 (0.188 to 17.215)
Web-based intervention	84/294 (28.6%)	—	0.93	0.914 (0.114 to 7.336)
Wii	2/10 (20%)	—	—	Reference
Masking		0.41	0.41	
Open label	86/319 (26.7%)	—	0.07	12.986 (0.786 to 213.344)
Single label	53/177 (29.9%)	—	0.12	9.041 (0.546 to 149.7930)
Double label	7/30 (23.3%)	—	0.07	15.213 (0.781 to 296.201)
Triple label	1/16 (6.3%)	—	0.99	>999.999 (0 to >999.999)[c]
Quadruple label	1/7 (14.3%)	—	0.17	13.859 (0.332 to 578.089)
Undefined	2/7 (28.6%)	—	—	Reference
Phases		0.01	0.004	
0/I	5/31 (16.1%)	—	0.08	3.112 (0.876 to 11.054)
I/II or II	8/56 (14.3%)	—	0.01	3.882 (1.460 to 10.318)
II/III, III, or IV	17/42 (40.5%)	—	0.13	0.512 (0.217 to 1.208)
Undefined	120/427 (28.1%)	—	—	Reference
Primary outcome measures		0.16	0.25	
Adherence to treatment	11/26 (42.3%)	—	0.69	0.761 (0.202 to 2.868)
Clinical evaluation	76/316 (24%)	—	0.42	1.386 (0.631 to 3.044)
Drug, tobacco, and alcohol use	10/41 (24.1%)	—	0.81	0.813 (0.148 to 4.475)
Physical activity and diet intake	9/30 (30%)	—	0.97	1.022 (0.330 to 3.161)
Process evaluation	13/58 (22.4%)	—	0.04	2.924 (1.036 to 8.250)
Undefined	1/3 (33.3%)	—	0.3	1.341 (0.782 to 2.297)
Vital measurement	30/82 (36.6%)	—	—	Reference

Trial characteristics	Unpublished RCTs[a]/Total RCTs[a], n (%)	P value[b]	Binary logistic regression	
			P value	Odds ratio (95% CI)
Prospective registration		0.006	0.29	
Retrospective	93/393 (23.7%)	—	0.29	1.341 (0.782 to 2.297)
Prospective	57/163 (35%)	—	—	Reference
Recruitment		<.001	<.001	
Active, not recruiting	0/1 (0%)	—	0.99	>999.999 (0 to >999.999)[c]
Completed	105/468 (22.4%)	—	0.002	3.303 (1.564 to 6.976)
Suspended	3/4 (75%)	—	0.21	0.188 (0.014 to 2.497)
Terminated	11/17 (64.7%)	—	0.21	0.403 (0.098 to 1.656)
Unknown status	21/56 (37.5%)	—	0.99	>999.999 (0 to >999.999)[c]
Withdrawn	10/10 (100%)	—	—	Reference
Start date by year[e]		0.71	0.99	
After 2008	109/413 (26.4%)	—	0.99	<.001 (<.001 to >999.999)[c]
On or Before 2008	41/142 (28.9%)	—	0.99	<.001 (<.001 to >999.999)[c]
Undefined	0/1 (0%)	—	—	Reference
Study arms		0.11	0.4	
One	8/18 (44.4%)	—	0.17	0.240 (0.032 to 1.820)
Two	101/410 (24.6%)	—	0.63	1.486 (0.296 to 7.459)
Three	27/75 (36%)	—	0.74	0.756 (0.143 to 3.999)
Four or more	11/38 (28.9%)	—	0.78	1.295 (0.219 to 7.646)
Undefined	3/15 (20%)	—	—	Reference
Study results reported		0.86	0.79	
No	133/495 (26.9%)	—	0.79	1.113 (0.512 to 2.420)
Yes	17/61 (27.9%)	—	—	Reference

[a]RCT: randomized controlled trial.

[b]*P* value from Pearson Chi-square test.

[c]Nonconvergence was reported after 20 iterations possibly due to quasicomplete separation. Logistic regression model was not appropriate for this variable level value.

[d]The median of the latest completion date year was 2012.

[e]The cut-off point for the year of start date was set at 2008, the year when the 7th Declaration of Helsinki was adopted.

Table 3. Results of the Pearson Chi-square test between start date of trials and prospective trial registration.

Trial start date	Prospective trial registrations/total, n (%)	*P* value
Before or on 2008	73/142 (51.4%)	<.001
After 2008	90/414 (21.7%)	<.001

Recruitment

Results of the Pearson Chi-square test showed a statistically significant relationship (*P*<.001) between the trial recruitment status and nonpublication rate. Similarly, the binary logistic regression test showed a significant relationship (*P*<.001) between the trial recruitment status and nonpublication rate, and the completed trials were 3.3 times more likely to be published (*P*=.002, OR=3.303, 95% CI: 1.564-6.976). Our results also showed that discontinued trials have higher nonpublication rates than completed or active trials. We referred to trials with withdrawn, suspended, and terminated recruitment statuses as discontinued trials. We extended our analysis to explore the reasons for trial discontinuation as potential contributors to higher nonpublication rate. We examined the reasons for discontinuation of 31 trials with withdrawn, suspended, and terminated recruitment statuses among the included trials (Table 4).

Table 4. Summary of reasons for discontinuation.

Reason for discontinuation	Trials (N=31), n (%)
Recruitment challenges	9 (29%)
Funding challenges	6 (19%)
New study priorities	3 (10%)
Primary investigator/staff attrition	2 (6%)
Drop out	2 (6%)
Technical challenges	2 (6%)
Primary investigator/staff attrition and funding challenges	2 (6%)
Not provided	5 (16%)

Our analysis showed that recruitment and funding challenges are major factors contributing to discontinuation of trials and their nonpublication rates. Details of the classification of discontinuation reasons are provided in Multimedia Appendix 4.

Reporting of Study Results

Results of the Pearson Chi-square test showed no statistically significant relationship (*P*=.86) between the primary investigators who reported the results in the ClinicalTrials.gov database and the publication of trial results.

Time to Publication

We aimed to analyze the duration required to publish trial results for the 556 included trials. We measured the time to publication as the duration in years between the start date of trials and their respective publication date, which we then reported along with the number of published trials and cumulative nonpublication rates on a biyearly scale (Table 5, Figure 3).

The majority of our 556 included trials were published within 6 and 8 years of the trial's start date (356 [64%] and 393 [70.7%], respectively). A total of 148 (26.6%) trials were published in the fourth year of the trial. We also observed that half of our included trials were published between the fourth and fifth year after the trial start date.

Trial Size

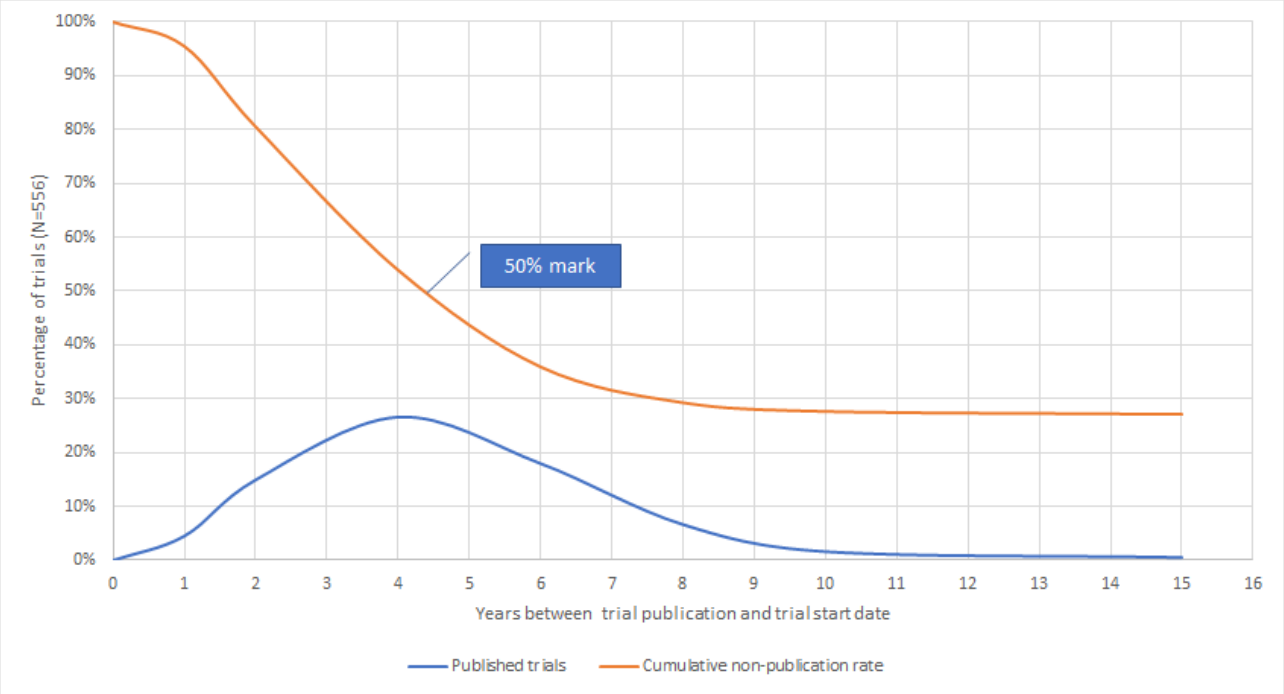
No enrollment values were identified for ten trials in the ClinicalTrials.gov database, and we could not identify any publications for these trials. We stratified all trials into four strata by size at the 5th, 50th, and 95th percentiles and found a statistically significant difference between the nonpublication rate of trials and trial size. The highest nonpublication rate was 51.7% for small trials that enrolled no more than 26 participants (at the 5th percentile), whereas the lowest nonpublication rate was 23.8% for trials that enrolled between 27 and 148 participants (between the 5th and 50th percentile).

The Pearson Chi-square test showed a statistically significant relationship between the nonpublication rate and trial size (*P*<.001). In addition, we found that half of the 546 randomized controlled trials that provided details of the trial size enrolled ≥148 participants (actual or intended). The cumulative enrolment in the 546 trials was 312,906 participants, split between 236,066 (75.44%) participants in published trials and 76,840 (24.56%) in unpublished trials. We found that the nonpublication rate was twice as high as that for trials below the 5th trial size percentile (≤26 participants) compared to other trials above the 5th trial size percentile (>26 participants).

Table 5. Analysis of trial publication cycles (duration).

Time to publication (start date to publication date), years	Published trials (N=556), n (%)	Cumulative nonpublication rate (N=556), %
2	108 (19.4%)	80.6
4	148 (26.6%)	54
6	100 (18%)	36
8	37 (6.7%)	29.3
10	9 (1.6%)	27.7
<15	3 (1%)	27.2

Figure 3. Time to publication of registered clinical trials in digital health.



Discussion

Overview

The research literature has identified the impact and risks of publication bias for researchers, clinicians, healthcare professionals, and health policy decision makers

as well as a number of factors contributing to nonpublication and discontinuation of clinical trials [21,30,60-63]. Recruitment challenges were the most-frequently reported factor contributing to clinical trial discontinuation [10], and clinical trials with larger numbers of participants or statistically significant positive outcomes were more likely to be published [6,31,64,65]. Funding sources, study language (in particular non-English language) and study design (single-center versus multicenter studies) were also identified as contributing factors for potential bias [21,64]. Authors and primary investigators reported a lack of time as the key factor for not publishing their results in a peer-reviewed journal along with other factors such as the lack of relevance and importance of their results and disagreement with coauthors [65,66].

In the domain of digital health, we analyzed the nonpublication rate among 556 randomized clinical trials that were registered in ClinicalTrials.gov, with the latest completion date between April 2010 and April 2013. We found that 27% of all included trials remain unpublished 5 years after the latest completion date. Our finding is in line with a similar study of large randomized clinical trials, with at least 500 enrolled participants, that reported a 29% nonpublication rate [31]. However, our reported nonpublication rate (27%) was considerably less than that reported in a few other similar studies with nearly half of the trials remaining unpublished [6,7,9]. We postulate that this difference may be explained by two major factors. First, the fast-paced technology involved in digital health trials could provide more extrinsic motivation for primary investigators to share and publish their results in order to become leaders in the field and stay ahead of the digital innovation curve. Second, digital health trials are likely to be sponsored by academic entities, such as universities, hospitals, and medical and research centers, that are more disciplined and obliged by scholarly ethics to publish their results. Industry sponsors and digital technology developers, on the other hand, are likely to be more driven by the scale and opportunity in the broader digital health marketplace, beyond the realm for academia and the complexity of randomized trials design.

As part of our publication-identification process, we compared the published outcomes and primary outcomes of trials indicated in the trial registration entries in ClinicalTrials.gov. Only 6 of the 556 (1.1%) published trials did not report any of the primary outcome measures indicated in the trial registration protocols. Our finding is substantially different and should not be compared to findings from other studies that reported that 40%–62% of clinical trials had at least one change in primary outcome when comparing trial publications and protocols [12,13,15]. The difference lies in our focus on identifying trial publications with at least one reported primary outcome from the trial protocol without measuring whether all, or a subset, of the primary outcomes outlined in the trial protocol were reported or examining if secondary outcomes were reported.

We reported a statistically significant relationship between the nonpublication rate and eight different condition groups in the Pearson Chi-square test (P=.005) and the binary logistic regression test (P=.01). The highest nonpublication rate was 45.2% for randomized clinical trials focusing on the “Cancer” condition. This relative underreporting suggests challenges in conducting digital health oncology trials. These challenges align with and may be explained by findings from other studies that reported several barriers to traditional oncology trials, such as recruitment, eligibility, follow-up, and oncologist and patient attitudes [67-69]. However, we suspect that there are explicit

barriers to digital health oncology trials, in particular, at the pre-enrollment and recruitment stages of the trial. Oncologists may be more inclined to enroll their patients in other traditional, nondigital health, oncology trials, where experimental drug treatment could have more tangible outcomes for their patients. Patients' perceptions and priorities to enroll in a trial could also be influenced by the preferences of their treating oncologists. In our study, only two trials were funded by the pharmaceutical industry: This clearly small number of pharmaceutical industry-funded trials supports our postulate of explicit pre-enrollment barriers to digital health oncology trials.

We also found that half of our included trials enrolled ≥148 participants, which is similar to other findings from two different studies: 46% of trials included ≥160 participants, and 45% of trials included ≥100 participants [8,70]. On comparing trial enrollment between US–based and international randomized controlled trials, we found that US–based trials had a cumulative enrolment of 228,479 participants as compared to 48,427 participants in international trials. This finding indicates that digital health trials within the United States enroll 4.7 times more participants than international trials; this value is higher than that in all clinical trials reported in a different study, which showed that US–based trials enroll only two-thirds of the number of participants enrolled in international trials [67]. The nonpublication rate was twice as high for trials with a trial size below the 5th percentile(≤26 participants) as compared to trials with a trial size above the 5th trial size percentile (>26 participants), which is consistent with the findings of similar studies reporting that clinical trials with a larger number of participants are more likely to be published [6,31].

Randomized clinical trials are usually conducted in a series of phases, 0 to IV, to examine the intervention efficacy, safety, and adverse events over various periods and sizes of population samples [53,71-74]. However, clinical studies focusing on medical devices or behavioral interventions might not be conducted in phases and did not report information in the phase field in the ClinicalTrials.gov database [55]. The finding of our study confirms this notion, as 427 (76.8%) of the 556 included randomized clinical trials reported no information on the trial phases in the ClinicalTrials.gov database. Our results showed that phase III/IV trials have the highest nonpublication rate (40.5%) among all other phase trials and are terminated and withdrawn four times more often than other phase trials. The fact that phase III/IV trials include a large group of participants may justify the higher nonpublication, termination, and withdrawal rates when considering recruitment and attrition challenges.

In our study, we reported a statistically significant relationship between the trial recruitment status and trial nonpublication rate, and completed trials were 3.3 times more likely to be published (*P*=.002, OR=3.303, 95% CI: 1.564-6.976). Our analysis of 31 discontinued trials (trials with withdrawn, suspended, and terminated recruitment statuses) showed that enrollment and funding challenges were major contributors to the higher nonpublication rate among our included trials. This finding is in line with that of another study indicating that recruitment challenges were the most-frequently reported factor contributing to discontinuation of clinical trials [10]. Another less-frequently reported reason for discontinuation of trials is new study priorities—when the primary investigator shifts his or her priority to a new trial. The fact that a primary investigator discontinues an existing registered trial to start another new, and perhaps, similar trial questions his or her commitment to the ethics of trial registration. It is important to understand the motivation behind the discontinuation of the existing trial and the interest in starting a new trial. Primary investigators should explain if the shift in priorities to a new trial was driven by implementation challenges of the existing trial (such as insignificant outcomes and adverse events) or the research perspective of the new trial (such as a new funding or collaboration opportunity).

We analyzed the nonpublication rate with regard to the start date year of trials, stratified according to their start before or after 2008, when the 7th revision of the Declaration of Helsinki was adopted [27]. We found that the nonpublication rate for trials started in or before 2008 was 3% higher than that for trials started after 2008, although the difference was not statistically significant.

We postulate that the nonpublication rate may be higher for trials registered prospectively, as the primary investigator would register a trial before the enrollment of any participant, without knowing if the trial would be completed successfully or the results would ultimately be published. The Pearson Chi-square test showed a statistically significant relationship (*P*=.006) between prospectively registered trials and nonpublication rates, with a higher nonpublication rate for prospectively registered trials (11.3%). We also expected to see an incremental trend in the prospective registration of trials after 2008, when the 7th revision of the Declaration of Helsinki was adopted to raise awareness of prospective trial registration within the scholar community [27]. Contrary to our expectation, the Pearson Chi-square test showed a statistically significant relationship (*P*<.001) between the prospective trial registration and the trial start date, with a lower number of prospective registrations for trials starting after 2008 (29.6%). This significant decline in prospective registration, compared to the influx in retrospective registration, may be explained by the general emphasis on trial registration after 2008. It is possible that the primary investigators of unregistered trials were increasingly required to register their trials retrospectively prior to publication by the editors or the submission guidelines of the scholarly journals. However, there are two major limitations to this finding in our study: the majority (74.3%) of our included trials started after 2008, and the study scope was limited to digital health trials. These two limitations can impact the internal and external validity of our analysis to evaluate the general impact of adoption of the 7th revision of the Declaration of Helsinki on the nonpublication rate of trials and prospective trial registrations.

Most of our included trials were published within 6 to 8 years after the trial start date (356 [64%] and 393 [70.7%], respectively). We also observed that half of our included trials were published between the fourth and fifth year of the trial start date. The timelines of our findings are comparable to those of a 2007 study that analyzed time to publication of clinical trials (also measured from the start to publication date) and

reported that clinical trials with statistically significant positive results were published 4-5 years after their start date, whereas trials with negative results were published in 6-8 years [75].

When we analyzed the funding sources of trials, we found that only a small number of trials (38 [6.8%] of our included trials) were funded by the industry. This finding is in contrast with the results of other studies, in which most included trials were funded by the industry. A study of delayed and nonpublication of randomized clinical trials on vaccines reported that 85% of their included trials were funded by the industry [9]. Another cross-sectional study of nonpublication of large randomized clinical trials found that 80% of the included trials were funded by the industry [31], whereas an observational study of discontinuation and nonpublication of surgical randomized controlled trials reported that 42% of the included trials were funded by the industry [11]. In our study, a majority (76.3%) of the 38 industry-sponsored trials were funded by a technology and service industry sponsor, and only two trials were funded by a pharmaceutical industry sponsor.

We observed a trend of 1.5 times higher nonpublication rates among industry-funded trials than among non-industry-funded trials. However, the trend was not statistically significant, which may be explained by the small sample size. We also found that the ratio of industry-funded trials in the United States is five times higher than that of international trials. Although these findings may be interpreted by the predominantly privately funded healthcare system in the United States, they could also be attributed to the scale of the digital health industry in the United States compared to the rest of the world, with US–based digital health startups holding 75% of the global market shares between 2013 and 2017 [76-78].

Limitations

Despite ICMJE–mandated trial registration since 2005, not all randomized trials are registered [79]. Therefore, in practice, the proportion of unreported trials, trials that failed, and publications that did not report the primary outcomes may be different.

In this study, the ClinicalTrials.gov database was the sole data source of trial registrations. The choice was driven by feasibility challenges with limited research resources available for this study initiative and broader and global adoption of the ClinicalTrials.gov registry within the biomedical research enterprise. There are many other trials registries such as the European Clinical Trials Registry [80] and the International Standard Registered Clinical/Social Study Number (ISRCTN) registry [81]. The exclusion of all trial registries other than ClinicalTrials.gov in our analysis may have impacted the external validity (generalizability) of our findings.

Our publication-identification process was conducted between June 29, 2016, and February 10, 2018, for all included 556 randomized clinical trials. Therefore, our findings did not include studies published after February 10, 2018. This study includes trials based on their completion date and primary completion date declared in the registry record in ClinicalTrials.gov. When not provided, we considered the latest completion date as described in Multimedia Appendix 1. These criteria assume that the primary investigators and study sponsors provided and updated trial details in the ClinicalTrials.gov database. However, this is a manual and voluntarily process that may not be fully complied with, given the competing priorities and limited resources available for the primary investigators and study sponsors. These limitations may impact the generalizability of our study results.

Conclusion

From our study of 556 randomized clinical trials in the field of digital health that are registered in the ClinicalTrials.gov database, we found that nonpublication of trials is prevalent, with almost a third (150, 27%) of all included trials remaining unpublished 5 years after their completion date. There are distinct differences in nonpublication rates between US- and non-US–based trials and according to the funding sources (industry sponsors vs non-industry sponsors). Further research is required to define the rationale behind the nonpublication rates from the perspectives of primary investigators and, more importantly, to articulate the impact and risk of publication bias in the field of digital health clinical trials. Future studies could also include nonrandomized trials such as projects published in protocols (such as JMIR Research Protocols).

It is not clear whether the research or technology failed, or if the results were disappointing and scholars did not write up a report, or if reports were rejected by journals; however, given the multitude of potential publication venues, and increased transparency in publishing, the former seems more likely. Scholarly communication is evolving, and short reports of failed trials may not always be published in peer-reviewed journals, but may be found in preprint servers. With the growing popularity of preprints, future analyses may also include searches for draft reports on preprint servers (such as preprints.jmir.org) to include unpublished reports, which may further shed light on why trials failed or remained unpublished. In the meantime, a general recommendation would be to conduct thorough formative research and pilot studies before conducting a full randomized controlled trial to reduce the risk of failure such as having insufficient power due to lack of participant engagement and nonuse attrition [82].

Conflicts of Interest

GE is the editor-in-chief of the Journal of Medical Internet Research (and publisher at JMIR Publications) but was not involved in the peer-review or decision-making process for this paper. The associate editor handling this manuscript and the reviewers were blinded and not aware of the co-authorship of GE. As owner of JMIR Publications, GE may benefit from increased publication rates of digital health trials. The other authors declare no conflicts of interests.

Multimedia Appendix 1

Evaluation of the latest completion date of trials.

[PDF File (Adobe PDF File), 16KB - jmir_v20i12e11924_app1.pdf]

Multimedia Appendix 2

Determination of search terms and phrases.

[PDF File (Adobe PDF File), 27KB - jmir_v20i12e11924_app2.pdf]

Multimedia Appendix 3

Classification of trial condition groups.

[PDF File (Adobe PDF File), 18KB - jmir_v20i12e11924_app3.pdf]

Multimedia Appendix 4

Classification of reasons for discontinuation of trials.

[PDF File (Adobe PDF File), 37KB - jmir_v20i12e11924_app4.pdf]

Multimedia Appendix 5

Classification of major technologies used in trials.

[PDF File (Adobe PDF File), 19KB - jmir_v20i12e11924_app5.pdf]

Multimedia Appendix 6

Identification of publications.

[PDF File (Adobe PDF File), 65KB - jmir_v20i12e11924_app6.pdf]

Multimedia Appendix 7

Global distribution of all included trials.

[PDF File (Adobe PDF File), 26KB - jmir_v20i12e11924_app7.pdf]

References

1. Rosenthal R. The file drawer problem and tolerance for null results. Psychological Bulletin 1979;86(3):638-641. [doi: 10.1037/0033-2909.86.3.638]
2. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA 1990 Mar 09;263(10):1385-1389. [Medline: 2406472]
3. Harrison BA, Mayo-Wilson E. Trial Registration. Research on Social Work Practice 2013 Nov 19;24(3):372-376. [doi: 10.1177/1049731513512374]
4. Wager E, Williams P, Project Overcome failure to Publish nEgative fiNdings Consortium. “Hardly worth the effort”? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study. BMJ 2013 Sep 06;347:f5248 [FREE Full text] [doi: 10.1136/bmj.f5248] [Medline: 24014339]
5. Simes RJ. Publication bias: the case for an international registry of clinical trials. J Clin Oncol 1986 Oct;4(10):1529-1541. [doi: 10.1200/JCO.1986.4.10.1529] [Medline: 3760920]
6. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. PLoS Med 2008 Sep 23;5(9):e191 [FREE Full text] [doi: 10.1371/journal.pmed.0050191] [Medline: 18816163]
7. Blümle A, Meerpohl JJ, Schumacher M, von Elm E. Fate of clinical research studies after ethical approval--follow-up of study protocols until publication. PLoS One 2014;9(2):e87184 [FREE Full text] [doi: 10.1371/journal.pone.0087184] [Medline: 24586265]
8. Ross JS, Tse T, Zarin DA, Xu H, Zhou L, Krumholz HM. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. BMJ 2012 Jan 03;344:d7292 [FREE Full text] [doi: 10.1136/bmj.d7292] [Medline: 22214755]
9. Manzoli L, Flacco ME, D'Addario M, Capasso L, De Vito C, Marzuillo C, et al. Non-publication and delayed publication of randomized trials on vaccines: survey. BMJ 2014 May 16;348:g3058 [FREE Full text] [doi: 10.1136/bmj.g3058] [Medline: 24838102]

10. Kasenda B, von Elm E, You J, Blümle A, Tomonaga Y, Saccilotto R, et al. Prevalence, characteristics, and publication of discontinued randomized trials. JAMA 2014 Mar 12;311(10):1045-1051. [doi: 10.1001/jama.2014.1361] [Medline: 24618966]
11. Chapman SJ, Shelton B, Mahmood H, Fitzgerald JE, Harrison EM, Bhangu A. Discontinuation and non-publication of surgical randomised controlled trials: observational study. BMJ 2014 Dec 09;349:g6870 [FREE Full text] [doi: 10.1136/bmj.g6870] [Medline: 25491195]
12. Chan A, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004 May 26;291(20):2457-2465. [doi: 10.1001/jama.291.20.2457] [Medline: 15161896]
13. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One 2008 Aug 28;3(8):e3081 [FREE Full text] [doi: 10.1371/journal.pone.0003081] [Medline: 18769481]
14. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, CONSORT. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10(1):28-55 [FREE Full text] [doi: 10.1016/j.ijsu.2011.10.001] [Medline: 22036893]
15. Dwan K, Gamble C, Williamson PR, Kirkham JJ, Reporting Bias Group. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One 2013;8(7):e66844 [FREE Full text] [doi: 10.1371/journal.pone.0066844] [Medline: 23861749]
16. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. BMC Med 2015 Nov 18;13:282 [FREE Full text] [doi: 10.1186/s12916-015-0520-3] [Medline: 26581191]
17. Ioannidis JP, Caplan AL, Dal-Ré R. Outcome reporting bias in clinical trials: why monitoring matters. BMJ 2017 Feb 14;356:j408. [doi: 10.1136/bmj.j408] [Medline: 28196819]
18. Eysenbach G. Tackling publication bias and selective reporting in health informatics research: register your eHealth trials in the International eHealth Studies Registry. J Med Internet Res 2004 Sep 30;6(3):e35 [FREE Full text] [doi: 10.2196/jmir.6.3.e35] [Medline: 15471761]
19. Dickersin K, Rennie D. Registering clinical trials. JAMA 2003 Jul 23;290(4):516-523. [doi: 10.1001/jama.290.4.516] [Medline: 12876095]
20. Chalmers I. Underreporting research is scientific misconduct. JAMA 1990 Mar 09;263(10):1405-1408. [Medline: 2304220]
21. Turner L, Boutron I, Hróbjartsson A, Altman DG, Moher D. The evolution of assessing bias in Cochrane systematic reviews of interventions: celebrating methodological contributions of the Cochrane Collaboration. Syst Rev 2013 Sep 23;2:79 [FREE Full text] [doi: 10.1186/2046-4053-2-79] [Medline: 24059942]
22. Horton R, Smith R. BMJ. Time to register randomised trials URL: https://doi.org/10.1136/bmj.319.7214.865 [accessed 2018-12-03] [WebCite Cache ID 74NkXINwa]
23. Lapierre Y, Mohr E. Can Med Assoc J. 1994. Clinical-trial registration URL: http://www.cmaj.ca/content/150/12/1939 [accessed 2018-12-03] [WebCite Cache ID 74NkboRFg]
24. ClinicalTrials.gov. Trends, Charts, and Maps URL: https://clinicaltrials.gov/ct2/resources/trends [accessed 2018-08-09] [WebCite Cache ID 71XIYK31a]
25. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, International Committee of Medical Journal Editors. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. Lancet 2004;364(9438):911-912. [doi: 10.1016/S0140-6736(04)17034-7] [Medline: 15364170]
26. Abbasi K. Compulsory registration of clinical trials. BMJ 2004 Sep 18;329(7467):637-638 [FREE Full text] [doi: 10.1136/bmj.329.7467.637] [Medline: 15374895]
27. Krleza-Jerić K, Lemmens T. 7th revision of the Declaration of Helsinki: good news for the transparency of clinical trials. Croat Med J 2009 Apr;50(2):105-110 [FREE Full text] [Medline: 19399942]
28. Huser V, Cimino JJ. Linking ClinicalTrials.gov and PubMed to track results of interventional human clinical trials. PLoS One 2013;8(7):e68409 [FREE Full text] [doi: 10.1371/journal.pone.0068409] [Medline: 23874614]
29. ClinicalTrials.gov. Home URL: http://clinicaltrials.gov/ [accessed 2006-04-08] [WebCite Cache ID 5Eymd12jE]
30. Carter AO, Griffin GH, Carter TP. A survey identified publication bias in the secondary literature. J Clin Epidemiol 2006 Mar;59(3):241-245. [doi: 10.1016/j.jclinepi.2005.08.011] [Medline: 16488354]
31. Jones CW, Handler L, Crowell KE, Keil LG, Weaver MA, Platts-Mills TF. Non-publication of large randomized clinical trials: cross sectional analysis. BMJ 2013 Oct 29;347(oct28 9):f6104-f6104. [doi: 10.1136/bmj.f6104]
32. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
33. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, 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] [Medline: 21212045]

34. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. 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] [Medline: 23151820]
35. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
36. Clough BA, Casey LM. Technological adjuncts to increase adherence to therapy: a review. Clin Psychol Rev 2011 Jul;31(5):697-710. [doi: 10.1016/j.cpr.2011.03.006] [Medline: 21497153]
37. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. J Health Commun 2015;20(1):4-34. [doi: 10.1080/10810730.2013.864735] [Medline: 24673171]
38. ClinicalTrials.gov. Map of All Studies on ClinicalTrials.gov URL: https://clinicaltrials.gov/ct2/search/map?map= [accessed 2018-12-03] [WebCite Cache ID 736HlK7Vd]
39. ClinicalTrials.gov. 2015. How Are the Primary Completion Date and the Study Completion Date Different? URL: https://medschool.duke.edu/sites/medschool.duke.edu/files/field/attachments/ctgov_what%20are%20completion%20dates%20Tip%20Sheet_20150519_1.pdf [WebCite Cache ID 71XIjegzT]
40. ClinicalTrials.gov. Help URL: https://ClinicalTrials.gov/ct2/help/last_updated [accessed 2018-08-09] [WebCite Cache ID 71XIsSnsZ]
41. Drion I, Pameijer LR, van Dijk PR, Groenier KH, Kleefstra N, Bilo HJG. The Effects of a Mobile Phone Application on Quality of Life in Patients With Type 1 Diabetes Mellitus: A Randomized Controlled Trial. J Diabetes Sci Technol 2015 May 11;9(5):1086-1091 [FREE Full text] [doi: 10.1177/1932296815585871] [Medline: 25963412]
42. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
43. Kohl LFM, 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]
44. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2):81-92. [doi: 10.1097/NCN.0b013e3182065eed] [Medline: 21164337]
45. Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten AAHA, et al. Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review. J Med Internet Res 2013;15(1):e17 [FREE Full text] [doi: 10.2196/jmir.2130] [Medline: 23353680]
46. Clough BA, Casey LM. Technological adjuncts to enhance current psychotherapy practices: a review. Clin Psychol Rev 2011 Apr;31(3):279-292. [doi: 10.1016/j.cpr.2010.12.008] [Medline: 21382535]
47. Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. Health Info Libr J 2007 Jun;24(2):78-94 [FREE Full text] [doi: 10.1111/j.1471-1842.2007.00713.x] [Medline: 17584211]
48. van den Berg MH, Schoones JW, Vliet Vlieland TPM. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
49. 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] [Medline: 21795238]
50. Revere D, Dunbar PJ. Review of computer-generated outpatient health behavior interventions: clinical encounters “in absentia”. J Am Med Inform Assoc 2001;8(1):62-79 [FREE Full text] [Medline: 11141513]
51. Younger P. Internet-based information-seeking behaviour amongst doctors and nurses: a short review of the literature. Health Info Libr J 2010 Mar;27(1):2-10 [FREE Full text] [doi: 10.1111/j.1471-1842.2010.00883.x] [Medline: 20402799]
52. ClinicalTrials.gov. Advanced Search Field Definitions URL: https://ClinicalTrials.gov/ct2/help/how-find/advanced/field-defs [accessed 2018-08-09] [WebCite Cache ID 71XIvs08o]
53. ClinicalTrials.gov. Glossary of Common Site Terms URL: https://ClinicalTrials.gov/ct2/about-studies/glossary#C [accessed 2018-08-09] [WebCite Cache ID 71XIdKBQK]
54. NCBI. Diabetes Mellitus Mesh Term - National Center for Biotechnology Information, U.S. National Library of Medicine URL: https://www.ncbi.nlm.nih.gov/mesh/68003920 [accessed 2018-08-09] [WebCite Cache ID 71XJF8VvK]
55. ClinicalTrials.gov. Protocol Registration Data Element Definitions for Interventional and Observational Studies URL: https://prsinfo.clinicaltrials.gov/definitions.html [accessed 2018-08-09] [WebCite Cache ID 71XIov38G]
56. Wood AJJ. Progress and deficiencies in the registration of clinical trials. N Engl J Med 2009 Feb 19;360(8):824-830. [doi: 10.1056/NEJMsr0806582] [Medline: 19228628]
57. FDAAA 801 Requirements. ClinicalTrials.gov URL: http://clinicaltrials.gov/ct2/manage-recs/fdaaa [accessed 2014-04-20] [WebCite Cache ID 6OyiQKYxq]
58. National Institutes of Health (NIH). HHS takes steps to provide more information about clinical trials to the public URL: https://www.nih.gov/news-events/news-releases/hhs-takes-steps-provide-more-information-about-clinical-trials-public [accessed 2018-08-09] [WebCite Cache ID 71XJ5CroN]

59. U.S. National Library of Medicine. Clinical Trial Registry Numbers in MEDLINE®/PubMed® Records URL: https://www.nlm.nih.gov/bsd/policy/clin_trials.html [accessed 2018-08-09] [WebCite Cache ID 71XJAHi35]
60. Liebeskind DS, Kidwell CS, Sayre JW, Saver JL. Evidence of publication bias in reporting acute stroke clinical trials. Neurology 2006 Sep 26;67(6):973-979. [doi: 10.1212/01.wnl.0000237331.16541.ac] [Medline: 17000963]
61. Moscati R, Jehle D, Ellis D, Fiorello A, Landi M. Positive-outcome bias: comparison of emergency medicine and general medicine literatures. Acad Emerg Med 1994;1(3):267-271 [FREE Full text] [Medline: 7621207]
62. Sterling TD, Rosenbaum WL, Weinkam JJ. Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa. The American Statistician 1995 Feb;49(1):108-112. [doi: 10.1080/00031305.1995.10476125]
63. Parekh-Bhurke S, Kwok CS, Pang C, Hooper L, Loke YK, Ryder JJ, et al. Uptake of methods to deal with publication bias in systematic reviews has increased over time, but there is still much scope for improvement. J Clin Epidemiol 2011 Apr;64(4):349-357. [doi: 10.1016/j.jclinepi.2010.04.022] [Medline: 20800992]
64. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess 2010 Feb;14(8):iii, ix-iixi, 1 [FREE Full text] [doi: 10.3310/hta14080] [Medline: 20181324]
65. Dickersin K, Min YI. NIH clinical trials and publication bias. Online J Curr Clin Trials 1993 Apr 28;Doc No 50:[4967 words; 53 paragraphs]. [Medline: 8306005]
66. Weber EJ, Callaham ML, Wears RL, Barton C, Young G. Unpublished research from a medical specialty meeting: why investigators fail to publish. JAMA 1998 Jul 15;280(3):257-259. [Medline: 9676674]
67. Institute of Medicine. Transforming Clinical Research in the United States: Challenges and Opportunities: Workshop Summary (Forum on Drug Discovery, Development, and Translation). Washington, DC: National Academies Press (US); 2010.
68. Unger JM, Cook E, Tai E, Bleyer A. The Role of Clinical Trial Participation in Cancer Research: Barriers, Evidence, and Strategies. Am Soc Clin Oncol Educ Book 2016;35:185-198 [FREE Full text] [doi: 10.1200/EDBK_156686] [Medline: 27249699]
69. Ajithkumar TV, Gilbert DC. Modern Challenges of Cancer Clinical Trials. Clin Oncol (R Coll Radiol) 2017 Dec;29(12):767-769. [doi: 10.1016/j.clon.2017.10.006] [Medline: 29066171]
70. Ross JS, Mulvey GK, Hines EM, Nissen SE, Krumholz HM. Trial publication after registration in ClinicalTrials.Gov: a cross-sectional analysis. PLoS Med 2009 Sep;6(9):e1000144 [FREE Full text] [doi: 10.1371/journal.pmed.1000144] [Medline: 19901971]
71. Kummar S, Rubinstein L, Kinders R, Parchment RE, Gutierrez ME, Murgo AJ, et al. Phase 0 clinical trials: conceptions and misconceptions. Cancer J 2008;14(3):133-137. [doi: 10.1097/PPO.0b013e318172d6f3] [Medline: 18536551]
72. Fromer MJ. FDA Introduces New Phase 0 for Clinical Trials. Oncology Times 2006;28(15):18-19. [doi: 10.1097/01.COT.0000293387.34124.a8]
73. U.S Food and Drug Administration. The Drug Development Process - Step 3: Clinical Research URL: https://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm [accessed 2017-10-09] [WebCite Cache ID 6u5fLb6O8]
74. Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: a narrative review. Postgrad Med 2011 Sep;123(5):194-204 [FREE Full text] [doi: 10.3810/pgm.2011.09.2475] [Medline: 21904102]
75. Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. Cochrane Database Syst Rev 2007 Apr 18(2):MR000011. [doi: 10.1002/14651858.MR000011.pub2] [Medline: 17443632]
76. National Health Expenditure Data, 2016. CMS.gov - Centers for Medicare & Medicaid Services URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet.html [accessed 2018-08-09] [WebCite Cache ID 71XJOPDlw]
77. OECD Health Data: Health expenditure and financing: Health expenditure indicators. Peterson-Kaiser Health System Tracker URL: https://www.healthsystemtracker.org/chart-collection/health-spending-u-s-compare-countries/#item-u-s-similar-public-spending-private-sector-spending-triple-comparable-countries [accessed 2018-08-09] [WebCite Cache ID 71XJSKRb6]
78. Healthy Planet: Mapping Digital Health Startups Across The Globe. CB INSIGHTS: Research Briefs URL: https://www.cbinsights.com/research/digital-health-startups-world-map/ [accessed 2018-08-09] [WebCite Cache ID 71XIAP8TX]
79. Harriman SL, Patel J. When are clinical trials registered? An analysis of prospective versus retrospective registration. Trials 2016 Apr 15;17:187 [FREE Full text] [doi: 10.1186/s13063-016-1310-8] [Medline: 27079379]
80. The European Union Clinical Trials Register. The European Union Clinical Trials Register URL: https://www.ClinicalTrialsregister.eu [accessed 2018-12-03] [WebCite Cache ID 736HlK7Vd]
81. The International Standard Randomised Controlled Trial Number Register. The International Standard Randomised Controlled Trial Number Register URL: https://www.isrctn.com [accessed 2018-12-03] [WebCite Cache ID 74NiwWo30]
82. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

Abbreviations

ICMJE: International Committee of Medical Journal Editors
OR: odds ratio
RCT: randomized controlled trial

Edited by P Wicks; submitted 12.08.18; peer-reviewed by R Kok, E Chiauzzi; comments to author 25.09.18; revised version received 09.11.18; accepted 28.11.18; published 18.12.18.

Please cite as:
Al-Durra M, Nolan RP, Seto E, Cafazzo JA, Eysenbach G
Nonpublication Rates and Characteristics of Registered Randomized Clinical Trials in Digital Health: Cross-Sectional Analysis
J Med Internet Res 2018;20(12):e11924
URL: http://www.jmir.org/2018/12/e11924/
doi:10.2196/11924
PMID:30485832

Review

Health Literacy in Web-Based Health Information Environments: Systematic Review of Concepts, Definitions, and Operationalization for Measurement

Anna-Maija Huhta[1,2], MA; Noora Hirvonen[1,2,3], PhD; Maija-Leena Huotari[1,2], PhD

[1]Department of Information Studies, Faculty of Humanities, University of Oulu, Oulu, Finland

[2]Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland

[3]Department of Information Studies, Faculty of Social Sciences, Business and Economics, Åbo Akademi University, Turku, Finland

Corresponding Author:
Anna-Maija Huhta, MA
Department of Information Studies
Faculty of Humanities
University of Oulu
PO Box 1000
Oulu, FI-90014
Finland
Phone: 358 504285175
Email: anna-maija.huhta@oulu.fi

Abstract

Background: Health literacy research seems to lack a consensus on what aspects to include into *literacy* in the context of health and on how to operationalize these concepts for measurement purposes. In addition to health literacy, several other concepts, such as electronic health (eHealth) literacy and mental health literacy, have been developed across disciplines. This study examines how these different concepts are used when studying health-related competencies in Web contexts.

Objective: This study systematically reviews health literacy concepts and definitions and their operationalization in studies focused on Web-based health information environments.

Methods: A systematic literature search was conducted in April 2016 in 6 electronic databases with a limitation to articles in English published between January 2011 and April 2016. Altogether, 1289 unique records were identified and screened according to the predefined inclusion criteria: (1) original, peer-reviewed research articles written in English; (2) the topic of the article concerned literacy in the context of health; (3) informants of the study were lay people, not health professionals or students of the field; and (4) the focus of the study was placed on an Web-based information environment. In total, 180 full texts were screened, of which 68 were included in the review. The studies were analyzed with an emphasis on the used health literacy concepts and measures.

Results: On the basis of the included studies, several concepts are in use when studying health-related literacy in Web environments, eHealth literacy and health literacy being the most common ones. The reviewed studies represent a variety of disciplines, but mostly medical sciences. Typically, quantitative research methods are used. On the basis of the definitions for health literacy, 3 thematic categories were identified: general and skill-based, multidimensional, and domain-specific health literacy. Most studies adopted a domain-specific concept, followed by the ones that used a general and skill-based concept. Multidimensional concepts occurred least frequently. The general health literacy concepts were usually operationalized with reading comprehension measures, the domain-specific concepts with self-efficacy measures, and multidimensional concepts with several types of measures. However, inconsistencies in operationalization were identified.

Conclusions: The results show that in studies conducted in Web-based information environments, several different health literacy concepts are in use, and there is no clear consensus on the definitions for these concepts. Future studies should place emphasis on the conceptual development of health literacy in Web contexts to gain better results on operationalization for measurement. Researchers are encouraged to provide clear operational definitions for the concepts they use to ensure transparency in reporting.

(J Med Internet Res 2018;20(12):e10273) doi:10.2196/10273

KEYWORDS

health literacy; consumer health information; internet; review, systematic

Introduction

Background

The contemporary digital information environment challenges our understanding of what it means to be literate. The fast and free flow of information on the internet offers multiple ways to communicate, but it can also challenge with overload of information and loss of authority and identity [1]. Exercising critical thinking and employing information and digital literacies are ways to reduce the effects of information overload [2]. These types of literacies usually refer to a diverse set of competencies, skills, and strategies vital for acting in multimodal and transforming information environments. In the context of Web-based health information, these competencies are essential as the amount of health information is rapidly increasing and the possibility to encounter misinformation is apparent.

The concept of health literacy has been widely used to address literacy competencies required in health settings. A recent definition [3] describes health literacy as a concept that recognizes people's different capacities to find, understand, and use health information as well as the different life experiences that shape peoples' willingness and confidence to do these tasks. According to the World Health Organization (WHO) [4], health literacy regards the environmental, political, and social factors that determine health, and it is gained through comprehensive health education at the individual and community levels. Both the concept of health literacy and the means to measure it have been under development for over three decades. Yet, the research on the phenomena seems to lack a consensus on what aspects to include into *literacy* in the context of health and on how to operationalize it for measurement purposes [5-8].

On the basis of earlier reviews, health literacy is typically understood as individuals' functional skills, such as reading comprehension and numeracy [9] that are assessed in clinical settings [5], and the research is conducted predominantly within medical sciences [10]. More recently, however, research on health literacy–associated issues has been conducted in several other disciplines and subconcepts and related concepts have emerged [11,12]. Although the definitions have unique elements, especially the most recent definitions for health literacy overlap substantially [3]. The digital context that has changed the ways people communicate has been taken into account in the definitions of the concept only recently and, thus, needs to be investigated further.

The aim of this study was to increase understanding of the health literacy concepts that are used as well as their definitions and operationalization in Web-based information environments. The purpose was to provide a synthesis of their use in a sample of studies published between the years 2011 and 2016.

From Health Literacy to Electronic Health Literacy

Contemporary discussion on health literacy reveals that there is no consensus on the definition of the concept [5,10,13,14]. For instance, the attributes included in the concept [10] and the distinction between basic functional health literacy, communicative or interactive health literacy, and critical health literacy have been debated [14]. Mårtensson and Hensing [9] note that the research on health literacy is heterogeneous and identify 2 perspectives: health literacy as a polarized phenomenon focused on the extremes of high and low and health literacy as a multidimensional concept that acknowledges the broadness of skills in interaction with social and cultural contexts. These definitions emphasize the interactive and critical skills needed to use information for making appropriate health decisions [9]. They also consider multiple settings and recognize that there are both social and individual components to the concept [3].

The internet and the new digital tools for seeking, communicating, and using information have become embedded in the social actions of people since the 1990s. Moreover, the growing interest in consumer health and digital solutions to tailor health information for electronic health (eHealth) purposes has increased research and generated new conceptualizations for health literacy. The concept of eHealth literacy by Norman and Skinner [15] was one of the first attempts to capture the meaning of health literacy in the digital context. The definition draws on Eng's [16] definition of eHealth as "the use of emerging information and communication technology, especially the internet, to improve or enable health and health care." However, Norman and Skinner [15] add to it by stating that "[c]onsumer eHealth requires basic reading and writing skills, working knowledge of computers, a basic understanding of science, and an appreciation of the social context that mediates how online health information is produced, transmitted, and received."

The definition of eHealth literacy by Norman and Skinner [15] has been criticized for not fully describing the competencies essential in digital environments [17-19]. Gilstad [18] notes that the concept lacks the notions of contextual and cultural literacy and communicative expertise as central literacy competencies. There are several new definitions proposed for the concept. For example, Griebel et al [19] recently proposed a definition of eHealth literacy that encompasses aspects of interactivity, the dynamic evolvement of literacy, changing information practices of individuals, and the integration of technology aspects. The authors note that there are several models describing eHealth literacy but also that there is a lot of research that deals with the themes related to eHealth literacy but uses other terms [19]. Typically, health literacy is seen as an umbrella concept that covers other concepts such as eHealth literacy and mental literacy. However, the hierarchy is not entirely clear. For example, health information literacy, a concept used in information science, can be seen as a related rather than a subconcept to health literacy as it combines the concepts of health literacy and information literacy [20]. In this study, we do not focus on the hierarchical relationships of these concepts and use the phrase *health literacy concepts* to refer to all health-related literacy concepts.

Measuring Health Literacy

The first health literacy assessment tools were designed to measure the functional health literacy of individuals in clinical settings [21]. The basis of these measures is on the definitions of health literacy that present individuals' reading comprehension and numeracy as central competencies when dealing with medical texts. Therefore, these measures have been criticized for capturing only a narrow spectrum of the conception of health literacy [5,11,22]. Another way to assess health literacy is to measure the level of health knowledge of individuals. Usually, these measures are content- and context-based knowledge tests that have been developed in and for the use of clinical settings [11]. The more recent measures for health literacy consider individuals' self-reported abilities or self-efficacy as an indicator of health literacy. These measures usually aim to detect the self-perceived abilities of the individual to, for example, collect, communicate, and evaluate health information [23] or to rate the individuals' ability to understand health-related material [24]. However, the risk of assessing merely self-efficacy or behavior instead of health literacy is considered to be a major disadvantage of self-reported health literacy measures [11].

Altin et al [25] reviewed generic health literacy instruments and categorized them by their measurement modes (print, oral, numeracy, and multimodal) and their measurement approaches (objective, subjective, mixed, and multidimensional construct). The review indicated that more than two-thirds of the generic health literacy instruments were based on multidimensional constructs of health literacy. Moreover, it was shown that there is a trend toward mixing objective and subjective measurement approaches. In addition, a third of the reviewed instruments were based on existing functional literacy screeners. O'Neill et al [26] reviewed self-administered health literacy instruments and discovered that the majority of the instruments measured general health literacy, whereas one-third of them measured condition- or context-specific health literacy (see also [22]). Therefore, it was suggested that for the instruments to progress, more research should be focused on the investigation and elaboration of the construct of health literacy itself [26,27]

A systematic review on eHealth literacy measures [28] found that all the identified measures were based on self-report and measured the self-efficacy of individuals. The authors identified 3 concept-based eHealth literacy measurement tools and 5 dual-design tools that comprised individual measures of health literacy and digital literacy. The dual-design measurement tools did not intend to measure eHealth literacy but ended up doing so by including the main components of the concept [28]. An overview of the recent eHealth literacy research [29] indicates that although international research has been conducted, the tools to measure eHealth literacy lack acknowledgment of different personal backgrounds influencing the measured competencies, such as social and cultural factors. Griebel et al [19] criticize the eHealth literacy community for missing an agreement on how to measure eHealth literacy. Accordingly, it is stated that the new tools should consider the earlier research and create a well-founded theoretical basis to place eHealth literacy into broader context [19].

Objectives

Earlier reviews have focused on: (1) the definitions and measures of the concepts of health literacy [5,6,9,11,12,25,26], eHealth literacy [28,29], and critical health literacy [13,14] and their (2) operationalization in a specific demographic group, for example, adolescents [30-32] and older adults [33], or in a specific context, for example, eHealth service use [34].

This systematic review contributes to these earlier reviews by synthesizing health literacy research conducted in *Web-based information environments* and in *different disciplines*. It differs from the earlier reviews as it reviews not only the definitions of different health literacy concepts but also the measures used to operationalize these concepts in empirical studies. By elaborating remarks made in previous literature about the conceptions of health literacy, the following objectives were set:

1. To categorize thematically the definitions of health literacy and related concepts used in empirical studies focused on Web-based information environments.
2. To examine the operationalization of the concepts within these thematic categories.

Methods

Data Sources and Search Strategy

This systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [35]. The review is interpretive [36] and emphasizes the integration of studies across different disciplines to create a synthesis of the data. A search strategy was developed to identify articles examining health literacy or related concepts in a Web-based information environment. Overall, 6 academic databases were searched on April 14, 2016. The databases were Library and Information Science Abstracts, Applied Social Sciences Index and Abstracts, Education Resources Information Center, US National Library of Medicine premier bibliographic database (MEDLINE), Library and Information Science and Technology Abstracts, and the Cumulative Index to Nursing and Allied Health Literature. The search terms used covered 3 domains, "web," "health," and "literacy," including related terms. The search was limited to title and abstract and to peer-reviewed articles published in English between years 2011 and 2016. This time span was chosen to provide a sample of studies published during a period within which Web information seeking [37] and the use of social media [38] have increased considerably. This tight time frame enabled reviewing a manageable sample of studies. A broader time frame would have required a narrower search strategy. The search strategy is reported in detail in Multimedia Appendix 1.

In addition, 1 academic journal (*Computers in Human Behavior*) was searched manually as it was not indexed in the searched databases but showed potential to finding relevant articles. Search from this journal was conducted by searching with the phrase "health" AND "literac*" OR "knowledge" from article titles and abstracts and within the same time frame as the database search. This search resulted in 4 relevant articles. In

total, 1289 articles were identified through the literature search, as presented in Figure 1.

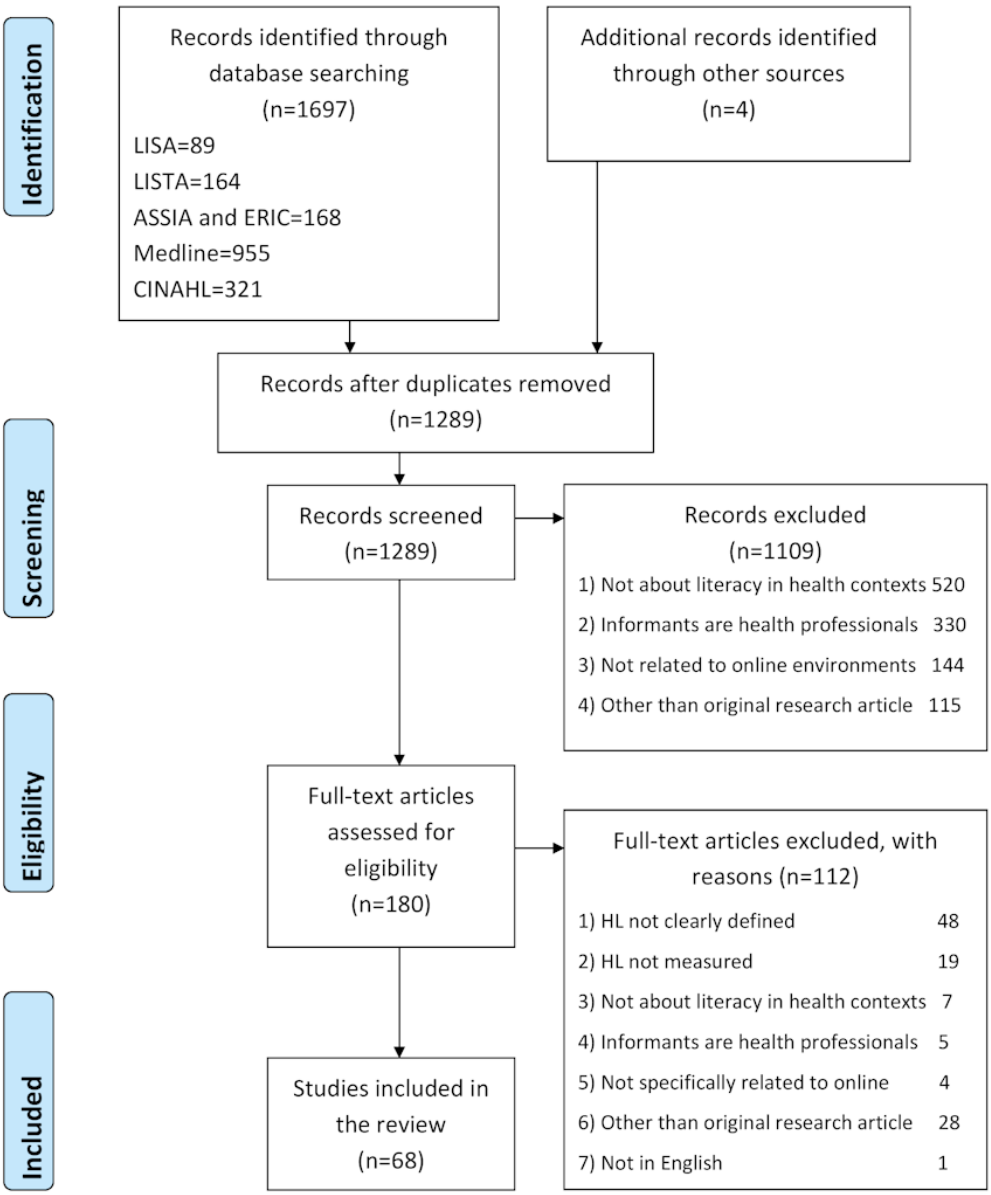
Study Selection and Extraction of Data

The screening process of the articles was 2-phased. In the first phase, the duplicates were removed and the titles and abstracts of the articles (n=1289) screened independently by the first author to identify eligible articles for full-text screening. A 10% random sample was screened by the second author with an interrater agreement rate of 93%. The articles chosen for the full-text screening had to fulfill the following inclusion criteria: (1) original, peer-reviewed full-text article written in English; (2) the topic of the article concerned literacy in the context of health; (3) informants of the study were lay people, not health professionals or students of the field; and (4) the focus of study was health literacy in a Web-based information environment. In the second phase of the selection process, 180 full-text articles were screened, 112 of which were excluded.

After the study selection process, 68 articles were included in the review. The following data were extracted from these articles:

1. Title
2. Authors
3. Publication title
4. Year of publication
5. Research area or discipline (according to the first authors' affiliation)
6. Aim or objective of the study
7. Method of data collection
8. Method of data analysis
9. Health literacy concept used
10. Definition of the concept
11. Measurement tool and its description.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the study selection process. LISA: Library and Information Science Abstracts; LISTA: Library and Information Science and Technology Abstracts; ASSIA: Applied Social Sciences Index and Abstracts; ERIC: Education Resources Information Center; CINAHL: Cumulative Index to Nursing and Allied Health Literature.



A detailed description of the study selection process is presented in the PRISMA chart (see Figure 1). The characteristics of the included studies can be found in Multimedia Appendix 2.

Results

Characteristics of the Included Studies

In total, 68 studies were included in the systematic review. The studies represent a variety of disciplines (based on the first author's affiliation), including medicine (n=13), health education and promotion or health communication (n=8), nursing (n=6), health sciences or public health (n=5), health policy (n=2), nutrition science (n=2), pharmacy (n=2), gerontology (n=1), biomedical informatics (n=1), communication or advertising (n=9), psychology (n=8), information science and information studies (n=8), sociology or social work (n=2), and behavioral sciences (n=1).

A total of 8 different health literacy concepts (Table 1) with 21 definitions (Multimedia Appendix 3) were identified. The most commonly used concepts were health literacy, which was referred to in 38 studies, and eHealth literacy, which was used in 37 studies. Other health-related literacy concepts that emerged were mental health literacy (n=3), oral health literacy (n=1), and *bad* health literacy (n=1). The concepts of health information literacy and everyday health information literacy were presented in 1 study. Refer to the study by Huhta et al [10] for a detailed description of the concepts and their definitions.

The most common method for data collection was a questionnaire survey, which was the only data collection method in 58 studies. There were 2 studies where interviews or focus groups were the only methods used. In 8 studies, several data collection methods were used. The analysis methods were predominantly quantitative (n=62). Mixed methods were applied in 4 studies and qualitative methods in 2 studies.

The included studies focused on different populations: patients or adults with risk factors for a disease (n=17), older adults or veterans (n=14), students (n=8), adults (n=8), and parents or caregivers (n=4). Other groups were participants with limited health literacy or computer literacy (n=2), middle-aged men (n=1), library users (n=1), members of an online support group (n=1), and the general public (n=12). The sample sizes ranged from 20 to 4368.

Categorization

The content analysis focused on the health literacy concepts along with their definitions and measures. On the basis of the *definitions* of the health literacy concepts identified in the included articles, the studies were grouped into 3 thematic categories: health literacy as (1) a general skill, (2) a multidimensional concept, and (3) as a domain-specific concept. The categorization is drawn from the data, and it follows remarks made on health literacy research in earlier literature [9,25]. In Table 1, the identified definitions are presented in these categories.

If several concepts were cited, the main concept of the included study was derived from the article title, or if it was not mentioned, from the abstract. A detailed description of all identified concepts and their definitions is provided in Multimedia Appendix 3.

Health Literacy as a General Skill

The definitions that describe health literacy as personal skills to utilize health information to gain better health were categorized as general and skill-based constructs. A general health literacy concept was adopted as the main concept in 23 studies. These studies referred to the health literacy definitions by Nutbeam [39], American Medical Association [40], Ratzan and Parker [41], Australian Bureau of Statistics [42], Rootman and Gordon-El-Bihbety [43], Berkman et al [44], The Patient Protection and Affordable Care Act [45], National Network of Libraries of Medicine [46], and the health information literacy definition by Shipman et al [20].

Table 1. Health literacy concepts identified in the included articles.

Thematic category	Concept	Defined by	Example of definition
General and skill-based	Health literacy; Health information literacy	Nutbeam [39], American Medical Association [40], Ratzan and Parker [41], Australian Bureau of Statistics [42], Rootman and Gordon-El-Bihbety [43], Berkman et al [44], The Patient Protection and Affordable Care Act [45], National Network of Libraries of Medicine [46]; Shipman et al [20]	Health literacy is "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions." Ratzan and Parker [41]
Multidimensional	Health literacy	Nutbeam [47], Zarcadoolas et al [48], Baker [49], Nutbeam [50], Sørensen et al [6]	"Health literacy is linked to literacy and entails people's knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course." Sørensen et al [6]
Domain-specific	eHealth literacy; Mental health literacy; Oral health literacy; *Bad* health literacy	Norman and Skinner [15], Bodie and Dutta [51], Norman [52]; Jorm et al [53]; US Department of Health and Human Services [54]; Schultz and Nakamoto [55]	"eHealth literacy is 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." Norman and Skinner [15]

The definition of health literacy as a capacity that individuals have in certain degrees by Ratzan and Parker [41] was cited in 24 studies [56-79]. Overall, it is the most often cited definition for health literacy in the included articles. Most of the articles cited a secondary source for the definition, such as that by the Healthy People 2010 initiative of the US Department of Health and Human Services [54]. The concept is process-oriented, focusing on obtaining basic health information and health services to make health decisions. A rather similar definition, but one with a wider scope including oral communication skills by Berkman et al [44] was cited in 5 studies [62,74,80-82]. This was the second most cited definition.

The health literacy definition adopted by the WHO and outlined by Nutbeam [39], stressing on both cognitive and social skills of an individual in the process of building motivation and understanding health information, was cited in 4 studies [83-86]. The health literacy definition by the American Medical Association [40] focused on individuals' skills to perform tasks on reading comprehension and numeracy. It was cited in 2 studies [87,88]. Other definitions for general health literacy were cited only once and were rather similar to each other with only minor differences. For example, the definition by Rootman and Gordon-El-Bihbety [43] includes the attribute of evaluation and presents health literacy as an ability that can be improved across the life-course.

The concept of health information literacy by the Medical Library Association [20] presents individuals' skills to recognize an information need, seek information, and use it as key competencies needed to make good health choices [20]. It was cited in 1 study [89]. In this definition, the focus is placed on the process of information seeking, described in more detail compared with the definitions for health literacy. The concept of health information literacy addresses also the individuals' ability to assess the found information critically and to evaluate its applicability to a specific situation. This critical attribute is not present in all the definitions for health literacy and related concepts [10] and thus distinguishes the concept from other, more functional health literacy definitions.

Common for these definitions of health literacy and health information literacy is the focus on individuals' abilities to obtain health information to make good health decisions. These definitions describe health literacy from 2 perspectives. First, health information is seen as general information obtained through information seeking. Second, health literacy is seen as a general skill set that an individual has to some degree and that it can be utilized universally in decision-making situations. Thus, health literacy is understood as a general, skill-based ability that can be applied to all kinds of situations that are related to health.

Operationalization of the General Health Literacy Concepts

In total, 11 studies in this category used 1 or several measurement tools with an aim to detect the functional reading skills and numeracy of the selected population (see Table 2). The most often used functional measurement tools were the Newest Vital Sign [90] used in 4 studies [57,62,73,87] and the Rapid Estimate in Adult Literacy in Medicine (REALM) [91] used in 3 studies [56,58,63]. Other measurement tools used were The Test for Functional Health Literacy in Adults (TOHFLA) [21] (cited in [59]) and its shorter version S-TOFHLA [92] (cited in [84]), Short Assessment of Health Literacy in Dutch [93] (cited in [81]), and Adult Literacy & Life Skills Survey [94] (cited in [67]). These tools were developed to detect limited health literacy among adult patients in clinical settings.

Self-efficacy measures of health literacy were used in 5 studies [66,68,70,83,96] that adopted a general health literacy concept. Of these studies, 3 used a self-efficacy measure with only few screening items. Kim [66] states that individuals with higher levels of health literacy are expected to search health-related information from the Web more efficiently, and thus, in the study, health literacy was measured by asking whether the respondents searched for health information from the Web. Lee et al [68] used a 1-item health literacy screener by Chew et al [24], and Mayberry et al [70] used a modified 3-item version of the screener. It consists of questions about reading problems and confidence in filling out medical forms [24].

Other self-efficacy measures used were a reading comprehension screener called Single Item Literacy Screener [97] (cited in [71]) and the Functional Communicative and Critical Health Literacy scale (FCCHL) [23] (cited in [61]), which is based on Nutbeam's [47] multidimensional definition of health literacy. FCCHL is a self-efficacy measure containing questions about the frequency of the patient's actions, such as how often the patient had problems to read and comprehend medical texts (functional health literacy); how often they collect information, communicate about medical conditions, and apply the found information (communicative health literacy); and how often they critically evaluate the found information (critical health literacy) [23]. Furnival et al [89] used the Everyday Health Information Literacy (EHIL) screening tool by Niemelä et al [98] to measure the study participants' health information literacy. The screening tool is based on the concept of health information literacy and was developed for studying "laypersons' general and nonprofessional abilities related to health information" [98].

In addition, 2 studies [65,80] measured health literacy with a knowledge test. Jiang and Beaudoin [65] referred to Ratzan's and Parker's [41] definition of health literacy in their study and operationally defined the concept as "one's knowledge and understanding on health-related issues." The test consists of self-reported knowledge about medical research (scientific literacy), beliefs about US tobacco regulation (civic literacy), and a numeracy section. The authors suggested that the used knowledge test aligns with the multidimensional model of health literacy developed by Zarcadoolas et al [48]. Lee et al [80] cited the health literacy definition by Berkman et al [44] and stated that health knowledge is seen as a subdimension or a proxy of health literacy. In their study, health knowledge was measured by asking respondents to indicate the plausibility of 7 health statements [44]. Other types of measures identified were a skill-based health literacy performance test [74] and qualitative assessment of health-related information literacy [86].

Table 2. Operationalization of health literacy concepts in selected studies (N=68).

Type of measure	Thematic category, n (%)		
	General and skill-based (n=23)	Multidimensional (n=6)	Domain-specific (n=39)
Reading comprehension and numeracy	11 (16.2)	0 (0)	1 (1.5)
Self-efficacy	6 (8.8)	2 (2.9)	34 (50.0)
Knowledge	2 (2.9)	0 (0)	3 (4.4)
Performance tasks	1 (1.5)	0 (0)	0 (0)
Qualitative assessment	1 (1.5)	0 (0)	0 (0)
Several	2 (2.9)	4 (5.9)	1 (1.5)

Moreover, 2 studies [64,99] used several types of measures to assess general health literacy. In both studies, health literacy is defined as a skill-based construct, and it is assessed with reading comprehension and self-efficacy measures [64] or additionally also with a knowledge test [99]. For example, in a study by Woods et al [99], the study participants completed 11 different questionnaires that measured health knowledge, health literacy, and internet and computer skills. In 1 study [86], a qualitative assessment of health and information literacy was conducted.

Almost all the studies that adopted a general health literacy concept screen participants' internet use [56,57,59,61-68,70,71,74,77,80,84,86,87], usually with a simple yes or no question. In 4 studies [60,61,66,68] computer or internet literacy was measured, although in 2 of these, this means screening the internet use of the participants. In fewer cases, measures also included access to internet [59,63,87], skills [56,68,85,100] or comfort [70] to use internet or a computer, and abilities to communicate with peer or health professionals and providers in the Web [57,68]. In 3 studies [58,73,81] internet, computer, or technology-related measures were not included.

Health Literacy as a Multidimensional Concept

Models that include several attributes, such as the social factors and cultural context into the definitions of health literacy, were categorized as multidimensional health literacy concepts. For example, the critical appraisal of found information is taken into account more thoroughly in these models. These multidimensional health literacy definitions and models by Nutbeam [47,50], Baker [49], Zarcadoolas et al [48], and Sørensen et al [6] were cited in 9 studies, the last 2 being the most used. In total, 6 studies chose the multidimensional construct as the main health literacy concept.

The health literacy definition by Zarcadoolas et al [48] was cited in 3 studies [72,83,88]. The definition includes the notion of health literacy as a lifelong learning process and sets the outcome of acquiring health literacy skills as an improved quality of life. This definition presents health and health literacy as the lifelong projects of *people*, not individuals. The model complementing the definition of health literacy by Zarcadoolas et al [48] is built around 4 central domains of literacy: fundamental, scientific, civic, and cultural. Of these, especially the domain of civic literacy represents the sociocultural aspect of literacy, as it includes "[u]nderstanding the relationship between one's actions and the larger social group." The civic literacy domain also stresses critical media literacy skills that include, for example, awareness of possible *biased authorities* in consumer advertising [48].

The health literacy definition by Sørensen et al [6] was cited in 3 studies [75,88,101]. Sørensen et al [6] reviewed health literacy research and created an integrated model with 6 dimensions of health literacy: (1) competence, skills, and abilities; (2) actions; (3) information and resources; (4) objective; (5) context; and (6) time. The definition considers individual capabilities, but it also aims to address the public health perspective [6].

Baker's [49] conceptual model of health literacy was cited in 2 studies [102,103]. It presents several domains that affect health literacy. In the model, prior knowledge, such as vocabulary and conceptual knowledge of health together with reading fluency, is seen as a resource for an individual for facilitating health literacy. Health-related print and oral literacy are seen as dimensions of holistic health literacy that can lead to improved health outcomes. In addition, influencing factors, such as culture and norms, and barriers, such as limited access to health care, can have an effect on health behavior change [49].

Nutbeam [47] continued his examination on health literacy by broadening the definition into a conceptual model. The model consists of 3 literacy concepts: functional health literacy relates to health education and learning of factual information on health risks and on how to use the health system. Interactive health literacy concerns improving personal capacity to act independently on knowledge. Critical health literacy regards cognitive and skills development outcomes that support effective social and political action. According to Nutbeam [47], the first 2 literacy dimensions are effective on an individual level, but the third can also be seen linked to population level benefits. The model is developed to address the challenges for health education, and therefore, it presents health literacy as an outcome of health promotion. In his more recent article, Nutbeam [50] suggests that instead of conceptualizing health literacy as a risk factor influencing clinical outcomes, it should be seen as an asset that can support individual and population level health outcomes, when improved through patient education.

Operationalization of the Multidimensional Health Literacy Concepts

In total, 6 studies [75,83,88,101,103,108] adopted a multidimensional health literacy concept as the central concept of the study. The operationalization of these concepts varied, and several types of measures were used, as seen in Table 2.

Rowsell et al [101] referred to the multidimensional health literacy definitions by Sørensen et al [6] and Nutbeam [50] and evaluated the level of health literacy with a single-item self-efficacy measure by Chew et al [24] with the aim to detect patients' difficulties in understanding written information. On the other hand, van der Vaart et al [103] adopted Baker's [49] health literacy definition as their main literacy concept and measured it with the FCCHL self-efficacy scale that includes several literacy domains.

In 4 studies, several types of measures were used. In a study by Tam et al [75], the combination of measures included a reading comprehension and numeracy measure the Rapid Estimate of Adult Literacy in Medicine and Dentistry measure (REALMD-20) [107], a 2-item self-efficacy measure by Chew et al [24], and a dental health knowledge test. In this study, oral health literacy was measured, although the authors did not provide a clear definition of the concept itself. Instead, the health literacy definition by Sorensen et al [6] and the concept of eHealth literacy [15] were discussed. In other studies that adopted a multidimensional health literacy concept, reading comprehension and numeracy [88], self-efficacy [83,88,108], knowledge [88,108], and performance [83] were measured.

Computer or internet literacy was not measured in studies that adopted a multidimensional concept of health literacy. Instead, internet access [103,108] and use [103] were screened. Subramania et al [88] included internet-related questions to their overall assessment of health literacy skills of the participants. Moreover, 3 studies [75,83,101] did not include any kinds of internet- or computer-related measures to their study.

Health Literacy as a Domain-Specific Concept

The health literacy concepts that focus on a specific context or target a specific patient group are categorized as domain-specific concepts of health literacy. In total, a domain-specific concept of health literacy was cited in 41 of the included studies. Of these, eHealth literacy by Norman and Skinner [15], Bodie and Dutta's [51] elaboration of the same concept, and Norman's [52] suggestion of eHealth literacy 2.0 definition are essentially targeted to address health literacy in Web contexts. Of these, Norman's and Skinner's definition was the most often cited definition in included studies. In several studies (n=11), in addition to eHealth literacy, also other health literacy concepts and definitions were discussed (see Multimedia Appendix 2).

In total, 39 studies adopted a domain-specific health literacy concept as the main concept of the study. In most of these studies (n=34), the main concept was eHealth literacy [60,69,72,76-79,82,85,95,96,100-102,104-106,109-126]. The concept of eHealth literacy is accompanied by the Lily model that consists of 6 literacies organized in 2 central types: analytic (traditional, media, and information) and context-specific (computer, scientific, and health). The analytic literacy types are described as skills that are applicable to a wide range of information sources [15]. The context-specific types involve skills that are applied in specific situations. According to Norman and Skinner [15], all these skills are required when engaging with electronic sources. In the definition of eHealth literacy, the electronic element of health information seeking seems to be addressed as a contrast to nonelectronic information seeking, although a deeper explanation of those electronic sources is absent in the definition [15].

Bodie and Dutta [51] present an elaborated definition for eHealth literacy that stresses the significance of the Web context in seeking, evaluating, and using health information. This definition was presented in 1 study [114]. Norman's [52] definition for eHealth literacy 2.0 was presented in 1 study [126]. With the definition, Norman attempts to emphasize the context of social media regarding eHealth literacy screening tool development by presenting social media relevant tasks and skills to the concept [52].

Other domain-specific health literacy concepts identified in the studies were mental health literacy used in 3 studies [127-129], oral health literacy used in 1 study [130], and *bad* health literacy used in 1 study [31]. The definition of mental health literacy by Jorm et al [53], unlike other health literacy definitions, also addresses beliefs and attitudes toward health issues. The definition of oral health literacy by the US Department of Health and Human Services [54] is based on the health literacy definition by Ratzan and Parker [41] and thus takes a skill-based approach to the concept. The concept of *bad* health literacy originally introduced by Schulz and Nakamoto [55] refers, according to Allam et al [131], to "the presence of the ability to understand medical information turned sour by the simultaneous absence of the ability to recognize it as false." In other words, the information seeker might be literate enough to find, understand, and process even low-quality information, obtained, for example, from electronic sources but is incapable to recognize it as false, irrelevant, or fraudulent [131].

Operationalization of the Domain-Specific Health Literacy Concepts

Within the studies that adopted a domain-specific concept as the main health literacy concept (n=39), the operationalization is more often done with a self-efficacy measurement tool than other types of measures, as seen in Table 2.

Most of the studies that adopted eHealth literacy as the main concept used the eHealth Literacy Scale (eHEALS) by Norman and Skinner [132] as the main measurement tool. In total, the eHEALS is used in 29 of the 39 studies in this category and as the only used tool in 25 of them [60,72,76,77,79,82,85,95,96,100,102,104-106,113,115-123,125]. The 8-item eHEALS scale aims to measure "consumers' combined knowledge, comfort, and perceived skills at finding, evaluating, and applying electronic health information to health problems." The scale is proposed to address the 6 literacy types of the Lily model [15]. In the included studies, the eHEALS is described in different ways. Typically, the scale is described as a measurement that detects consumers' perceived information technology or computer skills. In addition, the abilities to seek health information from the Web are seen as central attributes of the scale. Other studies that adopted the eHealth literacy as the main concept of the study also used other self-efficacy measures, such as EHIL [98] (cited in [110]) and Brief Health Literacy Screening Tool BRIEF [133] (cited in [112]). In addition, 2 studies [69,114] present a new eHealth literacy measure. Hsu et al [114] discuss eHealth literacy definitions by Norman and

Skinner [15] and Bodie and Dutta [51] and present a new eHealth literacy measure eHL that seeks to detect individuals' "ability to seek, find, understand, and evaluate health information from electronic sources and apply this knowledge to address or solve a health problem" [114]. The self-efficacy measure eSEARCH, eHealth Literacy Tool used in a study by Manafò et al [69], was developed to measure eHealth literacy skills of older adults.

Other types of measures used in the included articles that adopted eHealth literacy as the main concept were performance tests [109,126]; combined measures of reading comprehension, numeracy, and knowledge [124]; and self-efficacy [78]. In 1 study [111], eHealth literacy was assessed qualitatively based on focus group discussions of the participants.

Mental health literacy was measured in 3 studies [127-129] and oral health literacy [130] and *bad* health literacy [131] both in 1 study. In 2 of the studies that focused on mental health literacy [128,129], the concept was operationalized by measuring the participants' knowledge about and attitudes toward mental health issues. Li et al [127] used several types of measures. The 31-item questionnaire consists of questions about the participant's knowledge and self-efficacy on mental health issues. In a study by Tse et al [130], oral health literacy was measured with READ-30 [134], a word recognition instrument that requires participants to read aloud 30 oral health–related words. Allam et al [131] measured *bad* health literacy with a knowledge test focused on vaccine information.

Discussion

Principal Findings

The aim of this systematic review was to identify health literacy concepts and their definitions and operationalization in studies focused on Web-based information environments. The concept of eHealth literacy by Norman and Skinner [15] was used most often. However, the concept of health literacy was also used and a variety of definitions were presented for it in the selected studies. On the basis of the definitions for health literacy, 3 thematic categories were identified, namely, general and skill-based, multidimensional, and domain-specific. Most studies adopted a domain-specific concept, followed by the ones that used a general and skill-based concept. Multidimensional concepts occurred least frequently.

The general concept of health literacy was typically operationalized by using reading comprehension and numeracy measures. In turn, the domain-specific concepts were most often operationalized by using a self-efficacy measure. Several types of measures were used in studies that adopted multidimensional constructs of health literacy. Nevertheless, inconsistencies in the operationalization of the different concepts were identified.

Comparison With Prior Work

The lack of consensus in defining health literacy, as presented in several reviews [6,11,30], is supported by the results of this systematic review as several different definitions for the concept were identified in the included studies. The modern health literacy definitions are more often multidimensional than functional [3,9]. However, this systematic review shows that there is a tendency to refer to the early definitions of health literacy, which present a functional understanding of the concept. Within the studies that applied the concept of eHealth literacy, a more consistent understanding of the definition was detected as only 2 definitions for the concept were presented.

As earlier reviews indicate, the currently used measures of health literacy have focused on assessing individuals' reading comprehension and understanding of medical texts in clinical contexts [5,135]. In addition, within the studies conducted in Web-based information environments, general health literacy was measured with a widely used and validated functional measurement tool, although there are more recent and multidimensional measures available [25]. Pleasant et al [135] argue that the focus on measuring only the functional skills of individuals leaves important factors such as individual information and communication skills untested. Despite the trend of understanding health literacy as a multidimensional construct including contextual, cultural, and social factors [5], these were not acknowledged in the studies included in this systematic review.

The concept of eHealth literacy by Norman and Skinner [15] was clearly the most used concept in the included studies. As a domain-specific concept, eHealth literacy aims to address especially the literacy skills needed in Web contexts. However, in the included studies, the concept was described as the technological skills of the study subjects. Yet, it is clear that eHealth literacy competencies are more varied than the mere ability to use the internet or a computer efficiently. Addressing literacy skills or practices through domain-specific concepts offers an opportunity to express domain-specific issues, such as the importance of the technological skills as part of eHealth literacy competencies, or oral health knowledge as part of oral health literacy. However, the development of these concepts may be challenging, as the focus of research is fragmented in empirical studies and the conceptual development is scarce (See also [8]).

Measurement of eHealth literacy is more often focused on assessing the self-reported skills of individuals. Unlike in the systematic review by Karnoe and Kayser [28], dual-design eHealth literacy measures are not common in studies conducted in Web-based information environments, as only few studies included internet or digital literacy measures in their health literacy screening tools.

The trend toward mixing different measuring types, as indicated by Altin et al [25], was noted also within the studies conducted in Web-based information environments. The focus on clinical settings as a study context was not as clearly indicated as in the earlier reviews, and usually, the sample population was a certain age instead of patients.

Strengths and Limitations

To our knowledge, this study is among the first cross-disciplinary reviews of health literacy concepts, definitions, and their operationalization in Web contexts. The systematic process of this review enabled thorough investigation of the health literacy–related academic research focused on the context of Web-based information environments. The main

limitations of this review lie within the search strategy. Only studies written in English were included in the review, which excluded relevant studies in other languages. In addition, some studies may have been missed due to the restricted search terms and limited time frame.

Conclusions

This systematic review identified health literacy concepts, definitions, and operationalization used in research focusing on Web-based information environments. On the basis of the results, several concepts are being used, eHealth literacy and health literacy being the most common ones. In addition, 3 thematic categories of the different definitions were identified: general and skill-based, multidimensional, and domain-specific. Typically, general and skill-based health literacy was measured with reading comprehension or numeracy tests and domain-specific health literacy with self-efficacy tests. Multidimensional concepts were used less often and operationalized by using several types of measures. Future studies conducted in Web contexts should place emphasis on the conceptual development of health literacy. Researchers are encouraged to provide clear operationalization for the concepts they use to ensure transparency in reporting.

Acknowledgments

This study is part of the Academy of Finland project Cognitive Authorities in Everyday Health Information Environments of Young People (#2430227313) and funded by the Eudaimonia Doctoral Program in the University of Oulu, Finland.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 197KB - jmir_v20i12e10273_app1.pdf]

Multimedia Appendix 2

Characteristics of the included studies.

[PDF File (Adobe PDF File), 447KB - jmir_v20i12e10273_app2.pdf]

Multimedia Appendix 3

Health literacy concepts and definitions in the included studies.

[PDF File (Adobe PDF File), 384KB - jmir_v20i12e10273_app3.pdf]

References

1. Bawden D, Robinson L. The dark side of information: overload, anxiety and other paradoxes and pathologies. J Inf Sci 2009 Apr;35(2):180-191. [doi: 10.1177/0165551508095781]
2. Koltay T. The bright side of information: ways of mitigating information overload. J Doc 2017 Jul 10;73(4):767-775. [doi: 10.1108/JD-09-2016-0107]
3. Batterham RW, Beauchamp A, Osborne RH. Health Literacy. In: Quah SR, editor. International Encyclopedia Of Public Health, Second Edition. Kidlington, Oxford: Elsevier; 2017:428-437.
4. World Health Organization. Track 2: Health literacy and health behavior URL: http://www.who.int/healthpromotion/conferences/7gchp/track2/en/ [accessed 2018-02-06] [WebCite Cache ID 6x1poUDyA]
5. Guzys D, Kenny A, Dickson-Swift V, Threlkeld G. A critical review of population health literacy assessment. BMC Public Health 2015 Mar 04;15:215 [FREE Full text] [doi: 10.1186/s12889-015-1551-6] [Medline: 25885742]
6. Sørensen K, Van den Brouck S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
7. McCormack L, Haun J, Sørensen K, Valerio M. Recommendations for advancing health literacy measurement. J Health Commun 2013;18(Suppl 1):9-14 [FREE Full text] [doi: 10.1080/10810730.2013.829892] [Medline: 24093340]
8. Mackert M, Champlin S, Su Z, Guadagno M. The many health literacies: advancing research or fragmentation? Health Commun 2015;30(12):1161-1165. [doi: 10.1080/10410236.2015.1037422] [Medline: 26372026]
9. Mårtensson L, Hensing G. Health literacy – a heterogeneous phenomenon: a literature review. Scand J Caring Sci 2012 Mar;26(1):151-160. [doi: 10.1111/j.1471-6712.2011.00900.x] [Medline: 21627673]

10. Huhta AM, Hirvonen N, Huotari ML. Concepts related to health literacy in online information environments: a systematic review with an emphasis on approach to information. In: Information Literacy in the Workplace. Cham: Springer; 2018 Presented at: 5th European Conference on Information Literacy; Sep 18-21, 2017; Saint Malo, France p. 460-469. [doi: 10.1007/978-3-319-74334-9_48]
11. 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] [Medline: 21724626]
12. Mancuso JM. Assessment and measurement of health literacy: an integrative review of the literature. Nurs Health Sci 2009 Mar;11(1):77-89. [doi: 10.1111/j.1442-2018.2008.00408.x] [Medline: 19298313]
13. Chinn D. Critical health literacy: a review and critical analysis. Soc Sci Med 2011 Jul;73(1):60-67. [doi: 10.1016/j.socscimed.2011.04.004] [Medline: 21640456]
14. Sykes S, Wills J, Rowlands G, Popple K. Understanding critical health literacy: a concept analysis. BMC Public Health 2013 Feb 18;13:150 [FREE Full text] [doi: 10.1186/1471-2458-13-150] [Medline: 23419015]
15. Norman CD, Skinner HA. eHealth literacy: Essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
16. Eng TR. 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.
17. Bautista J. From solving a health problem to achieving quality of life: redefining eHealth literacy. J Lit and Technol 2015;16(2):33-54 [FREE Full text]
18. Gilstad H. Toward a comprehensive model of eHealth literacy. In: Proceedings of the 2nd European Workshop on Practical Aspects of Health Informatics (PAHI 2014).: CEUR Workshop Proceedings; 2014 Presented at: 2nd European Workshop on Practical Aspects of Health Informatics; May 19, 2014; Trondheim, Norway.
19. Griebel L, Enwald H, Gilstad H, Pohl A, Moreland J, Sedlmayr M. eHealth literacy research-Quo vadis? Inform Health Soc Care 2018 Dec;43(4):427-442. [doi: 10.1080/17538157.2017.1364247] [Medline: 29045164]
20. Shipman JP, Kurtz-Rossi S, Funk CJ. The health information literacy research project. J Med Libr Assoc 2009 Oct;97(4):293-301 [FREE Full text] [doi: 10.3163/1536-5050.97.4.014] [Medline: 19851494]
21. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med 1995 Oct;10(10):537-541. [Medline: 8576769]
22. Haun JN, Valerio MA, McCormack LA, Sørensen K, Paasche-Orlow MK. Health literacy measurement: an inventory and descriptive summary of 51 instruments. J Health Commun 2014;19(Suppl 2):302-333. [doi: 10.1080/10810730.2014.936571] [Medline: 25315600]
23. Ishikawa H, Takeuchi T, Yano E. Measuring functional, communicative, and critical health literacy among diabetic patients. Diabetes Care 2008 May;31(5):874-879. [doi: 10.2337/dc07-1932] [Medline: 18299446]
24. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
25. Altin SV, Finke I, Kautz-Freimuth S, Stock S. The evolution of health literacy assessment tools: a systematic review. BMC Public Health 2014;14:1207 [FREE Full text] [doi: 10.1186/1471-2458-14-1207] [Medline: 25418011]
26. Gonçalves D, Ricci-Cabello I, Ziebland S, Valderas J. An overview of self-administered health literacy instruments. PLoS One 2014;9(12):e109110 [FREE Full text] [doi: 10.1371/journal.pone.0109110] [Medline: 25478813]
27. Nguyen TH, Paasche-Orlow MK, McCormack LA. The state of the science of health literacy measurement. Stud Health Technol Inform 2017;240:17-33. [Medline: 28972507]
28. Karnoe A, Kayser L. How is eHealth literacy measured and what do the measurements tell us? A systematic review. Knowl Manag E-Learn Int J 2015 Dec 26;7(4):576-600 [FREE Full text]
29. Pohl A, Griebel L, Trill R. Contemporary eHealth literacy research. An overview with focus on Germany. In: Proceedings of the 3rd European Workshop on Practical Aspects of Health Informatics (PAHI 2015).: CEUR Workshop Proceedings; 2016 Presented at: 3rd European Workshop on Practical Aspects of Health Informatics; Oct 27; Elgin, Scotland, UK.
30. Ormshaw MJ, Paakkari LT, Kannas LK. Measuring child and adolescent health literacy: a systematic review of literature. Health Education 2013 Aug 23;113(5):433-455. [doi: 10.1108/HE-07-2012-0039]
31. Bröder J, Okan O, Bauer U, Bruland D, Schlupp S, Bollweg TM, et al. Health literacy in childhood and youth: a systematic review of definitions and models. BMC Public Health 2017 Dec 26;17(1):361 [FREE Full text] [doi: 10.1186/s12889-017-4267-y] [Medline: 28441934]
32. Okan O, Lopes E, Bollweg TM, Bröder J, Messer M, Bruland D, et al. Generic health literacy measurement instruments for children and adolescents: a systematic review of the literature. BMC Public Health 2018 Dec 22;18(1):166 [FREE Full text] [doi: 10.1186/s12889-018-5054-0] [Medline: 29357867]
33. Watkins I, Xie B. eHealth literacy interventions for older adults: a systematic review of the literature. J Med Internet Res 2014;16(11):e225 [FREE Full text] [doi: 10.2196/jmir.3318] [Medline: 25386719]
34. Kim H, Xie B. Health literacy in the eHealth era: a systematic review of the literature. Patient Educ Couns 2017 Jun;100(6):1073-1082. [doi: 10.1016/j.pec.2017.01.015] [Medline: 28174067]
35. 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] [Medline: 19631508]

36. Harden A, Thomas J. Mixed methods systematic reviews: examples and emerging issues. In: Tashakkori A, Teddlie C, editors. SAGE handbook of mixed methods in social & behavioral research. 2nd ed. Thousand Oaks, California: SAGE; 2010:749-774.
37. Fox S, Duggan M. Pew Research Center. 2013. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2018-08-28] [WebCite Cache ID 720KD06v1]
38. Pew Research Center. Social Media Fact Sheet 2018 URL: http://www.pewinternet.org/fact-sheet/social-media/ [accessed 2018-08-28] [WebCite Cache ID 720MwFtrn]
39. Nutbeam D. Health promotion glossary. Health Promot Int 1998 Jan 01;13(4):349-364. [doi: 10.1093/heapro/13.4.349]
40. Health literacy: report of the Council on Scientific Affairs. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. J Am Med Assoc 1999 Feb 10;281(6):552-557. [Medline: 10022112]
41. Ratzan S, Parker R. National Library of Medicine Current Bibliographies in Medicine: Health Literacy. Washington, DC: National Institutes of Health, US Department of Health and Human Services; 2000.
42. Australian Bureau of Statistics. 4233.0 - Health Literacy, Australia, 2006. Canberra: Australian Bureau of Statistics; 2006.
43. Rootman I, Gordon-El-Bihbety D. A vision for health literate Canada: report of the expert panel on health literacy. Ottawa, Ontario, Canada: Canadian Public Health Association; 2006.
44. Berkman ND, Davis TC, McCormack L. Health literacy: what is it? J Health Commun 2010;15(Suppl 2):9-19. [doi: 10.1080/10810730.2010.499985] [Medline: 20845189]
45. The Patient Protection Affordable Care Act (PPACA). Congress. 2010. Public Law 111–148 URL: https://www.congress.gov/111/plaws/publ148/PLAW-111publ148.pdf [accessed 2018-11-30] [WebCite Cache ID 74JrZuk38]
46. National Network of Libraries of Medicine. 2018. Health Literacy URL: https://nnlm.gov/initiatives/topics/health-literacy [accessed 2018-10-01] [WebCite Cache ID 72qMDhSsf]
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]
48. Zarcadoolas C, Pleasant A, Greer DS. Understanding health literacy: an expanded model. Health Promot Int 2005 Jun;20(2):195-203 [FREE Full text] [doi: 10.1093/heapro/dah609] [Medline: 15788526]
49. Baker DW. The meaning and the measure of health literacy. J Gen Intern Med 2006 Aug;21(8):878-883 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00540.x] [Medline: 16881951]
50. Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008 Dec;67(12):2072-2078. [doi: 10.1016/j.socscimed.2008.09.050] [Medline: 18952344]
51. 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] [Medline: 18935884]
52. Norman C. 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] [Medline: 22193243]
53. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": a survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. Med J Aust 1997 Feb 17;166(4):182-186. [Medline: 9066546]
54. US Department of Health Human Services. Healthy People 2010: Understanding and Improving Health, 2nd ed. Washington, DC: US Government Printing Office; 2010:21-39.
55. Schulz P, Nakamoto K. "Bad" literacy, the internet, and the limits of patient empowerment. In: AAAI Spring Symposium Series: Artificial Intelligence & Health Communication. 2011 Presented at: AAAI Spring Symposium Series: Artificial Intelligence & Health Communication; Mar 21-23, 2011; Palo Alto, CA.
56. Agree EM, King AC, Castro CM, Wiley A, Borzekowski DL. "It's got to be on this page": age and cognitive style in a study of online health information seeking. J Med Internet Res 2015;17(3):e79 [FREE Full text] [doi: 10.2196/jmir.3352] [Medline: 25831483]
57. Bailey SC, O'Conor 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] [Medline: 25363660]
58. Bickmore TW, Utami D, Matsuyama R, Paasche-Orlow MK. Improving access to online health information with conversational agents: a randomized controlled experiment. J Med Internet Res 2016;18(1):e1 [FREE Full text] [doi: 10.2196/jmir.5239] [Medline: 26728964]
59. Chisolm DJ, Hardin DS, McCoy KS, Johnson LD, McAlearney AS, Gardner W. Health literacy and willingness to use online health information by teens with asthma and diabetes. Telemed J E Health 2011 Nov;17(9):676-682 [FREE Full text] [doi: 10.1089/tmj.2011.0037] [Medline: 21943161]
60. Chung S, Nahm E. 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] [Medline: 25783223]
61. Cnossen IC, van Uden-Kraan CF, Eerenstein SE, Jansen F, Witte BI, Lacko M, et al. An online self-care education program to support patients after total laryngectomy: feasibility and satisfaction. Support Care Cancer 2016 Mar;24(3):1261-1268 [FREE Full text] [doi: 10.1007/s00520-015-2896-1] [Medline: 26306518]

62. Crook B, Stephens KK, Pastorek AE, Mackert M, Donovan EE. Sharing health information and influencing behavioral intentions: The role of health literacy, information overload, and the Internet in the diffusion of healthy heart information. Health Commun 2016;31(1):60-71. [doi: 10.1080/10410236.2014.936336] [Medline: 25668744]
63. Fagnano M, Halterman JS, Conn KM, Shone LP. Health literacy and sources of health information for caregivers of urban children with asthma. Clin Pediatr (Phila) 2012 Mar;51(3):267-273 [FREE Full text] [doi: 10.1177/0009922811421413] [Medline: 21911409]
64. Ghaddar SF, Valerio MA, Garcia CM, Hansen L. Adolescent health literacy: the importance of credible sources for online health information. J Sch Health 2012 Jan;82(1):28-36. [doi: 10.1111/j.1746-1561.2011.00664.x] [Medline: 22142172]
65. Jiang S, Beaudoin CE. Health literacy and the internet: an exploratory study on the 2013 HINTS survey. Comput Human Behav 2016;58:240-248. [doi: 10.1016/j.chb.2016.01.007]
66. Kim Y. Does online searching cause or enforce health information disparity? J Info Know Mgmt 2013 Dec;12(04):1350032. [doi: 10.1142/S0219649213500329]
67. Kobayashi LC, Wardle J, von Wagner C. Internet use, social engagement and health literacy decline during ageing in a longitudinal cohort of older English adults. J Epidemiol Community Health 2015 Mar;69(3):278-283 [FREE Full text] [doi: 10.1136/jech-2014-204733] [Medline: 25428933]
68. Lee YJ, Boden-Albala B, Jia H, Wilcox A, Bakken S. The association between online health information-seeking behaviors and health behaviors among hispanics in New York city: a community-based cross-sectional study. J Med Internet Res 2015;17(11):e261 [FREE Full text] [doi: 10.2196/jmir.4368] [Medline: 26611438]
69. Manafò E, Wong S. eSEARCH©: a tool to promote the eHealth literacy skills of older adults. J Consum Health Internet 2013 Jul;17(3):255-271. [doi: 10.1080/15398285.2013.812914]
70. Mayberry LS, Kripalani S, Rothman RL, Osborn CY. Bridging the digital divide in diabetes: family support and implications for health literacy. Diabetes Technol Ther 2011 Oct;13(10):1005-1012 [FREE Full text] [doi: 10.1089/dia.2011.0055] [Medline: 21718098]
71. McCleary-Jones V, Scheideman-Miller C, Rev DJ, Johnson B, Overall M, Dwyer K. Health information technology use and health literacy among community-dwelling African Americans. ABNF J 2013;24(1):10-16. [Medline: 23589966]
72. Paek H, Hove T. Social cognitive factors and perceived social influences that improve adolescent eHealth literacy. Health Commun 2012;27(8):727-737. [doi: 10.1080/10410236.2011.616627] [Medline: 22452551]
73. Smith SG, O'Conor R, Aitken W, Curtis LM, Wolf MS, Goel MS. Disparities in registration and use of an online patient portal among older adults: findings from the LitCog cohort. J Am Med Inform Assoc 2015 Jul;22(4):888-895 [FREE Full text] [doi: 10.1093/jamia/ocv025] [Medline: 25914099]
74. Sun X, Chen J, Shi Y, Zeng Q, Wei N, Xie R, et al. Measuring health literacy regarding infectious respiratory diseases: a new skills-based instrument. PLoS One 2013;8(5):e64153 [FREE Full text] [doi: 10.1371/journal.pone.0064153] [Medline: 23724029]
75. Tam A, Yue O, Atchison KA, Richards JK, Holtzman JS. The association of patients' oral health literacy and dental school communication tools: a pilot study. J Dent Educ 2015 May;79(5):530-538 [FREE Full text] [Medline: 25941146]
76. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
77. Xie B. Older adults, e-health literacy, and collaborative learning: an experimental study. J Am Soc Inf Sci 2011 Mar 14;62(5):933-946. [doi: 10.1002/asi.21507]
78. Xie B. Improving older adults' e-health literacy through computer training using NIH online resources. Libr Inf Sci Res 2012 Jan 01;34(1):63-71 [FREE Full text] [doi: 10.1016/j.lisr.2011.07.006] [Medline: 22639488]
79. Xie B. Effects of an eHealth literacy intervention for older adults. J Med Internet Res 2011 Nov;13(4):e90 [FREE Full text] [doi: 10.2196/jmir.1880] [Medline: 22052161]
80. Lee C, Niederdeppe J, Freres D. Socioeconomic disparities in fatalistic beliefs about cancer prevention and the Internet. J Commun 2012 Dec;62(6):972-990 [FREE Full text] [doi: 10.1111/j.1460-2466.2012.01683.x] [Medline: 25530627]
81. Meppelink CS, Bol N. Exploring the role of health literacy on attention to and recall of text-illustrated health information: an eye-tracking study. Comput Human Behav 2015 Jul;48:87-93. [doi: 10.1016/j.chb.2015.01.027]
82. Wenhong C, Kye-Hyoung L. More than search? Informational and participatory eHealth behaviors. Comput Human Behav 2014 Jan;30:103-109. [doi: 10.1016/j.chb.2013.07.028]
83. Austvoll-Dahlgren A, Bjørndal A, Odgaard-Jensen J, Helseth S. Evaluation of a web portal for improving public access to evidence-based health information and health literacy skills: a pragmatic trial. PLoS One 2012;7(5):e37715 [FREE Full text] [doi: 10.1371/journal.pone.0037715] [Medline: 22701531]
84. Echt KV, Burridge AB. Predictors of reported Internet use in older adults with high and low health literacy: the role of socio-demographics and visual and cognitive function. Phys Occup Ther Geriatr 2011 Mar 02;29(1):23-43. [doi: 10.3109/02703181.2010.547657]
85. Manafò E, Wong S. Assessing the eHealth literacy skills of older adults: a preliminary study. J Consum Health Internet 2012 Oct;16(4):369-381. [doi: 10.1080/15398285.2012.701163]

86. Marshall A, Henwood F, Guy ES. Information and health literacy in the balance: findings from a study exploring the use of ICTs in weight management. Library Trends 2012;60(3):479-496. [doi: 10.1353/lib.2012.0006]
87. Gutierrez N, Kindratt TB, Pagels P, Foster B, Gimpel NE. Health literacy, health information seeking behaviors and internet use among patients attending a private and public clinic in the same geographic area. J Community Health 2014 Feb;39(1):83-89. [doi: 10.1007/s10900-013-9742-5] [Medline: 23900880]
88. Subramaniam M, St. Jean B, Taylor NG, Kodama C, Follman R, Casciotti D. Bit by bit: using design-based research to improve the health literacy of adolescents. JMIR Res Protoc 2015 May 29;4(2):e62 [FREE Full text] [doi: 10.2196/resprot.4058] [Medline: 26025101]
89. Furnival A, Silva-Jerez N. The general public's access and use of health information: a case study from Brazil. In: Proceedings of ISIC, the Information Behaviour Conference Part 1. 2014 Presented at: Information Seeking in Context: The Information Behavior Conference ISIC; Sept 2-5; Leeds, UK p. 2-5 URL: http://InformationR.net/ir/19-4/isic/isic15.html
90. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
91. Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med 1993 Jun;25(6):391-395. [Medline: 8349060]
92. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. Patient Educ Couns 1999 Sep;38(1):33-42. [Medline: 14528569]
93. Pander MH, Essink-Bot M, Leenaars KE, Fransen MP. A short assessment of health literacy (SAHL) in the Netherlands. BMC Public Health 2014 Sep 23;14:990 [FREE Full text] [doi: 10.1186/1471-2458-14-990] [Medline: 25246170]
94. Thorn W. OECD Education Working Papers, No. 26. Paris: OECD Publisher; 2009. International Adult Literacy and Basic Skills Surveys in the OECD Region URL: https://www.oecd-ilibrary.org/docserver/221351213600.pdf?expires=1542639491&id=id&accname=guest&checksum=F651F2E595F3CEF9FF360A18B5EFAB2B [accessed 2018-11-19]
95. Hu X, Bell RA, Kravitz RL, Orrange S. The prepared patient: information seeking of online support group members before their medical appointments. J Health Commun 2012;17(8):960-978. [doi: 10.1080/10810730.2011.650828] [Medline: 22574697]
96. 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] [Medline: 26715363]
97. Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. BMC Fam Pract 2006 Mar 24;7:21 [FREE Full text] [doi: 10.1186/1471-2296-7-21] [Medline: 16563164]
98. Niemelä R, Ek S, Eriksson-Backa K, Huotari ML. A screening tool for assessing everyday health information literacy. Int J Lib Inf Sci 2012;62(2):125-134. [doi: 10.1515/libri-2012-0009]
99. Woods SP, Iudicello JE, Morgan EE, Cameron MV, Doyle KL, Smith TV, HIV Neurobehavioral Research Program (HNRP) Group. Health-related everyday functioning in the internet age: HIV-associated neurocognitive disorders disrupt online pharmacy and health chart navigation skills. Arch Clin Neuropsychol 2016 Mar;31(2):176-185 [FREE Full text] [doi: 10.1093/arclin/acv090] [Medline: 26743327]
100. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
101. Rowsell A, Muller I, Murray E, Little P, Byrne CD, Ganahl K, et al. Views of people with high and low levels of health literacy about a digital intervention to promote physical activity for diabetes: a qualitative study in five countries. J Med Internet Res 2015 Oct 12;17(10):e230 [FREE Full text] [doi: 10.2196/jmir.4999] [Medline: 26459743]
102. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth literacy scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
103. van der Vaart R, Drossaert CH, Taal E, van de Laar MA. Patient preferences for a hospital-based rheumatology interactive health communication application and factors associated with these preferences. Rheumatology (Oxford) 2011 Sep;50(9):1618-1626 [FREE Full text] [doi: 10.1093/rheumatology/ker161] [Medline: 21551221]
104. Blackstock OJ, Cunningham CO, Haughton LJ, Garner RY, Norwood C, Horvath KJ. Higher eHealth literacy is associated with HIV risk behaviors among HIV-infected women who use the internet. J Assoc Nurses AIDS Care 2016;27(1):102-108 [FREE Full text] [doi: 10.1016/j.jana.2015.09.001] [Medline: 26456347]
105. Milne RA, Puts MT, Papadakos J, Le LW, Milne VC, Hope AJ, et al. Predictors of high eHealth literacy in primary lung cancer survivors. J Cancer Educ 2015 Dec;30(4):685-692. [doi: 10.1007/s13187-014-0744-5] [Medline: 25355524]
106. 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] [Medline: 22357448]
107. Gironda M, Der-Martirosian C, Messadi D, Holtzman J, Atchison K. A brief 20-item dental/medical health literacy screen (REALMD-20). J Public Health Dent 2013;73(1):50-55 [FREE Full text] [doi: 10.1111/jphd.12005] [Medline: 23293880]

108. Duncan M, Vandelanotte C, Kolt GS, Rosenkranz RR, Caperchione CM, George ES, et al. Effectiveness of a web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. J Med Internet Res 2014;16(6):e136 [FREE Full text] [doi: 10.2196/jmir.3107] [Medline: 24927299]
109. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. J Med Internet Res 2011 Nov;13(4):e94 [FREE Full text] [doi: 10.2196/jmir.1750] [Medline: 22094891]
110. Chang F, Chiu C, Chen P, Miao N, Lee C, Chiang J, et al. Relationship between parental and adolescent eHealth literacy and online health information seeking in Taiwan. Cyberpsychol Behav Soc Netw 2015 Oct;18(10):618-624. [doi: 10.1089/cyber.2015.0110] [Medline: 26375050]
111. Connolly KK, Crosby ME. Examining e-Health literacy and the digital divide in an underserved population in Hawai'i. Hawaii J Med Public Health 2014 Feb;73(2):44-48 [FREE Full text] [Medline: 24567867]
112. 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] [Medline: 26690761]
113. Hogan TP, Hill JN, Locatelli SM, Weaver FM, Thomas FP, Nazi KM, et al. Health information seeking and technology use among veterans with spinal cord injuries and disorders. PM R 2016 Feb;8(2):123-130. [doi: 10.1016/j.pmrj.2015.06.443] [Medline: 26164352]
114. Hsu W, Chiang C, Yang S. The effect of individual factors on health behaviors among college students: the mediating effects of eHealth literacy. J Med Internet Res 2014;16(12):e287 [FREE Full text] [doi: 10.2196/jmir.3542] [Medline: 25499086]
115. Knapp C, Madden V, Marcu M, Wang H, Curtis C, Sloyer P, et al. Information seeking behaviors of parents whose children have life-threatening illnesses. Pediatr Blood Cancer 2011 May;56(5):805-811. [doi: 10.1002/pbc.22674] [Medline: 21370415]
116. Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. J Med Internet Res 2011;13(3):e75 [FREE Full text] [doi: 10.2196/jmir.1697] [Medline: 21960017]
117. Koo M, Norman CD, Chang HM. Psychometric evaluation of a Chinese version of the eHealth literacy scale (eHEALS) in school age children. Int Electronic J Health Educ 2012;15:29-36 [FREE Full text]
118. Li N, Orrange S, Kravitz RL, Bell RA. Reasons for and predictors of patients' online health information seeking following a medical appointment. Fam Pract 2014 Oct;31(5):550-556. [doi: 10.1093/fampra/cmu034] [Medline: 24963151]
119. Mitsutake S, Shibata A, Ishii K, Oka K. Association of eHealth literacy with colorectal cancer knowledge and screening practice among internet users in Japan. J Med Internet Res 2012;14(6):e153 [FREE Full text] [doi: 10.2196/jmir.1927] [Medline: 23149453]
120. Nahm E, Resnick B, Orwig D, Magaziner J, Bellantoni M, Sterling R, et al. A theory-based online hip fracture resource center for caregivers: effects on dyads. Nurs Res 2012;61(6):413-422 [FREE Full text] [doi: 10.1097/NNR.0b013e318270712a] [Medline: 23052420]
121. Ossebaard HC, Seydel ER, van Gemert-Pijnen L. Online usability and patients with long-term conditions: a mixed-methods approach. Int J Med Inform 2012 Jun;81(6):374-387. [doi: 10.1016/j.ijmedinf.2011.12.010] [Medline: 22261086]
122. Park H, Cormier E, Gordon G, Baeg JH. Identifying health consumers' eHealth literacy to decrease disparities in accessing eHealth information. Comput Inform Nurs 2016 Feb;34(2):71-99. [doi: 10.1097/CIN.0000000000000205] [Medline: 26657619]
123. Park H, Moon M, Baeg JH. Association of eHealth literacy with cancer information seeking and prior experience with cancer screening. Comput Inform Nurs 2014 Sep;32(9):458-463. [doi: 10.1097/CIN.0000000000000077] [Medline: 25105588]
124. Reininger B, Mecca LP, Stine KM, Schultz K, Ling L, Halpern D. A type 2 diabetes prevention website for african americans, caucasians, and mexican americans: formative evaluation. JMIR Res Protoc 2013;2(2):e24 [FREE Full text] [doi: 10.2196/resprot.2573] [Medline: 23846668]
125. Sheng X, Simpson PM. Seniors, health information, and the internet: motivation, ability, and internet knowledge. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):740-746. [doi: 10.1089/cyber.2012.0642] [Medline: 23679569]
126. van der Vaart R, Drossaert CH, de Heus M, Taal E, van de Laar MA. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. J Med Internet Res 2013 Feb;15(2):e27 [FREE Full text] [doi: 10.2196/jmir.2428] [Medline: 23399720]
127. Li TM, Chau M, Wong PW, Lai ES, Yip PS. The association between online health information-seeking behaviors and health behaviors among hispanics in New York city: a community-based cross-sectional study. J Med Internet Res 2013;15(5):e80 [FREE Full text] [doi: 10.2196/jmir.2316] [Medline: 23676714]
128. Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CE, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. Clin Psychol Psychother 2013;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]

129. Taylor-Rodgers E, Batterham PJ. Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: randomised controlled trial. J Affect Disord 2014 Oct 15;168:65-71. [doi: 10.1016/j.jad.2014.06.047] [Medline: 25038293]
130. Tse CK, Bridges SM, Srinivasan DP, Cheng BS. Social media in adolescent health literacy education: a pilot study. JMIR Res Protoc 2015 Mar 09;4(1):e18 [FREE Full text] [doi: 10.2196/resprot.3285] [Medline: 25757670]
131. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: two experiments manipulating Google output. J Med Internet Res 2014;16(4):e100 [FREE Full text] [doi: 10.2196/jmir.2642] [Medline: 24694866]
132. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
133. Haun J, Noland D, Graham-Pole J, Rienzo B, Donaldson P. Testing a health literacy screening tool: implications for utilization of a BRIEF health literacy indicator. Fed Pract 2009;26(12):24-31.
134. Lee JY, Rozier RG, Lee SD, Bender D, Ruiz RE. Development of a word recognition instrument to test health literacy in dentistry: the REALD-30 - a brief communication. J Public Health Dent 2007;67(2):94-98. [Medline: 17557680]
135. Pleasant A, McKinney J, Rikard RV. Health literacy measurement: a proposed research agenda. J Health Commun 2011;16(Suppl 3):11-21. [doi: 10.1080/10810730.2011.604392] [Medline: 21951240]

Abbreviations

BRIEF: Brief Health Literacy Screening Tool
eSEARCH: eHealth Literacy Tool
eHealth: electronic health
eHEALS: eHealth Literacy Scale
EHIL: Everyday Health Information Literacy
FCCHL: Functional Communicative and Critical Health Literacy scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
REALM: Rapid Estimate in Adult Literacy in Medicine
TOHFLA: Test for Functional Health Literacy in Adults
WHO: World Health Organization

Edited by G Eysenbach; submitted 02.03.18; peer-reviewed by B Xie, K Sørensen, I Mircheva; comments to author 12.04.18; revised version received 07.09.18; accepted 25.09.18; published 19.12.18.

Please cite as:
Huhta AM, Hirvonen N, Huotari ML
Health Literacy in Web-Based Health Information Environments: Systematic Review of Concepts, Definitions, and Operationalization for Measurement
J Med Internet Res 2018;20(12):e10273
URL: http://www.jmir.org/2018/12/e10273/
doi:10.2196/10273
PMID:30567690

Original Paper

How Returning Aggregate Research Results Impacts Interest in Research Engagement and Planned Actions Relevant to Health Care Decision Making: Cohort Study

Elissa R Weitzman[1,2,3], ScD, MSc; Kara M Magane[1], MS; Lauren E Wisk[1,3], PhD

[1]Division of Adolescent/Young Adult Medicine, Boston Children's Hospital, Boston, MA, United States
[2]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, United States
[3]Department of Pediatrics, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Elissa R Weitzman, ScD, MSc
Division of Adolescent/Young Adult Medicine
Boston Children's Hospital
300 Longwood Avenue
BCH 3187
Boston, MA, 02115
United States
Phone: 1 617 355 3538
Fax: 1 617 730 0699
Email: elissa.weitzman@childrens.harvard.edu

Abstract

Background: Collection of patient-reported outcomes measures (PROs) may augment clinical data and inform health research, improving care, yet approaches to sustaining interest among patient cohorts in research participation are needed. One approach may involve returning aggregate research results (ARRs), which may help patients contextualize personal experiences, prompt conversations with providers or family, and encourage information seeking. This model has been demonstrated for Web-based patient-centered registries. Studies with clinical cohorts may further elucidate the model, its impacts on interest in research participation and planned actions, and potential for participants to experience this as helpful or harmful—gap areas.

Objective: We sought to investigate the impacts of returning ARRs comprising summaries of PROs and clinical metrics to parents of children with rheumatic disease, assessing interest in future research participation among parents who viewed ARRs and plans for acting on returned information. Further, we sought to investigate reactions to viewing ARRs and how these reactions impacted planned actions.

Methods: Clinical and PRO data were obtained about children in a national clinical disease registry, summarized, and processed into annotated infographics, comprising ARRs for children's parents. Parents who viewed ARRs (n=111) were surveyed about the information's perceived value and their reactions. Reaction patterns were summarized using principal components analysis (PCA), and associations among reaction patterns and interest in research participation and planned actions were estimated using multivariate logistic regression.

Results: Parental endorsement of the value of ARRs for understanding their child's condition and making care decisions was high (across 10 topics for which ARRs were shared, 42.2%-77.3% of the parents reported information was "very valuable"). Most (58/111, 52.3%) parents reported being more interested in participating in research after viewing ARRs, with the remainder reporting that their interest levels were unchanged. Reactions to viewing ARRs reflected experiencing validation/affirmation and information burden based on PCA. Reactions were not associated with child demographic or clinical characteristics and PROs, except that parents from households with less education reported greater information burden than those from more educated households (P=.007). In adjusted models, parents with higher validation/affirmation scores had increased odds of reporting heightened interest in research participation (adjusted odds ratio [AOR] 1.97, 95% CI 1.18-3.30), while higher information burden scores were associated with decreased odds of planned discussions with their child (AOR 0.59, 95% CI 0.36-0.95) and increased odds of planned discussions with providers (AOR 1.75, 95% CI 1.02-3.00).

Conclusions: Returning ARRs may foster a "virtuous cycle" of research engagement, especially where ARRs are experienced favorably and affect plans to share and discuss ARRs in support of a child's chronic disease care and treatment. Reactions to ARRs vary with education level, underscoring the need for attention to equity for this model.

(J Med Internet Res 2018;20(12):e10647) doi:10.2196/10647

KEYWORDS

aggregate research results; decision making; juvenile idiopathic arthritis; patient engagement; patient-reported outcome measures; rheumatic disease

Introduction

Growing evidence supports the importance of engaging patients in research to share information about their disease and treatment experiences and health-related quality of life [1,2]. Where shared data flow into health care systems and clinical epidemiologic studies, better symptom management, improved treatments, and better outcomes result [3]. In support of engaging patients in sharing high-quality data, considerable investment has been made in developing standardized patient-reported outcomes measures (PROs) that characterize aspects of patients' physical, mental, and social health [3-6]. PROs enable scientific rigor and help capture the patient voice, consistent with the paradigm shifting efforts to advance patient-centered outcomes research [7-10]. Yet, open questions remain about how to motivate ongoing patient engagement in research, and this question is central to ambitious efforts to activate vast cohorts in donating PROs and clinical data [11].

One possible approach is to operationalize a process whereby patients donate health data that are subsequently processed and returned in aggregate. Here, the return of aggregate research results (ARRs) about the cohort is hypothesized to motivate a virtuous cycle of data donation that can help grow the evidence base to advance more acceptable, effective therapies [12]. Viewing ARRs may be motivating for research participation if the returned information helps patients appraise personal experiences of disease and treatment [13] and informs conversations with health care providers and family members, factors that are relevant to health care decision making [14]. These actions, which reflect an engaged and activated patient and socially embedded nature of health care decision making, are central to models of chronic illness care [15-17]. These factors are also consistent with survey reports about what motivates sharing of personal health information [18-20] and are reflected in the appeal of Web-based patient-centered health information repositories [17]. Additional empirical testing of this "closed-loop" model is essential for ascertaining whether it fosters interest in research participation among clinical cohorts and to better understand its potential for being experienced as helpful or harmful—gap areas that are central to ensuring equipoise.

Returning ARRs may be informative and reassuring for some research participants, providing a normalizing context around experiences; ARRs may also be overwhelming and disquieting for others, including if ARRs show evidence of problems experienced by others with the same condition. Other factors, including the level of education and health literacy, might also affect acceptability. Concern about the balance of benefit or harm experienced when viewing ARRs may be especially acute for conditions that are rare, have treatments that rest on an immature evidence base or incompletely ameliorate symptoms or health-related quality of life, and pose risks for side effects [21]. Arguably, motivating ongoing research engagement for such conditions is especially important because gaps in knowledge might be filled, driving improvements in therapies at the system level and decision making at the patient or family level.

Pediatric-onset rheumatic disease (RD) is a trenchant case for examining these issues. Among youth, RD is rare but rising in incidence [22], with affected children facing significant hurdles regarding health-related quality of life due to the chronic relapsing nature of RD, unpredictable disease course, and difficult treatments [1-3]. Lack of a mature evidence base guiding RD care makes maintaining patient engagement in research especially important for gathering information about disease and treatment experiences to improve care and outcomes. For example, in a prior study focused on youth with RD, we found that treatment-related problems for standard RD therapies were common and contributed to poor health-related quality of life even after controlling for patient clinical characteristics and ameliorative effects of these treatments on symptoms, such as pain [23].

This study aims to investigate impacts of returning ARRs on interest in participating in future research within a larger project focused on investigating pediatric-onset RD. We engaged parents of children with RD in donating PROs about their child's health and treatment experiences and then returned cohort-level summaries of clinical measures and PROs to parents, testing whether viewing these ARRs increased parents' interest in future research participation and their intentions to discuss ARRs with others or seek further information. Such discussions might encourage shared decision making, consistent with theory [6,24]. We hypothesized that the receipt of ARRs would be highly motivating for future research engagement and that reactions to and planned actions arising from viewing ARRs would vary with participants' experiences of benefit and burden.

Methods

Study Design and Setting

We obtained data from The Learning Cohort (TLC) study [23], which surveyed consented parents of youth with a pediatric-onset RD enrolled in the Childhood Arthritis and Rheumatology Research Alliance (CARRA) Registry [25] during subspecialty care visits. Surveys included PROs to capture aspects of disease or treatment experience and

well-being [23]. PRO survey measures were programmed in the Research Electronic Data Capture system [25], from whence they could flow into the registry using a modular, ontology-based, federated informatics infrastructure constructed from open-source software; this infrastructure provides research investigators full ownership and access to their contributed data while supporting permissioned and robust data-sharing across federated sites [26]. In total, 4 CARRA Registry sites participated in this study, with Institutional Review Board approval from each. All participants provided informed consent. Details of the TLC study are published elsewhere [23].

Toward the close of the study period, the research team aggregated PRO data from the full TLC cohort (N=202 dyads) and clinical metrics from the CARRA Registry. These results were summarized by the research team into a curated, annotated set of ARRs, reflecting areas of leading concern to parents of registry-involved youth based on an initial survey of their information needs conducted when they enrolled [23]. ARRs comprised 68 unique slides (including titles and section headers) delivered on a tablet computer at a routine visit and were delivered as static infographics (visuals, figures, and charts). Topics covered study methods, clinical and treatment characteristics of children in the cohort, including medications used, patterns of health-related quality of life, experiences of pain and morning stiffness, and treatment problems (Figure 1). All materials were pretested, including the ARR slides, and the survey that was to be administered to parents to elicit reactions to returned data. The process was iterative to address all concerns. Pretests for accessibility were conducted with 5 parent volunteers and a representative of a family-based disease advocacy group. Pretests for accuracy and safety were conducted with 6 pediatric rheumatologists and 1 pediatric emergency room physician/ informatician.

From August 2015 to February 2016, *in lieu* of collecting additional PROs during clinic visits, an approximately 50% convenience sample of participating parents viewed ARRs (119 were approached, 115 consented; 96.6% consent rate), after which they completed a survey about their reactions to these materials; 111 parents provided complete data on their reactions to ARRs.

Measures

Demographic, Clinical and Health Characteristics

Parents reported their child's age; sex, race/ethnicity; diagnosis (juvenile idiopathic arthritis, systemic lupus erythematosus, or mixed connective tissue disease); overall health status; health-related quality of life [27]; pain interference [28]; morning stiffness; experience of serious side effects from a medication; methotrexate intolerance [29]; and highest education attained in the family. Disease duration was obtained from the CARRA Registry. Time in cohort was calculated as the number of days from the enrollment date (initial PRO collection date) to the final PRO collection date. Sample mode and mean imputation were used for 9 participants with incomplete data on demographic or clinical characteristics.

Perceived Value of and Reactions to the Return of Aggregate Research Results

Novel measures were developed and used to assess parents' reactions to the return of ARRs. Perceived value of the return of ARRs was ascertained with the question, "Overall, how valuable might this summary information be when understanding and making decisions about your child's condition and care?" asked for each of 10 topics that were shared in ARRs. Responses were given using a 3-point Likert scale (very valuable, somewhat valuable, and not valuable) and reported as frequencies. Sample size for each item ranged from 106 to 111 due to participant nonresponse on select items. Parents' reactions to seeing ARRs were determined by the extent to which they agreed or disagreed with the following 6 statements, each rated on a 5-point Likert scale (strongly agree, agree, undecided, disagree, strongly disagree): "Seeing summary information about the experiences of other study participants was comforting because it made me feel like my experiences are shared and validated as real;" "Overall, seeing summary information about other patients' experiences help me understand my child's experience;" "Reviewing this type of information is within my comfort zone;" "Overall, this type of information raises more questions than it answers;" "This type of information requires more knowledge or expertise than I have to understand it;" and "I prefer to let my rheumatologist digest this type of information." The mean of nonmissing items within the scale was used to impute missing response for 5 participants missing 1 or 2 (of 6) items.

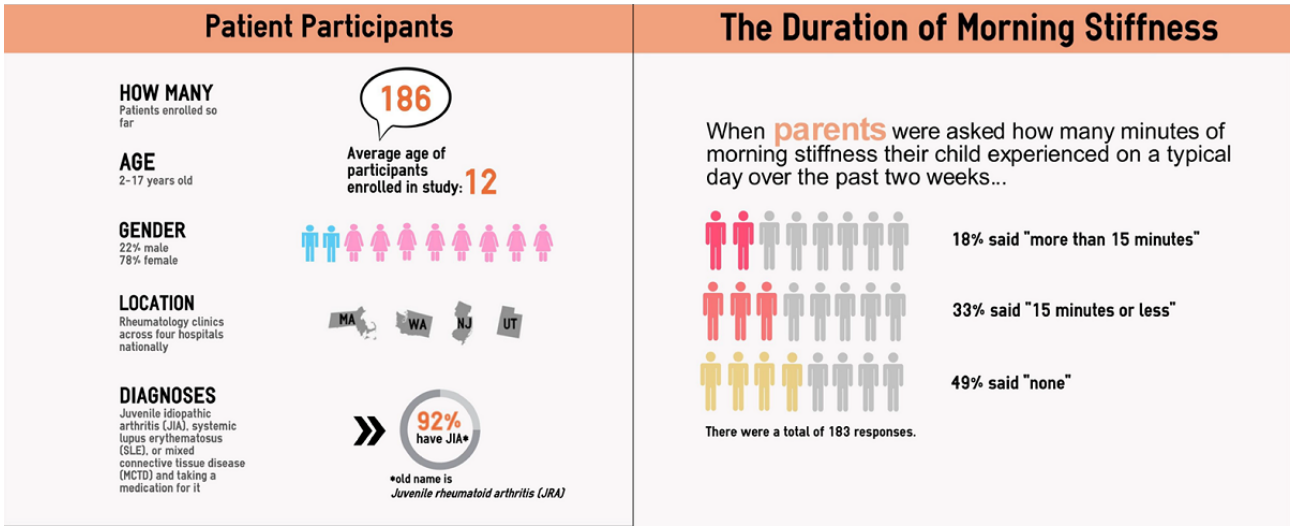
Engagement Outcomes

Parents were asked to report their interest in participating in research studies after seeing ARRs; options included more interested, less interested, and not any more or less interested. As one respondent endorsed "less interested," this response was combined with "not any more or less interested" to create a dichotomous variable. Parents were asked to report their planned actions after reviewing the summary information provided in ARRs by selecting all that applied from the following list: (1) discuss contents of the slide-deck with my child's health care provider; (2) discuss contents of the slide-deck with my child; (3) explore different medications; (4) look up information about something I saw in this slide-deck; and (5) do something else.

Statistical Analyses

All analyses were conducted using SAS 9.4 software (SAS Institute, Inc, Cary, North Carolina). Summary statistics were generated to describe sample characteristics; differences in demographic and clinical characteristics between parents who did (n=111) or did not (n=91) receive ARRs were evaluated using *t* tests, Kruskal-Wallis, or chi-square (χ^2) tests, as appropriate. Principal components analysis (PCA) was conducted to investigate the commonalities between ARR reactions and generate summary variables based on individual reaction measures. All factors with eigenvalues >1 were retained consistent with the standard practice [30,31], leading to a 2-factor solution, and an orthogonal rotation was applied to generate 2 uncorrelated scales (hereafter referred to as "validation/affirmation" and "information burden").

Figure 1. Example content from aggregate research results returned to parents.



Validation/affirmation reflects reactions to ARRs that reflect experiences of recognizing the legitimacy of personal experiences, greater insight into their child's condition, and level of comfort with ARRs. Information burden reflects reactions to ARRs that reflect experiences akin to being over one's head or uncomfortable with information in ARRs owing to perceived lack of personal expertise, preferences for their provider to digest the information and uncertainty. Pearson correlations, *t* tests, and analysis of variance (as appropriate) were used to examine bivariate relationships between each of the 2 ARR reaction constructs and demographic and clinical characteristics. Subsequently, multivariate logistic regression was used to model engagement outcomes as predicted by the 2 ARR reaction constructs; both unadjusted (controlling only for both factors simultaneously) as well as models adjusting for child age, race/ethnicity, and parent education were used.

Results

Sample Characteristics

Children in this cohort were 12.0 years of age on average (SD 3.6) and predominantly female (n=161, 79.7%), diagnosed with juvenile idiopathic arthritis (n=187, 92.6%), white/non-Hispanic individuals (n=152, 75.2%), and had parents with any college education (n=144, 71.3%); the average disease duration was 7.7 (SD 3.5) years. No differences were observed in demographic or clinical characteristics between those who did or did not receive ARRs (Table 1).

Perceived Value of and Reactions to the Return of Aggregate Research Results

The proportion of parents finding the ARR topics to be "very valuable" ranged from 42.2% (for experiences of morning stiffness) to 77.3% (for medication problems; Figure 2).

PCA identified a 2-factor solution with high loadings for all 6 ARR reaction items (Figure 3).

The validation/affirmation construct was defined by 3 items reflecting having one's experience validated (factor loading=0.873), improved understanding of their child's condition (factor loading=0.865), and feeling that the ARR materials are within their comfort zone (factor loading=0.564). The information burden construct was defined by reports of requiring more knowledge to understand ARRs (factor loading=0.856), preferring a physician to "digest" ARRs (factor loading=0.717), and viewing ARRs raising more questions than were answered (factor loading=0.682). Parents from households with less education reported greater information burden than those from more educated households (*P*=.007; Table 2); factors were not significantly associated with other demographic or clinical characteristics.

Engagement Outcomes

The majority of the parents (58/111, 52.3%) reported that after seeing ARR, they were more interested in participating in research. Higher validation/affirmation scores were associated with nearly a 2-fold increase in the odds of reporting more interest in research participation (adjusted odds ratio [AOR]: 1.97, 95% CI 1.18-3.30; Table 3). Nearly one-third (35/111, 31.5%) of the parents reported that they would discuss the contents of ARRs with their child, nearly one-fifth (20/111, 18.0%) reported they would look up information, 15.3% (17/111) said that they would discuss ARR contents with their child's health care provider, and approximately one-tenth (11/111, 9.9%) reported they would explore different medications. Parents who reported greater information burden were less likely to report wanting to discuss the contents of ARRs with their child (AOR 0.59, 95% CI 0.36-0.95) but were more likely to report wanting to discuss the contents of ARRs with their child's health care provider (AOR 1.75, 95%CI 1.02-3.00).

Table 1. The parent-reported sample characteristics by receipt of aggregate research results (ARRs).

Characteristic	Total (N=202)	Received ARRs (n=111)	Did not receive ARRs (n=91)	*P* value
Demographics				
Child age (years), mean (SD)	12.0 (3.6)	11.7 (3.6)	12.3 (3.7)	.29
Child sex, n (%)				.87
Female	161 (79.7)	88 (79.3)	73 (80.2)	—[a]
Male	41 (20.3)	23 (20.7)	18 (19.8)	—
Child race or ethnicity, n (%)				.63
White, non-Hispanic	152 (75.2)	85 (76.6)	67 (73.6)	—
Non-white or Hispanic	50 (24.8)	26 (23.4)	24 (26.4)	—
Highest level of parent education, n (%)				.79
High school graduate or less	58 (28.7)	31 (27.9)	27 (29.7)	—
Any college	144 (71.3)	80 (72.1)	64 (70.3)	—
Clinical characteristics				
Diagnosis, n (%)				.50
Juvenile idiopathic arthritis	187 (92.6)	104 (93.7)	83 (91.2)	—
Systematic lupus erythematosus or mixed connective tissues disease	15 (7.4)	7 (6.3)	8 (8.8)	—
Disease duration (in years), mean (SD)	7.7 (3.5)	7.7 (3.5)	7.7 (3.6)	.99
Methotrexate use and intolerance, n (%)				.65
No methotrexate use	103 (51.0)	59 (53.2)	44 (48.4)	—
Use with no intolerance	60 (29.7)	30 (27.0)	30 (33.0)	—
Methotrexate intolerance	39 (19.3)	22 (19.8)	17 (18.7)	—
Overall health[b], mean (SD)	8.1 (2.0)	8.2 (1.8)	8.0 (2.3)	.45
Typical morning stiffness in the past 2 weeks (minutes)[c]				.30
>15	36 (17.8)	17 (15.3)	19 (20.9)	—
≤15	166 (82.2)	94 (84.7)	72 (79.1)	—
Lifetime serious medication side effects				.64
One or more	52 (25.7)	30 (27.0)	22 (24.2)	—
None	150 (74.3)	81 (73.0)	69 (75.8)	—
Pediatric Quality of Life Inventory 4.0 scores, mean (SD)				
Total score	75.7 (18.3)	74.9 (18.4)	76.7 (18.4)	.50
Psychosocial score	76.1 (18.0)	75.2 (18.2)	77.1 (17.8)	.46
Physical score	75.1 (22.6)	74.4 (22.3)	75.9 (23.1)	.65
Patient-Reported Outcomes Measurement Information System Pain Interference[e], mean (SD)	50.5 (10.9)	51.2 (11.0)	49.7 (10.9)	.35
Time in cohort (days), mean (SD)	95.8 (100.9)	105.1 (100.7)	84.3 (100.4)	.14

[a]Not applicable.

[b]Parents' rating of their child's overall health from 1 to 10, where higher scores indicate better health.

[c]Parents' report of the number of minutes of morning stiffness their child experiences on a typical day over the past 2 weeks.

[d]The possible range of scores is from 0 to 100, with higher score indicating better quality of life.

[e]Raw pain interference scores were transformed into a "*t* score" for each participant. The *t* score rescales the raw score into a standardized score with a mean of 50, SD of 10, and the possible range of 38-78, with higher score indicating more pain interference.

Figure 2. Parents' ratings of the value of each patient-reported topic presented in aggregate research results for understanding and making decisions regarding their child's condition and care.

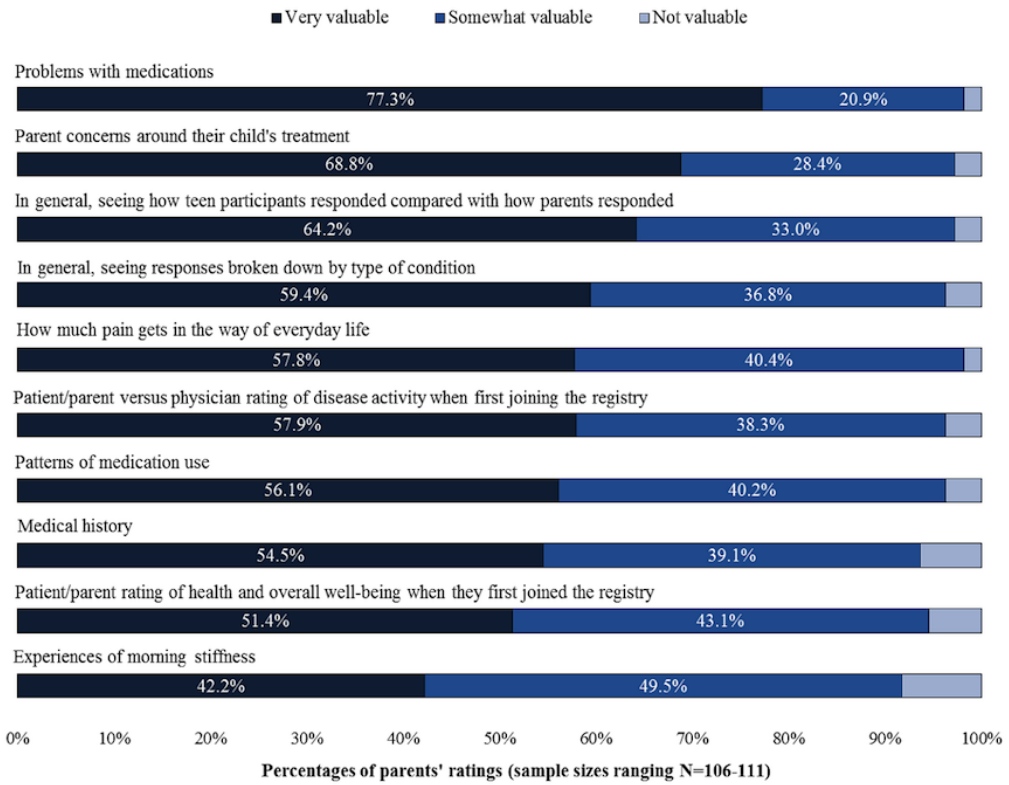


Figure 3. Percentages and factor loadings of response items regarding reactions to the return of aggregate research results.

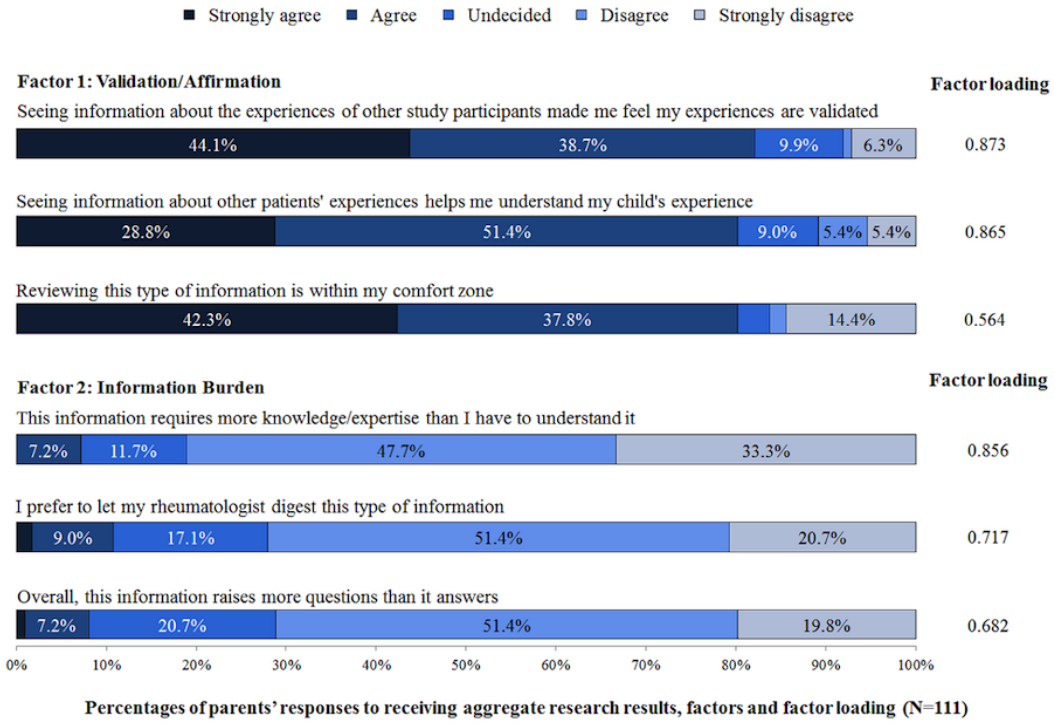


Table 2. Reactions to aggregate research results (ARRs) by demographic and clinical characteristics (n=111).

Characteristic	Factor 1: Validation/affirmation[a]	*P* value	Factor 2: Information burden[b]	*P* value
Demographic characteristics				
Child age (years), correlation coefficient	0.12	.21	0.00	.98
Child sex, mean (SD)				
Female	–0.05 (1.02)	.29	0.07 (1.02)	.15
Male	0.20 (0.92)	—[c]	–0.27 (0.89)	—
Child race or ethnicity, mean (SD)				
White, non-Hispanic	–0.01 (1.05)	.92	–0.06 (0.96)	.24
Non-white or Hispanic	0.02 (0.81)	—	0.20 (1.11)	—
Highest level of parent education, mean (SD)				
High school graduate or less	0.07 (0.89)	.67	0.41 (1.10)	.007
Any college	–0.03 (1.04)	—	–0.16 (0.92)	—
Clinical characteristics				
Diagnosis, mean (SD)				
Juvenile idiopathic arthritis	–0.02 (1.02)	.52	0.01 (0.99)	.69
Systematic lupus erythematosus or mixed connective tissues disease	0.24 (0.70)	—	–0.15 (1.15)	—
Disease duration (years), correlation coefficient	0.11	.27	0.05	.57
Methotrexate use and intolerance, mean (SD)				
No methotrexate use	–0.01 (0.96)	.27	0.07 (0.97)	.54
Use with no intolerance	0.20 (0.82)	—	–0.17 (1.01)	—
Methotrexate intolerance	–0.25 (1.28)	—	0.04 (1.08)	—
Overall health[d], correlation coefficient	−0.07	.48	−0.03	.73
Typical morning stiffness in the past 2 weeks (minutes)[e], mean (SD)				
>15	0.11 (1.05)	.61	0.05 (1.02)	.83
≤15	–0.02 (1.00)	—	–0.01 (1.00)	—
Lifetime serious medication side effects, mean (SD)				
One or more	0.27 (0.95)	.09	–0.25 (0.78)	.10
None	–0.10 (1.01)	—	0.09 (1.06)	—
Pediatric Quality of Life Inventory 4.0 scores[f], correlation coefficient				
Total	–0.09	.34	–0.11	.24
Psychosocial Score	–0.11	.25	–0.10	.29
Physical Score	–0.05	.61	–0.11	.26
Patient-Reported Outcomes Measurement Information System Pain Interference[g], correlation coefficient	0.06	.54	–0.02	.82
Time in Cohort (days), correlation coefficient	–0.02	.80	0.08	.43

[a]Reflects the extent to which parents feel their experience is validated or affirmed when viewing ARRs; higher scores indicate greater agreement.

[b]Reflects the extent to which parents experienced information burden when viewing; higher scores indicate greater agreement.

[c]Not applicable.

[d]Parent's rating of their child's overall health from 1 to 10, where higher scores indicate better health.

[e]Parent's report of the number of minutes of morning stiffness their child experiences on a typical day over the past 2 weeks.

[f]The possible range of scores is from 0 to 100, with higher score indicating better quality of life.

[g]Raw pain interference scores were transformed into a "*t* score" for each participant. The *t* score rescales the raw score into a standardized score with a mean of 50, SD of 10, and the possible range of 38-78. A higher score indicates more pain interference.

Table 3. Associations between reactions to aggregate research results and engagement outcomes.

Outcome[a]	Outcome prevalence, n (%)	Unadjusted models, OR (95% CI)	Adjusted models, OR (95% CI)
More interest in participating in research			
Validation/affirmation	58 (52.3)	1.97 (1.21-3.18)	1.97 (1.18-3.30)
Information burden		1.33 (0.89-1.98)	1.36 (0.89-2.09)
Planned actions			
Discuss with child			
Validation/affirmation	35 (31.5)	1.22 (0.79-1.87)	1.18 (0.75-1.86)
Information burden		0.69 (0.45-1.07)	0.59 (0.36-0.95)
Look up information			
Validation/affirmation	20 (18.0)	0.93 (0.58-1.47)	0.87 (0.54-1.41)
Information burden		0.68 (0.39-1.17)	0.64 (0.36-1.15)
Discuss with providers			
Validation/affirmation	17 (15.3)	0.69 (0.43-1.11)	0.69 (0.42-1.13)
Information burden		1.63 (0.98-2.71)	1.75 (1.02-3.00)
Explore different medications			
Validation/affirmation	11 (9.9)	1.03 (0.54-1.98)	1.06 (0.53-2.10)
Information burden		1.15 (0.63-2.11)	1.31 (0.69-2.52)

[a]Frequency and unadjusted prevalence of engagement outcomes among those who received aggregate research results (n=111). Planned action prevalence did not sum to 100% as participants could endorse multiple actions. "Unadjusted" models controlled for validation/affirmation and information burden scales only. Adjusted models controlled for the child's age, race/ethnicity, and highest education attained in the family in addition to both validation/affirmation and information burden scales.

Discussion

This study provides evidence that returning ARRs to study participants who have donated data is highly motivating for ongoing research participation. By investigating effects of sharing ARRs on interest in research participation and planned actions stemming from this experience, findings extend the understanding of benefits of patient engagement in health care research [32]. More than half of the parents in the cohort exposed to ARRs reported increased interest in research participation as a function of receiving ARRs—good news for participatory research models, including those predicated on engaging cohorts in donating data. Moreover, parents endorsed plans for proactive engagement in their child's treatment and outcomes after receiving ARRs, with specific plans varying as a function of validation/affirmation and information burden. Overall, reactions to ARRs aligned with hopes for fostering motivation for research participation and activation in the health care process [33,34].

As hypothesized, the experience of viewing ARRs was not "one size fits all"—participants' responses to the model reflected both experiences of validation/affirmation, wherein parents gained value from contextualizing disease or treatment experiences through viewing cohort-level data, and information burden, wherein aggregate results may be cognitively overwhelming. Yet, as greater information burden was associated with plans to discuss results with a provider, even parent participants who may have felt overwhelmed by ARRs might benefit if they are stimulated to talk with health care providers to understand findings and discuss any implications for their child. Such activities could foster improved patient-provider partnership and shared decision making. Notably, experiencing greater information burden was evident among participants who reported lower levels of educational attainment. As such, achieving goals of optimizing and equalizing health benefits and reducing the potential for disparities stemming from this model may require additional support around interpreting and processing ARRs [34]. Nationally, low levels of health literacy and numeracy indicate that large segments of the US population may face barriers to understanding health data, underscoring the importance of attending to these issues [35,36].

This investigation was undertaken with a cohort situated in a well-defined national multisite disease registry, whose members' diagnoses were clinically confirmed—significant strengths. Still, several limitations merit mention. First, findings reflect the experiences of engaged parents of registry participants who viewed a specific set of ARRs and may not generalize to other populations that may differ in conditions, concerns, and experiences (including the history of benefit or harm from research and care); the ARR contents are also specific to the population. Second, this study leveraged the visit structure of a registry-engaged clinical cohort, layering data collection and return onto the natural visit cycle of this cohort. Additional research using a randomized trial design would inform comparisons of engagement outcomes for participants who did and did not receive ARRs over time. Third, participants shared structured PROs, some validated, others novel; items describing motivation to participate in research and experiences of

receiving ARRs are not validated but capture important patient-centered dimensions of research experience. Fourth, parent proxy reports of child clinical characteristics and PROs and parent reactions to the return of ARRs may differ from those of the child [9,37], and results from this study should not be construed as reflecting child (ie, patient) reactions to the return of ARRs or reactions of another parent or guardian. Relatedly, excepting the measure of the highest parent education attained, demographic characteristics describe the child not parents. Congruence between child and parent experiences of this model and further investigation of effects of parent demographics on outcomes may further inform this work and merit future study. Lastly, self-reported data are subject to recall and social desirability biases; however, the use of structured and validated measures and electronic data collection help protect against known validity threats.

In sum, viewing ARRs increased motivation for research participation among a majority of study participants and shows promise for driving greater patient activation and engagement. Results of this study are encouraging in light of national plans for engaging volunteers in donating personally generated data, including PROs, to drive precision medicine and comparative effectiveness research [7,11,38,39]. To the extent that these efforts utilize a closed loop approach in which ARRs are returned to a learning and sharing cohort, they may thrive. Protecting against the potential for unintentionally worsening health disparities is vital as results show that participants from households with less parent education were more likely to experience information burden from viewing ARRs. Should this lead to the differential engagement or attrition of less educated participants, biases in study findings could be introduced and the ultimate fairness and beneficence of the model undermined.

Acknowledgments

This study was funded by grants from the National Library of Medicine award #R01LM011185; the National Institute of Arthritis and Musculoskeletal and Skin Diseases award #U19AR069522 and #R21AR070944; and the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resources and Services Administration award #T71MC00009. The funding bodies were not involved in the design of the study or collection, analysis, and interpretation of data, nor in writing, review or approval of the manuscript.

Conflicts of Interest

None declared.

References

1. Manary M, Boulding W, Staelin R, Glickman S. The patient experience and health outcomes. N Engl J Med 2013 Jan 17;368(3):201-203. [doi: 10.1056/NEJMp1211775] [Medline: 23268647]
2. Washington AE, Lipstein SH. The Patient-Centered Outcomes Research Institute--promoting better information, decisions, and health. N Engl J Med 2011 Oct 13;365(15):e31. [doi: 10.1056/NEJMp1109407] [Medline: 21992473]
3. Basch E. Patient-Reported Outcomes - Harnessing Patients' Voices to Improve Clinical Care. N Engl J Med 2017 Jan 12;376(2):105-108. [doi: 10.1056/NEJMp1611252] [Medline: 28076708]
4. Concannon T, Fuster M, Saunders T, Patel K, Wong J, Leslie L, et al. A systematic review of stakeholder engagement in comparative effectiveness and patient-centered outcomes research. J Gen Intern Med 2014 Dec;29(12):1692-1701 [FREE Full text] [doi: 10.1007/s11606-014-2878-x] [Medline: 24893581]
5. Asch D, Rosin R. Engineering Social Incentives for Health. N Engl J Med 2016 Dec 29;375(26):2511-2513. [doi: 10.1056/NEJMp1603978] [Medline: 28029924]
6. Lavallee DC, Chenok KE, Love RM, Petersen C, Holve E, Segal CD, et al. Incorporating Patient-Reported Outcomes Into Health Care To Engage Patients And Enhance Care. Health Aff (Millwood) 2016 Apr;35(4):575-582. [doi: 10.1377/hlthaff.2015.1362] [Medline: 27044954]
7. Ahmed S, Berzon RA, Revicki DA, Lenderking WR, Moinpour CM, Basch E, International Society for Quality of Life Research. The use of patient-reported outcomes (PRO) within comparative effectiveness research: implications for clinical practice and health care policy. Med Care 2012 Dec;50(12):1060-1070. [doi: 10.1097/MLR.0b013e318268aaff] [Medline: 22922434]
8. Levine DR, Mandrell BN, Sykes A, Pritchard M, Gibson D, Symons HJ, et al. Patients' and Parents' Needs, Attitudes, and Perceptions About Early Palliative Care Integration in Pediatric Oncology. JAMA Oncol 2017 Sep 01;3(9):1214-1220 [FREE Full text] [doi: 10.1001/jamaoncol.2017.0368] [Medline: 28278329]
9. Brudvik C, Moutte S, Baste V, Morken T. A comparison of pain assessment by physicians, parents and children in an outpatient setting. Emerg Med J 2017 Mar;34(3):138-144 [FREE Full text] [doi: 10.1136/emermed-2016-205825] [Medline: 27797872]
10. Selby JV, Forsythe L, Sox HC. Stakeholder-Driven Comparative Effectiveness Research: An Update From PCORI. JAMA 2015 Dec 01;314(21):2235-2236. [doi: 10.1001/jama.2015.15139] [Medline: 26595059]
11. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795 [FREE Full text] [doi: 10.1056/NEJMp1500523] [Medline: 25635347]

12. Califf R, Robb M, Bindman A, Briggs J, Collins F, Conway P, et al. Transforming Evidence Generation to Support Health and Health Care Decisions. N Engl J Med 2016 Dec 15;375(24):2395-2400. [doi: 10.1056/NEJMsb1610128] [Medline: 27974039]
13. Oerlemans S, Arts L, Horevoorts N, van de Poll-Franse LV. “Am I normal?” The Wishes of Patients With Lymphoma to Compare Their Patient-Reported Outcomes With Those of Their Peers. J Med Internet Res 2017 Aug 15;19(8):e288 [FREE Full text] [doi: 10.2196/jmir.7079] [Medline: 28811271]
14. Dwamena F, Holmes-Rovner M, Gaulden C, Jorgenson S, Sadigh G, Sikorskii A, et al. Interventions for providers to promote a patient-centred approach in clinical consultations. Cochrane Database Syst Rev 2012 Dec 12;12:CD003267. [doi: 10.1002/14651858.CD003267.pub2] [Medline: 23235595]
15. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. Health Aff (Millwood) 2001;20(6):64-78. [doi: 10.1377/hlthaff.20.6.64] [Medline: 11816692]
16. Gee P, Greenwood D, Paterniti D, Ward D, Miller L. The eHealth Enhanced Chronic Care Model: a theory derivation approach. J Med Internet Res 2015 Apr 01;17(4):e86 [FREE Full text] [doi: 10.2196/jmir.4067] [Medline: 25842005]
17. Bradley M, Braverman J, Harrington M, Wicks P. Patients' motivations and interest in research: characteristics of volunteers for patient-led projects on PatientsLikeMe. Res Involv Engagem 2016;2:33 [FREE Full text] [doi: 10.1186/s40900-016-0047-6] [Medline: 29507767]
18. Sanderson S, Linderman M, Suckiel S, Diaz G, Zinberg R, Ferryman K, et al. Motivations, concerns and preferences of personal genome sequencing research participants: Baseline findings from the HealthSeq project. Eur J Hum Genet 2016 Jan;24(1):14-20 [FREE Full text] [doi: 10.1038/ejhg.2015.118] [Medline: 26036856]
19. Weitzman ER, Kelemen S, Kaci L, Mandl KD. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. BMC Med Inform Decis Mak 2012 May 22;12:39 [FREE Full text] [doi: 10.1186/1472-6947-12-39] [Medline: 22616619]
20. Weitzman E, Kaci L, Mandl K. Sharing medical data for health research: the early personal health record experience. In: J Med Internet Res. Sharing medical data for health research: The early personal health record experience. J Med Internet Res Internet Journal of Medical Internet Research; May 25, 2010:e14.
21. Politi MC, Lewis CL, Frosch DL. Supporting shared decisions when clinical evidence is low. Med Care Res Rev 2013 Feb;70(1 Suppl):113S-128S. [doi: 10.1177/1077558712458456] [Medline: 23124616]
22. Helmick C, Felson D, Lawrence R, Gabriel S, Hirsch R, Kwoh C, National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis Rheum 2008 Jan;58(1):15-25 [FREE Full text] [doi: 10.1002/art.23177] [Medline: 18163481]
23. Weitzman E, Wisk L, Salimian P, Magane K, Dedeoglu F, Hersh A, et al. Adding patient-reported outcomes to a multisite registry to quantify quality of life and experiences of disease and treatment for youth with juvenile idiopathic arthritis. J Patient Rep Outcomes 2018;2:1-11 [FREE Full text] [doi: 10.1186/s41687-017-0025-2] [Medline: 29645010]
24. Wagner E. Chronic disease management: what will it take to improve care for chronic illness? Eff Clin Pract 1998;1(1):2-4. [Medline: 10345255]
25. Ringold S, Beukelman T, Nigrovic P, Kimura Y, CARRA Registry Site Principal Investigators. Race, ethnicity, and disease outcomes in juvenile idiopathic arthritis: a cross-sectional analysis of the Childhood Arthritis and Rheumatology Research Alliance (CARRA) Registry. J Rheumatol 2013 Jun;40(6):936-942. [doi: 10.3899/jrheum.121147] [Medline: 23588937]
26. Natter M, Quan J, Ortiz D, Bousvaros A, Ilowite N, Inman C, et al. An i2b2-based, generalizable, open source, self-scaling chronic disease registry. J Am Med Inform Assoc 2013 Jan 01;20(1):172-179 [FREE Full text] [doi: 10.1136/amiajnl-2012-001042] [Medline: 22733975]
27. Varni J, Seid M, Rode C. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care 1999 Feb;37(2):126-139. [Medline: 10024117]
28. Amtmann D, Cook K, Jensen M, Chen W, Choi S, Revicki D, et al. Development of a PROMIS item bank to measure pain interference. Pain 2010 Jul;150(1):173-182 [FREE Full text] [doi: 10.1016/j.pain.2010.04.025] [Medline: 20554116]
29. Bulatović M, Heijstek M, Verkaaik M, van Dijkhuizen EHP, Armbrust W, Hoppenreijs E, et al. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: development and validation of a methotrexate intolerance severity score. Arthritis Rheum 2011 Jul;63(7):2007-2013 [FREE Full text] [doi: 10.1002/art.30367] [Medline: 21437879]
30. Jolliffe IT. Principal Component Analysis, Second Edition. New York, USA: Springer-Verlag New York; 2002.
31. Dunteman G. Principal components analysis. Beverly Hills: Sage Publications; 1989:65-69.
32. Domecq J, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014 Feb 26;14:89 [FREE Full text] [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
33. Hibbard J, Greene J, Overton V. Patients with lower activation associated with higher costs; delivery systems should know their patients' 'scores'. Health Aff (Millwood) 2013 Feb;32(2):216-222. [doi: 10.1377/hlthaff.2012.1064] [Medline: 23381513]
34. Hibbard J. Patient Activation and Health Literacy: What's the Difference? How Do Each Contribute to Health Outcomes. Stud Health Technol Inform 2017;240:251-262. [Medline: 28972522]
35. Keim-Malpass J, Letzkus L, Kennedy C. Parent/caregiver health literacy among children with special health care needs: a systematic review of the literature. BMC Pediatr 2015 Aug 05;15:92 [FREE Full text] [doi: 10.1186/s12887-015-0412-x] [Medline: 26242306]

36. Friis K, Lasgaard M, Osborne R, Maindal H. Gaps in understanding health and engagement with healthcare providers across common long-term conditions: a population survey of health literacy in 29,473 Danish citizens. BMJ Open 2016 Jan 14;6(1):e009627 [FREE Full text] [doi: 10.1136/bmjopen-2015-009627] [Medline: 26769783]
37. Rescorla L, Ginzburg S, Achenbach T, Ivanova M, Almqvist F, Begovac I, et al. Cross-informant agreement between parent-reported and adolescent self-reported problems in 25 societies. J Clin Child Adolesc Psychol 2013;42(2):262-273. [doi: 10.1080/15374416.2012.717870] [Medline: 23009025]
38. Devine E, Alfonso-Cristancho R, Devlin A, Edwards T, Farrokhi E, Kessler L, CERTAIN Collaborative. A model for incorporating patient and stakeholder voices in a learning health care network: Washington State's Comparative Effectiveness Research Translation Network. J Clin Epidemiol 2013 Aug;66(8 Suppl):S122-S129 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.04.007] [Medline: 23849146]
39. Beskow LM, Burke W, Fullerton SM, Sharp RR. Offering aggregate results to participants in genomic research: opportunities and challenges. Genet Med 2012 Apr;14(4):490-496 [FREE Full text] [doi: 10.1038/gim.2011.62] [Medline: 22261761]

Abbreviations

AOR: adjusted odds ratio
ARR: aggregate research result
CARRA: Childhood Arthritis and Rheumatology Research Alliance
PCA: principal components analysis
PRO: patient-reported outcome measure
RD: rheumatic disease
TLC: The Learning Cohort

Edited by G Eysenbach; submitted 02.04.18; peer-reviewed by S Rush, A Hoffman; comments to author 24.07.18; revised version received 15.08.18; accepted 10.09.18; published 21.12.18.

Please cite as:
Weitzman ER, Magane KM, Wisk LE
How Returning Aggregate Research Results Impacts Interest in Research Engagement and Planned Actions Relevant to Health Care Decision Making: Cohort Study
J Med Internet Res 2018;20(12):e10647
URL: http://www.jmir.org/2018/12/e10647/
doi:10.2196/10647
PMID:30578228

Original Paper

Pediatric Web-Based Chat Services for Caregivers of Children: Descriptive Study

Anu Kaskinen[1], MD, PhD; Benjamin Ayeboa-Sallah[2], MD, MSc; Tiina Teivaanmäki[1], MD, PhD; Elina Wärnhjelm[1], MD; Liisa Korhonen[1], MD; Otto Helve[1,2], MD, PhD

[1]Pediatric Research Center, New Children's Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[2]National Institute for Health and Welfare, Helsinki, Finland

Corresponding Author:
Anu Kaskinen, MD, PhD
Pediatric Research Center
New Children's Hospital
University of Helsinki and Helsinki University Hospital
Stenbäckinkatu 9
PB 347
Helsinki, 00029 HUS
Finland
Phone: 358 94711
Fax: 358 947171977
Email: anu.kaskinen@helsinki.fi

Abstract

Background: Pediatric physician-led Web-based chat services offer a novel, low-threshold communication channel between caregivers and physicians.

Objective: Our aim was to describe chat conversations between caregivers and physicians in a Web-based chat service to determine the factors that should be considered when planning a similar chat service. We also aimed to evaluate whether caregivers considered the consultations helpful, whether physicians considered they could answer caregivers' questions, and whether further face-to-face medical contact was needed.

Methods: In September 2015, a private medical center for children in the greater Helsinki area initiated a Web-based chat service, accessible via any device with an internet connection, open from 9 am to 9 pm local time. Four residents in pediatrics, who had performed at least 60% of their 6-year residency program, served as the physicians responsible for chat consultations with caregivers of children. Between October 2015 and March 2016, 343 consecutive consultations were immediately evaluated by a chat physician. On average, caregivers were followed up by email questionnaire 7-14 days later, which 98 caregivers answered a median of 11 (interquartile range, IQR, 7-20) days later.

Results: The age of the children whose caregivers contacted the chat service was a median of 2.1 (IQR 0.83-4.69) years, and 29.8% (102/342) of the children were less than 1 year old. The majority (119/343, 34.7%) of the chat conversations took place from 9 am to noon, and infections were the most common concern in over half of cases (189/343, 55.1%). Chat physicians recommended a face-to-face appointment with a physician for that same day in 13.7% (47/343) of the cases. A face-to-face exam was recommended for that same day more often if the chat concerned infection (36/189, 19.0% cases) compared with other reasons (11/154, 7.1%, cases; P=.001). Physicians felt capable of answering caregivers' questions in 72.6% (249/343) of the cases, whereas 93% (91/98) of caregivers considered physicians' answers helpful. Whether caregivers had to take their children to see a physician that same day or whether caregivers' main concern was infection was not found to be associated with whether caregivers considered physicians' answers helpful or not. However, physicians felt more capable of answering caregivers' questions when the main concern was infection.

Conclusions: Parental consultations via Web-based chat service often take place before noon and focus on infection-related issues as well as on the health and illness of very young children. These factors should be considered when planning or setting up such a service. Based on the high satisfaction with the chat service by both physicians and caregivers, Web-based chat services may be a useful way to help caregivers with concerns about their child's health or illness.

(J Med Internet Res 2018;20(12):e10165) doi:10.2196/10165

KEYWORDS

chat service; health information; internet; Web-based resources; pediatrics; social media

Introduction

The availability of Web-based health information and usage of the internet to research health-related problems have increased during the last decade and still continue to grow [1-3]. In an Austrian study from 2015, more than 90% of caregivers visiting an outpatient clinic with their children reported that they collect health information from the internet [4]. Furthermore, seeking Web-based health information can reduce unnecessary contact with health care professionals [5].

The quality of social media sites as a source of health information is often variable as opinions are often biased and presented as facts, and it may be difficult to find reliable websites [6-8]. However, social media has shown potential in sharing child health information by perceived experts to caregivers [5,9]. Furthermore, digital communication channels between caregivers and health care professionals may improve family involvement in the health management of children [10]. In fact, the American Academy of Pediatrics Committee on Pediatric Emergency Medicine already agreed a decade ago that the practice of telemedicine should be supported to optimize the delivery of care for services that can be delivered via telemedicine [11]. Although the number of electronic consulting services has increased and emerging studies suggest that electronic consulting seems a feasible alternative to medical specialist's face-to-face appointments, only a few studies have focused on pediatric physician-led Web-messaging or chat-based consultation services [10,12-15].

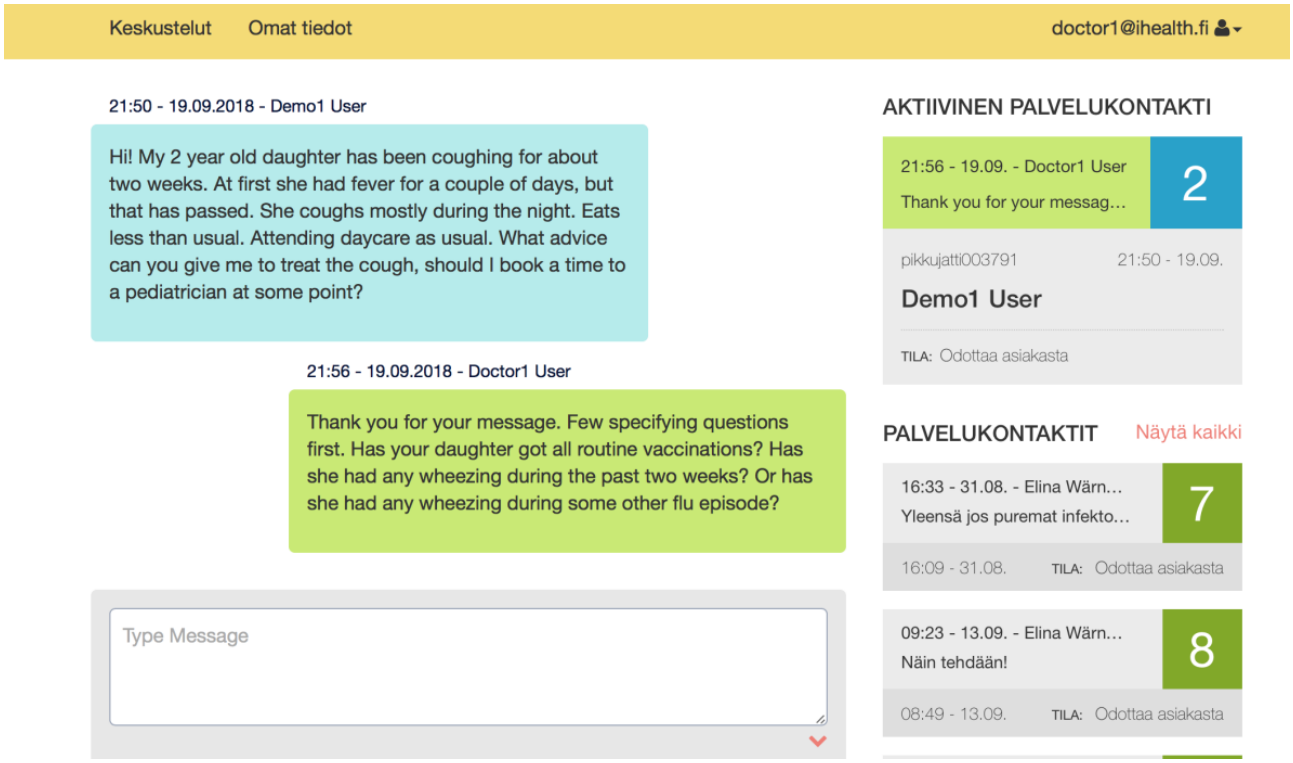
In this study, we aimed to describe Web-based chat conversations between caregivers and physicians to help others allocate human resources in an efficient way when planning a similar service. Furthermore, we evaluated whether physicians and caregivers considered the Web-based chat service helpful for caregivers and whether further face-to-face medical contact after Web-based chatting was needed. We hypothesized that most chat questions would be infection-focused, as described previously [5], that health information could be delivered to caregivers, and that satisfaction of the caregivers would be inversely associated with the need for further face-to-face medical contact.

Methods

A private medical center for children and youth with approximately 110,000 face-to-face appointments in 2015 established a Web-based chat service for caregivers of children in the greater Helsinki area in September 2015. The chat service was based on a secure Web platform accessible via any device with an internet connection. A video clip or picture could be attached to the chat message (Figure 1). The chat service was intended mainly for consultation and not necessarily to substitute an outpatient visit. During the study period, a chat consultation cost €19 per consultation. The fee was not tied to the duration of the chat or the advice given. Afterwards, all parents could get a partial reimbursement from the Social Insurance Institution of Finland and some of the patients also from their insurance companies. The chat service was open from 9 am to 9 pm local time, and one resident in pediatrics at a time was responsible for the Web-based chat consultations. For the study period of 6 months, 4 residents in pediatrics altogether worked for the chat service as a part-time job in their off-duty hours from the public hospital, where they were residents-in-training. In Finland, specializing for pediatrics includes a 6-year residency period (half in secondary and half in tertiary care hospitals) and a board exam. All 4 residents had completed at least 60% of their residency, and thus, all the chat physicians were considered to be qualified to answer questions from caregivers. The local ethics committee of the Helsinki University Hospital approved the study.

The study comprised 343 consecutive Web-based chat consultations and was started 1 month after launching the service to avoid the effects of possible initial technical problems. Chat consultations were immediately evaluated by the 4 chat physicians, and the following data were collected: child's age, time and total length of a chat, the caregiver's main concern for consultation, need for further face-to-face medical contact, whether the physician felt capable of answering the caregiver's question, and whether a prescription was given. Because all questions did not include sufficient information to set a specific diagnosis, the main concerns were grouped into 8 diagnosis groups based on common cases presenting at pediatric outpatient departments. Diagnosis groups were allergy, dermatology, endocrinology or growth problem, gastroenterology, infection, neonatology, neurology, nutrition, and trauma. In addition, an email including a link to a Web-based questionnaire was sent by authors to all caregivers 7-14 days after the chat consultation. The gap of at least a week was chosen because the timeline between a chat consultation and further face-to-face medical contact may vary in accordance with questions and concerns of caregivers. The questionnaire was answered by 28.6% (98/343) of caregivers a median of 11 (interquartile range, IQR, 7-20) days after consultation, and their responses could be matched with physicians' responses for analyses. The questionnaire was created specifically for this study, and the following data were collected from all the caregivers who answered the questionnaire: need for further face-to-face medical contact, how well a chat physician could answer a caregiver's question, and whether the caregiver's primary source of child health information was the internet. Both physicians' and caregivers' questions were answered on a 5-point scale (eg, caregivers and physicians judged whether caregivers' questions were answered very well, well, not well or poorly, quite poorly, or very poorly). However, the scale was then adjusted to a 2-point scale for studying the factors affecting how chat physicians could answer caregivers' questions.

Figure 1. Screenshot of a demo chat consultation from the Web-based chat service called "Pikkutsätti" initiated by Pikkujätti Medical Centre for Children and Youth.



Variables on a qualitative scale are presented as numbers with percentages and compared using chi-square test or Fisher's exact test, as appropriate. Variables on a continuous scale were visually assessed for normality using Kolmogorov-Smirnov test, described as mean (SD) or median with IQR, and compared using Student's *t* test or Mann-Whitney U test, as appropriate. A *P* value of ≤.05 was considered significant for all statistical analyses. Statistical analyses were performed using SPSS 21.0 (IBM Corp).

Results

The median age of the children of the caregivers who contacted the chat service was 2.1 (IQR 0.83-4.69) years, and 29.7% (102/343) of children were less than 1 year old (Figure 2). A mother contacted the chat service in 79.9% (274/343) of the cases, a father in 18.7% (64/343) of cases, and another caregiver in 1.5% (5/343) of cases. Of all chat conversations, 7.6% (26/343) lasted less than 15 minutes, 36.7% (126/343) lasted 15-30 minutes, 26.5% (91/343) lasted 30-45 minutes, and 29.2% (100/343) lasted more than 45 minutes. All chat consultations were initially responded to within 15 minutes of the first message from the caregiver, and the average response time was 5 (SD 2) minutes. The majority (119/343, 34.7%) of chat conversations took place from 9 am to noon, 21.6% (74/343) from noon to 3 pm, 21.9% (75/343) from 3 pm to 6 pm, and 21.0% (72/343) from 6 pm to 9 pm. The primary source of child health information was the internet for 57% (56/98) of the caregivers.

The most common concern for consultation was infection in 55.1% (189/343) of cases (Figure 3). When infection as a concern for consultation was compared with other reasons, no difference was apparent in child's age (*P*=.30) or in length (*P*=.22) or time (*P*=.43) of chat conversation. The chat physician gave a prescription in 20.7% (71/343) of cases, and there was no difference between the number of prescriptions given after chat conversations concerning infections (44/189, 23.2% cases) and those given after other concerns (27/154, 17.5% cases; *P*=.19).

Chat physicians recommended face-to-face medical contact on that same day in 13.7% (47/343) of cases, later (ie, after that day) in 15.2% (52/343) of cases, and if symptoms would worsen, for nearly half the cases (164/343, 47.8%). Face-to-face medical contact was recommended more often on that same day if the chat conversation concerned infection (36/189, 19.0% cases) compared with other concerns (11/154, 7.1% cases; *P*=.001). Chat physicians did not recommend more same day face-to-face medical contact for children younger than 12 months compared with that for older children (16/102, 15.7% vs 31/241,12.9%, *P*=.49). The chat physician did not recommend more same day face-to-face medical contacts after chat consultations lasting longer than 45 minutes compared with shorter chat consultations (36/243, 14.8% vs 11/100, 11.0%, *P*=.35).

Of all caregivers, 43% (42/98) took their children to a physician with respect to their main concern for chat consultation. In 57% (24/42) of these cases, the chat physician had advised face-to-face medical contact that same day or later. A total of 11% (11/98) of caregivers took their children to see a physician that same day, and 8 (73%) of those caregivers were advised to do so by a chat physician.

Physicians felt capable of answering caregivers' questions very well or well in nearly three-quarters (249/343, 72.6%) of cases. As for caregivers, 93% (91/98) received very good or good answers, and in 99% (90/91) of these cases, physicians gave a

similar assessment. However, 7% (7/98) of cases showed a discrepancy in response, that is, the physician felt capable of answering the caregiver's question very well or well, but the caregiver felt he or she received a poor answer. Whether caregivers had to take their children to see a physician that same day, whether the caregiver's main concern was infection, whether the child was younger than 12 months, or whether the chat conversation lasted longer than 45 minutes did not associate with whether caregivers thought the physician answered their question well or poorly (Table 1). However, physicians felt capable of answering caregivers' questions better when the main concern was infection (P=.02) and when the chat conversation lasted less than 45 minutes (P<.001; Table 1).

Figure 2. Age distribution of patients.

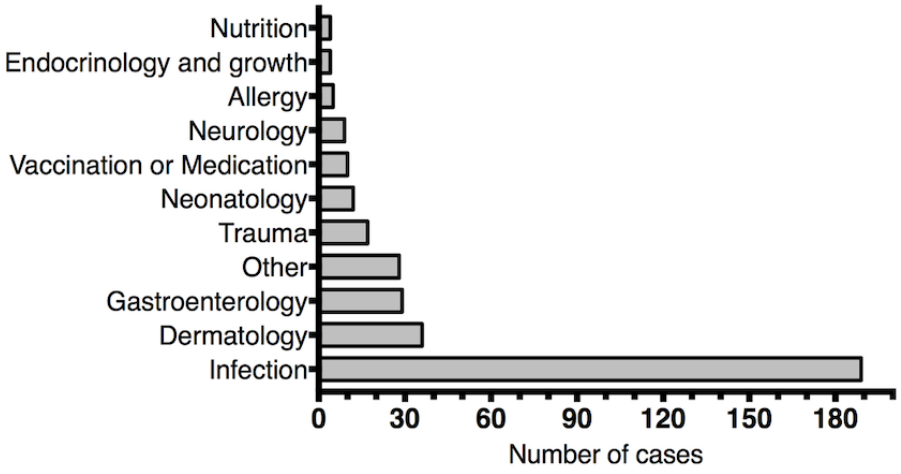


Figure 3. The distribution of main concerns leading to chat consultations.

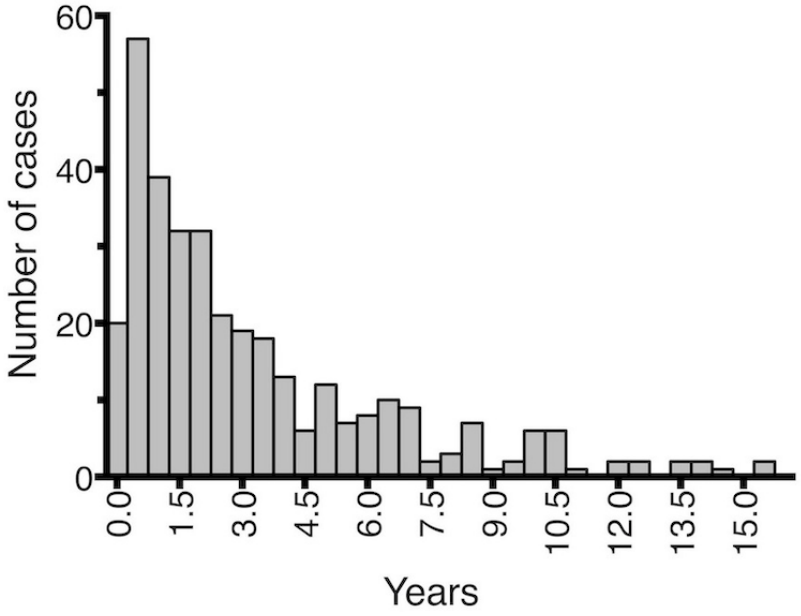


Table 1. Assessment of chat conversations by caregivers (n=98) and by physicians (N=343 chats).

Assessment of chat conversations	Yes, n (%)	No, n (%)	*P* value
Caregivers thought physicians gave a helpful answer			
Child <1-year-old	26 (96.3)	1 (3.7)	.67
Child ≥1-year-old	65 (91.5)	6 (8.5)	
Infection as a main concern	57 (95.0)	3 (5.0)	.43
Other concerns	34 (89.5)	4 (10.5)	
Length of chat >45 min	26 (89.7)	3 (10.3)	.42
Length of chat ≤45 min	65 (94.2)	4 (5.8)	
Need to visit a physician that same day	10 (90.9)	1 (9.1)	.58
No need to visit a physician that same day	81 (93.1)	6 (6.9)	
Physicians felt capable of giving a good answer			
Recommended face-to-face medical contact for that same day	44 (93.6)	3 (6.4)	.47
Did not recommend face-to-face medical contact for that same day	283 (95.6)	13 (4.4)	
Child <1-year-old	95 (93.1)	7 (6.9)	.26
Child ≥1-year-old	232 (96.3)	9 (3.7)	
Infection as a main concern	185 (97.9)	4 (2.1)	.02
Other concern	142 (92.2)	12 (7.8)	
Length of chat >45 min	89 (89.0)	11 (11.0)	<.001
Length of chat ≤45 min	238 (97.9)	5 (2.1)	

Discussion

Principal Findings

This descriptive study offers novel information about pediatric physician-led Web-based chat services, which may provide an easy e-Consultation channel for caregivers with a variety of concerns about their child's health or illness. A thorough description of chat conversations between caregivers and physicians is essential to improve chat services and to help others allocate sufficient and focused human resources when setting up a similar service.

In our study, most chat consultations took place during the morning between 9 am and noon. In line with our hypothesis, infections were the most common concern for Web-based chat consultations. Similarly, infections were the most common topic for parental concerns at a child health social media site, which was based on a question-and-answer service produced by a pediatrician [5]. This is logical because infections are a common cause for both ambulatory and emergency pediatrician visits. A recent study showed that healthy children have an average of 9 parent-reported infections during the first 2 years of life [16]. This high infection rate in children under the age of 2 may explain, at least partly, the low age of the children in this study [16]. Furthermore, previous studies have shown that the young age of children predisposes caregivers to use the internet to gather health information [4,17].

Half of the chat conversations lasted more than 30 minutes, which is consistent with a previous study reporting that the median length of a nurse-led direct Web-based chat triage service launched by the United Kingdom National Health Institute was 30 minutes [13]. Although the United Kingdom National Health Institute considered Web-based chat conversations too long and, therefore, too expensive, the patients were satisfied with the chat length [13]. Despite the time-consuming nature, several chat conversations can be managed at the same time and the physician can perform other tasks while on chat duty. Furthermore, chat connections can be used almost anywhere, be dropped briefly for a more urgent issue, and then be returned to a moment later compared with videoconference appointments or phone calls. Therefore, we conjecture that in the nonacute, low-level medical consultation need of a caregiver, a chat could be a more easily adopted form of first-line communication than a video call. However, in certain cases, a switch to video could be useful, especially for families living in rural and medically underserved communities [18].

In a previous study, 95% of caregivers for children with hemangioma found Web-based electronic health interventions for infantile hemangiomas reliable, and 98% of the caregivers would recommend the intervention to other parents [19]. Correspondingly, the vast majority of caregivers in our study reported that they received good or very good answers to their questions. Therefore, a Web-based chat service may be a useful way to help caregivers with various concerns about their child's health or illness. However, the proportion of physicians who felt capable of answering caregiver's question very well or well was somewhat lower. This discrepancy may have resulted mainly from physicians' self-criticism [20]. Physicians felt capable of answering caregivers' questions better when the chat consultation was focused on infection compared with that on other concerns, but parental satisfaction was not linked with the

subject of the chat consultation. Physicians also felt capable of answering caregivers' questions better when the chat conversation lasted less than 45 minutes. Although the time spent waiting for a response may lengthen chat conversations, the length may also reflect the complexity of the parental concern. Furthermore, the fact that the chat physicians were residents in pediatrics instead of licensed pediatricians may also have lengthened discussions. However, the length of chat conversation was not linked with parental satisfaction.

In 2007, half of European people rated the internet as an important source of health information, preceded only by face-to-face contact with health care professionals [3]. In our study, a decade later, the primary source of child health information was the internet for more than half of caregivers. In previous studies, nearly all caregivers reported some use of the internet for their children's health information, and a fifth of caregivers reported use before attending a pediatric outpatient clinic [4,17,21].

Based on a recent study, 30% of e-Consultations may lead to face-to-face medical contact [22]. In this study, chat physicians recommended face-to-face medical contact for that same day in only about a tenth of the cases, but in 15% of the cases, face-to-face contact was recommended later. If the chat conversation concerned infection, face-to-face medical contact was recommended in a fifth of the cases. Because the majority of caregivers were not advised to or did not seek further face-to-face consultation at all, some families may have avoided needing face-to-face medical contact thanks to the chat service. However, this study cannot reliably answer the question of whether a Web-based chat consultation can replace face-to-face medical contact. In addition, whether caregivers had to take their children to see a physician that same day was not associated with whether caregivers thought the physicians answered their question well or poorly. Therefore, it is likely that the caregivers may contact Web-based chat services not only to replace a face-to-face consultation but also to gather information [17].

Limitations

There are some limitations to our study, the first being the low response rate from parents. A gap of 1-2 weeks before sending the survey to caregivers may have lowered the response rate. Furthermore, since only a quarter of caregivers completed the questionnaire, it is possible that unsatisfied or highly satisfied parents are overrepresented. However, with a high satisfaction rate, we believe our findings show the general applicability of a Web-based chat service for health information distribution to caregivers. Second, we relied on parental reporting on subsequent visits to physicians, therefore, using subjective data. However, we believe that caregivers would have taken their children to a physician by the time they answered the questionnaire. This is further supported by a recent report showing that most consultations via telephone, email, or e-Consult are followed by another consultation within 14 days [22]. Third, in addition to the initial response time, further delay in both physician and caregiver responses was not recorded and may have inappropriately lengthened chat conversations. Fourth, although a private medical center for children established this chat service, we conjecture that similar chat services could be useful in both private and public sectors.

Conclusions

Both caregivers and physicians considered that the concerns of caregivers were well handled and the vast majority of caregivers' questions could be well answered in a Web-based chat. Thus, a pediatric Web-based chat service provided for caregivers of children may be a useful way to help caregivers with concerns about their child's health or illness. However, there are a number of factors that should be taken into account when setting up a Web-based physician-led chat service: physicians should have enough time for chat consultations, especially in morning hours, and they should have sufficient knowledge, especially of pediatric infections and the health of very young children.

Authors' Contributions

AK conceptualized and designed the study, collected data, carried out statistical analyses, drafted the initial manuscript, and revised the manuscript. BAS carried out statistical analyses, drafted the initial manuscript, and reviewed and revised the manuscript. EW, LK, and TT conceptualized and designed the study, collected data, and reviewed and revised the manuscript. OH conceptualized and designed the study, drafted the initial manuscript, and reviewed and revised the manuscript.

Conflicts of Interest

OH is a shareowner of iHealth Finland Ltd, which originally constructed the technical chat platform. AK acts as a guarantor for the integrity of the data collection and analyses.

References

1. Bujnowska-Fedak MM. Trends in the use of the Internet for health purposes in Poland. BMC Public Health 2015 Feb 27;15:194 [FREE Full text] [doi: 10.1186/s12889-015-1473-3] [Medline: 25886280]
2. Greenberg AJ, Serrano KJ, Thai CL, Blake KD, Moser RP, Hesse BW, et al. Public use of electronic personal health information: Measuring progress of the Healthy People 2020 Objectives. Health Policy Technol 2017 Mar;6(1):33-39 [FREE Full text] [doi: 10.1016/j.hlpt.2016.08.003] [Medline: 28480160]
3. Kummervold P, Chronaki C, 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 Nov 17;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]

4. Sebelefsky C, Karner D, Voitl J, Klein F, Voitl P, Böck A. Internet health seeking behaviour of parents attending a general paediatric outpatient clinic: A cross-sectional observational study. J Telemed Telecare 2015 Oct;21(7):400-407. [doi: 10.1177/1357633X15583431] [Medline: 26026180]
5. Helve O. A medical consultation service on Facebook: descriptive analysis of questions answered. J Med Internet Res 2014 Sep 04;16(9):e202 [FREE Full text] [doi: 10.2196/jmir.3194] [Medline: 25189182]
6. 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] [Medline: 23153164]
7. Khoo K, Bolt P, Babl FE, Jury S, Goldman RD. Health information seeking by parents in the Internet age. J Paediatr Child Health 2008;44(7-8):419-423. [doi: 10.1111/j.1440-1754.2008.01322.x] [Medline: 18564080]
8. Vance K, Howe W, Dellavalle RP. Social internet sites as a source of public health information. Dermatol Clin 2009 Apr;27(2):133-6, vi. [doi: 10.1016/j.det.2008.11.010] [Medline: 19254656]
9. Stroever SJ, Mackert MS, McAlister AL, Hoelscher DM. Using social media to communicate child health information to low-income parents. Prev Chronic Dis 2011 Nov;8(6):A148 [FREE Full text] [Medline: 22005641]
10. Armoiry X, Sturt J, Phelps EE, Walker C, Court R, Taggart F, et al. Digital Clinical Communication for Families and Caregivers of Children or Young People With Short- or Long-Term Conditions: Rapid Review. J Med Internet Res 2018 Jan 05;20(1):e5 [FREE Full text] [doi: 10.2196/jmir.7999] [Medline: 29305339]
11. American Academy of Pediatrics Committee on Pediatric Emergency Medicine, Yamamoto L. Access to optimal emergency care for children. Pediatrics 2007 Jan;119(1):161-164. [doi: 10.1542/peds.2006-2900] [Medline: 17200284]
12. Bergmo TS, Wangberg SC, Schopf TR, Solvoll T. Web-based consultations for parents of children with atopic dermatitis: results of a randomized controlled trial. Acta Paediatr 2009 Feb;98(2):316-320. [doi: 10.1111/j.1651-2227.2008.01033.x] [Medline: 18795905]
13. Binford Hopf RB, Le Grange D, Moessner M, Bauer S. Internet-based chat support groups for parents in family-based treatment for adolescent eating disorders: a pilot study. Eur Eat Disord Rev 2013 May;21(3):215-223. [doi: 10.1002/erv.2196] [Medline: 22949134]
14. Eminovic N, Wyatt J, Tarpey A, Murray G, Ingrams G. First evaluation of the NHS direct online clinical enquiry service: a nurse-led web chat triage service for the public. J Med Internet Res 2004 Jun 02;6(2):e17 [FREE Full text] [doi: 10.2196/jmir.6.2.e17] [Medline: 15249266]
15. Zandbelt LC, de Kanter FE, Ubbink DT. E-consulting in a medical specialist setting: Medicine of the future? Patient Educ Couns 2016 May;99(5):689-705. [doi: 10.1016/j.pec.2015.11.005] [Medline: 26597384]
16. Rosendahl J, Valkama S, Holmlund-Suila E, Enlund-Cerullo M, Hauta-Alus H, Helve O, et al. Effect of Higher vs Standard Dosage of Vitamin D3 Supplementation on Bone Strength and Infection in Healthy Infants: A Randomized Clinical Trial. JAMA Pediatr 2018 Jul 01;172(7):646-654. [doi: 10.1001/jamapediatrics.2018.0602] [Medline: 29813149]
17. Sebelefsky C, Voitl J, Karner D, Klein F, Voitl P, Böck A. Internet use of parents before attending a general pediatric outpatient clinic: does it change their information level and assessment of acute diseases? BMC Pediatr 2016 Dec 18;16:129 [FREE Full text] [doi: 10.1186/s12887-016-0677-8] [Medline: 27538782]
18. Marcin JP, Ellis J, Mawis R, Nagrampa E, Nesbitt TS, Dimand RJ. Using telemedicine to provide pediatric subspecialty care to children with special health care needs in an underserved rural community. Pediatrics 2004 Jan;113(1 Pt 1):1-6. [Medline: 14702439]
19. de Graaf M, Totte J, Breugem C, van Os-Medendorp H, Pasmans S. Evaluation of the Compliance, Acceptance, and Usability of a Web-Based eHealth Intervention for Parents of Children With Infantile Hemangiomas: Usability Study. JMIR Res Protoc 2013 Dec 17;2(2):e54 [FREE Full text] [doi: 10.2196/resprot.2897] [Medline: 24345450]
20. Violato C, Lockyer J. Self and peer assessment of pediatricians, psychiatrists and medicine specialists: implications for self-directed learning. Adv Health Sci Educ Theory Pract 2006 Aug;11(3):235-244. [doi: 10.1007/s10459-005-5639-0] [Medline: 16832707]
21. Pehora C, Gajaria N, Stoute M, Fracassa S, Serebale-O'Sullivan R, Matava CT. Are Parents Getting it Right? A Survey of Parents' Internet Use for Children's Health Care Information. Interact J Med Res 2015 Jun 22;4(2):e12 [FREE Full text] [doi: 10.2196/ijmr.3790] [Medline: 26099207]
22. Atherton H, Brant H, Ziebland S, Bikker A, Campbell J, Gibson A, et al. The potential of alternatives to face-to-face consultation in general practice, and the impact on different patient groups: a mixed-methods case study. Health Services and Delivery Research 2018:6. [doi: 10.3310/hsdr06200] [Medline: 29889485]

Abbreviations

IQR: interquartile range

Edited by G Eysenbach; submitted 22.02.18; peer-reviewed by S Badawy, M Hawkins, S Porter; comments to author 03.08.18; revised version received 26.09.18; accepted 10.10.18; published 14.12.18.

Please cite as:
Kaskinen A, Ayeboa-Sallah B, Teivaanmäki T, Wärnhjelm E, Korhonen L, Helve O
Pediatric Web-Based Chat Services for Caregivers of Children: Descriptive Study
J Med Internet Res 2018;20(12):e10165
URL: https://www.jmir.org/2018/12/e10165/
doi:10.2196/10165
PMID:30552087

Original Paper

Meeting Sexual Partners Through Internet Sites and Smartphone Apps in Australia: National Representative Study

Lucy Watchirs Smith[1], MPH; Rebecca Guy[1], PhD; Louisa Degenhardt[2], PhD; Anna Yeung[3], PhD; Chris Rissel[4], PhD; Juliet Richters[1], PhD; Bette Liu[5], DPhil

[1]Kirby Institute, University of New South Wales, Sydney, Sydney, Australia

[2]National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Sydney, Australia

[3]Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada

[4]Sydney School of Public Health, University of Sydney, Sydney, Australia

[5]School of Public Health and Community Medicine, University of New South Wales, Sydney, Sydney, Australia

Corresponding Author:
Lucy Watchirs Smith, MPH
Kirby Institute
University of New South Wales, Sydney
Level 6, Wallace Wurth Building
Sydney, 2052
Australia
Phone: 61 293859254 ext 9254
Fax: 61 293850920
Email: lwatchirs-smith@kirby.unsw.edu.au

Abstract

Background: Studies have reported on the proportion of the population looking for potential sexual partners using internet sites and smartphone apps, but few have investigated those who have sex with these partners, arguably a more important target group for health promotion.

Objective: This study aimed to determine the proportion of people who have had sex with someone they met on an internet site or a smartphone app in the previous year.

Methods: We analyzed data from the 2012-2013 Second Australian Study of Health and Relationships, a nationally representative telephone survey of Australian residents aged 16-69 years (N=20,091). The participation rate for the telephone survey was 66.22%. The prevalence of looking for a potential partner, physically meeting, and having sex with someone first met through an internet site or a smartphone app was estimated. Multivariate logistic regression was used for men and women separately to determine demographic and behavioral factors associated with having had sex with someone met on an internet site or a smartphone app in the last year.

Results: Overall, 12.09% of respondents had looked for potential partners using these technologies and 5.40% had done so in the last year. In the last year, 2.98% had met someone in person and 1.95% reported having had sex with someone first met on an internet site or a smartphone app. The prevalence of all behaviors was greater in men than in women and in younger respondents than in older respondents. Among sexually active men, factors associated with having had sex with someone met using internet sites or smartphone apps included identifying as gay or bisexual (adjusted odds ratio, AOR: 15.37, 95% CI 8.34-28.35), having either 2-3 or >3 sexual partners in the last year (AOR: 9.20, 95% CI 9.20-34.68 and AOR: 35.77, 95% CI 18.04-70.94, respectively), having had a sexually transmissible infection (STI) test in the past year (AOR: 2.02, 95% CI 1.21-3.38), or an STI in the last year (AOR: 3.15, 95% CI 1.25-7.97). Among sexually active women, factors associated with having had sex with someone met on an internet site or a smartphone app were as follows: having either 2-3 or >3 sexual partners in the last year (AOR: 32.01, 95% CI 13.17-77.78 and AOR: 71:03, 95 % CI 27.48-183.57, respectively), very low and low income (vs very high AOR: 3.40, 95% CI 1.12-10.35), and identifying as lesbian or bisexual (AOR: 2.27, 95% CI 1.04-4.49).

Conclusions: More than a third of adults who had looked for potential partners using websites and apps each year had sex with such partners, and those who had done so were more sexually active, suggesting that dating and hookup websites and applications are suitable settings for targeted sexual health interventions.

(J Med Internet Res 2018;20(12):e10683) doi:10.2196/10683

KEYWORDS

dating websites; internet; mobile phone; sexually transmitted diseases; health risk behaviors

Introduction

A range of dating websites, accessible via the internet or smartphone apps, are now available to search for potential sexual partners. These sites first surfaced in 1995 with match.com and eHarmony in 2000; in 2009, Grindr was launched, targeting men who have sex with men, followed by Tinder in late 2012 (targeted more toward heterosexually active adults). Since then, increasing numbers of these apps have become available. People have various motivations for using these sites and apps; some may be searching for a life partner and others for just a one-off encounter. The platforms enable selection of partners based on preferred personal characteristics, and some sites use geospatial technology to allow the user to determine the geographical proximity of a potential partner (eg, both Grindr and Tinder are location-based hookup apps). Sites are also available for particular cultural groups, and some focus on certain sexual preferences. The sites are generally open to people aged ≥18 years.

Beyond that, dating sites have the potential to provide sexual health promotion interventions. However, there is little available information on how many people access these sites and what their characteristics are. To date, most studies of meeting partners online have recruited specific populations and used convenience-sampling strategies, such as targeting online users, gay venues, or health care settings [1-3]. However, these settings are not representative and may result in an overestimation of prevalence.

Furthermore, earlier studies reported on the proportion and characteristics of people who looked for partners using Web-based technologies (but may not have intended to have sex with them or actually have done so). However, the characteristics of people who have sex with these partners are of the greatest relevance for health promotion. Of the two population-level studies conducted to date, a study among Norwegian young people (aged 15-20 years) in 2009 found that 30% reported ever having had sex with someone they met online (but did not ask about the last year), and a British survey of adults (aged 16-74 years), conducted from 2010 to 2012, focused only on respondents looking for sexual partners in the last year, not whether respondents had sex with them [4,5].

In 2012-2013, the Second Australian Study of Health and Relationships (ASHR2) survey was conducted just after the introduction of Tinder and other geosocial dating apps [6]. ASHR2 is a national representative survey of the Australian population covering demographics, knowledge, attitudes, behaviors, and experiences related to sexual health. A series of questions about looking for, physically meeting, and having sex with people met on websites or smartphone apps were also asked [7,8], providing an opportunity to determine both the prevalence of Australian adults who had looked for potential partners on websites and apps and the characteristics of people who had sex with these partners in the past year.

Methods

Study Population

This study was a cross-sectional analysis of data from ASHR2. The methods of ASHR2 have been described elsewhere [7]. In brief, ASHR2 is a national survey of 20,091 Australian residents aged 16-69 years. Data were collected in 2012 and 2013 via a computer-assisted telephone survey by trained interviewers. The study sample was selected using a modified random digit dialing sampling frame, which combined random digit dialing of landlines with that of cell phones. The overall participation rate among eligible people was 66.22%; the study population has been shown to be broadly representative of the Australian population, except for an overrepresentation of people with postgraduate degrees [7].

To allocate resources efficiently and gather more information from those with potentially higher HIV and sexually transmissible infection (STI) risk, we administered interviews in two forms [7]. All respondents who reported no sexual partners or >1 sexual partner in the previous year or any lifetime same-sex experience were given a long-form interview, as were a 20.00% random sample of survey respondents who had reported having 1 partner in the previous year and no same same-sex experience; the remaining 80.00% of one-partner respondents were given the short-form interview. Questions asked only in the long-form interview included those on meeting and having sex with a partner met on websites and smartphone apps.

Statistical Analysis

The estimates of prevalence included sexually active and sexually inactive respondents. However, the predictor analysis of factors associated with meeting and having had sex with someone met on websites or smartphone apps was restricted to sexually active respondents because many sexual health outcomes were queried only of sexually active survey respondents. For this study, people were considered sexually active if they had ≥1 partners (for vaginal or anal intercourse or oral or manual sex) of the same or other sex in the previous 12 months. Respondents who reported no lifetime sexual experience were coded as not sexually active.

Textbox 1. Questions related to use of internet and smartphone apps from the Second Australian Study of Health and Relationships.

1. Have you ever used an internet site or smartphone application to look for potential partners? Have you done so in the past year?
2. In the last year have you met someone in person that you first met on an internet site?
3. And did you have sex with that person?

Outcome Measure

The primary outcome of this study was having sex with someone met on a website or a smartphone app in the past year. We also calculated the proportion of people looking for partners and meeting someone in person who they first met on a website or mobile app. These outcomes were ascertained using the following questions collected in the long-form questionnaire. The questions' exact wording is shown in Textbox 1.

The proportion of people searching for, meeting, and having sex with partners on websites and smartphone apps were calculated separately using descriptive statistics. Data were weighted according to the Australian population and the probability of being selected for the long-form questionnaire. The characteristics of respondents who reported searching for partners using these technologies were compared with respondents who reported having sex with someone met on a website or a smartphone app. A chi-square test was used to compare differences in distributions between groups for a range of covariates.

Univariate and multivariate logistic regression, weighted in accordance with study procedures, were used to examine factors associated with having had sex with someone met on a website or a smartphone app in the last year. All data were analyzed using Stata statistical software version 14. Variables significant at the $P<.1$ level in the univariate analysis were included in multivariate logistic regression analyses. Backward elimination of variables was then used to determine the final adjusted model.

The demographic covariates included in the models were age group (16-29 years, and then 10-year age groups up to 69 years), language spoken at home (English or other), annual household income: very low or low (<Aus $52,000), middle (Aus $52,001-Aus $83,000), high (Aus $83,001-Aus $125,000), and very high (>Aus $125,000), and area of residence (urban or rural and remote) according to the Accessibility/Remoteness Index of Australia [9].

The behavioral covariates included in models were levels of alcohol consumption (high or not, with high alcohol consumption classified as >28 standard drinks per week for men and >14 standard drinks per week for women), injecting drug use in the last year (yes or no), smoking status (never and former, or current), sexual identity (heterosexual or gay, lesbian, bisexual, and other), condom use at last event (used or did not use condoms), STI history the last year (no STI test, STI test, or STI diagnosis), and sexual partner numbers in the previous year (1, 2-3, or >3). The numbers of sexual partners included both male and female partners. In relation to STIs, respondents were asked whether they had had an STI in the past year and whether they had an STI test; these two questions were combined to provide the composite variable. The STIs included were pubic lice, genital warts, chlamydia, genital herpes, gonorrhea, and syphilis; in addition, for women, warts virus on Pap smear, pelvic inflammatory disease, bacterial vaginosis or gardnerella, and trichomoniasis were included and for men, nonspecific urethritis and anal warts were included [10].

Ethical Approval

The study received La Trobe University's (HEC 11-040) ethical approval, which was ratified by the ethics committees of the University of New South Wales, the University of Sydney, and the University of Sussex.

Results

Prevalence of Looking for Potential Partners on Websites and Smartphone Apps and Meeting Them in Person

Overall, 12.09% (2346/19,398) of respondents reported ever searching for potential partners on websites and smartphone apps (13.52% men [1320/9761], 10.65% women [1026/9637]) and 5.40% (1048/19,398) of respondents (7.01% men [685/9637] and 3.77% women [364/9636]) reported doing so in the last year. Table 1 shows that searching for potential partners using smartphone apps and websites in the last year was most common among people aged 16-29 years (8.42%, 435/5169) and decreased with increasing age to 1.87% (52/2785) among people aged 60-69 years. Furthermore, 4.92% (815/16586) of sexually active respondents used websites and smartphone apps to look for potential partners in the last year, and 8.28% (233/2811) of sexually inactive respondents did. The activity was more common among gay, lesbian, and bisexual respondents (25.32%, 172/680) than among heterosexual respondents (4.68%, 876/19715).

Table 1. Prevalence of looking for potential partners on websites and smartphone apps and prevalence of having sex with these partners.

Characteristics	Searched for potential partner (ever), %[a] (95% CI)	Searched for potential partner (last year), %[a] (95% CI)	Met in person (last year), %[a] (95% CI)	Had sex (last year), %[a] (95% CI)
All participants[b,c]	12.09 (11.24-13.00)	5.40 (4.89-5.97)	2.98 (2.63-3.37)	1.95 (1.69-2.25)
Sex				
Men	13.52 (12.26-14.88)	7.01 (6.20-7.93)	3.70 (3.14-4.36)	2.54 (2.14-3.02)
Women	10.65 (9.54-11.87)	3.77 (3.17-4.48)	2.19 (1.81-2.65)	1.35 (1.04-1.74)
Sexually active in last year				
Yes	11.41 (10.46-12.44)	4.92 (4.35-5.55)	2.97 (2.58-3.42)	N/A[d]
No	16.10 (14.58-17.74)	8.28 (7.16-9.56)	3.01 (2.36-3.85)	N/A
Age group (years)				
16-29	14.24 (12.25-16.50)	8.42 (7.05-10.02)	4.73 (3.85-5.79)	3.03 (2.41-3.81)
30-39	15.89 (13.71-18.35)	6.30 (5.07-7.80)	3.57 (2.68-4.73)	2.25 (1.66-3.05)
40-49	11.83 (10.21-13.67)	4.78 (3.92-5.82)	2.40 (1.87-3.08)	1.79 (1.33-2.42)
50-59	10.0 (8.54-11.67)	3.52 (2.78-4.45)	2.00 (1.42-2.80)	1.23 (0.79-1.92)
60-69	5.84 (4.74-7.17)	1.87 (1.43-2.44)	0.98 (0.66-1.47)	0.54 (0.29-0.99)
Sexual identity				
Heterosexual	10.98 (10.12-11.90)	4.68 (4.17-5.25)	2.40 (2.07-2.78)	1.45 (1.21-1.73)
Homosexual or lesbian or bisexual	42.80 (37.97.0-47.78)	25.32 (21.40-29.69)	18.9 (15.48-22.87)	15.8 (12.66-19.55)
Language spoken at home				
English only	12.14 (11.27-3.06)	5.26 (4.76-5.81)	2.98 (2.63-3.37)	2.02 (1.75-2.29)
Other	11.44 (7.95-16.19)	7.47 (4.66-11.76)	2.94 (1.37-6.17)	0.89 (0.380-2.10)
Annual household income				
Very low or low	11.89 (10.08-13.98)	6.60 (5.21-8.32)	3.98 (2.43-3.19)	2.49 (1.77-3.46)
Middle	10.48 (8.57-12.75)	4.65 (3.62-5.96)	1.92 (1.39-2.64)	1.12 (0.75-1.73)
High	9.52 (7.72-11.68)	2.43 (1.65-3.55)	0.94 (0.62-1.40)	0.55 (0.33-0.93)
Very high	9.76 (7.98-11.)	2.70 (1.94-3.75)	1.58 (1.04-2.41)	1.03 (0.72-1.46)
Area of residence[e]				
Urban	12.92 (11.82-14.09)	5.62 (4.97-6.36)	3.37 (2.90-3.91)	2.11 (1.78-2.51)
Regional or remote	10.25 (8.98-11.7)	4.94 (4.14-5.88)	2.13 (1.71-2.64)	1.52 (1.17-1.96)
High alcohol consumption				
No	12.47 (11.45-13.56)	5.51 (4.89-6.21)	2.91 (2.51-3.36)	1.86 (1.57-2.20)
Yes	11.20 (9.72-12.87)	5.14 (4.27-6.18)	3.15 (2.28-3.99)	2.10 (1.61-2.75)
Injected drugs in last year				
No	11.93 (11.07-12.85)	5.28(4.76-5.86)	2.93 (2.57-3.33)	191 (1.65-2.22)
Yes	19.08 (14.21-25.13)	10.50 (7.43-14.63)	4.90 (3.00-7.90)	3.6 (2.00-6.43)
Smoking status				
Never smoked/former	11.09 (10.15-12.10)	4.65 (4.10-5.27)	2.6 (2.25-3.01)	1.79 (1.51- 2.12)
Current smoker	16.75 (14.63-19.12)	9.07 (7.63-10.74)	4.86 (3.79-6.21)	2.88 (2.20-3.77)
STI[f] testing in last year				
No test	9.13 (8.20-10.14)	3.32 (2.81-3.92)	1.70 (1.38-2.09)	1.18 (0.94-1.48)
STI test	23.86 (20.38-27.73)	13.57 (11.20-16.35)	9.55 (7.73-11.76)	7.24 (5.84-8.93)
STI diagnosis	36.96 (26.46-17.39)	26.00 (18.05-35.92)	21.13 (14.21-30.23)	16.88 (10.89-25.25)

Characteristics	Searched for potential partner (ever), %[a] (95% CI)	Searched for potential partner (last year), %[a] (95% CI)	Met in person (last year), %[a] (95% CI)	Had sex (last year), %[a] (95% CI)
Condom use with most recent partner				
Used condoms	14.69 (12.50-17.18)	9.35 (7.79-11.17)	5.78 (4.69-7.13)	4.52 (3.55-5.74)
Did not use	12.11 (10.49-13.94)	4.51 (3.66-5.54)	2.70 (2.12-3.45)	1.92 (1.51-2.44)
Number of sexual partners in last year				
1	9.66 (8.80-10.61)	3.0 (2.6-3.6)	1.20 (0.93-1.54)	0.47 (0.31-0.69)
2-3	33.99 (30.57-37.51)	25.4 (22.3-28.7)	16.9 (14.31-19.85)	12.40 (10.19-15.00)
>3	39.44 (34.53-44.58)	33.0 (28.4-38.0)	27.32 (22.99-32.12)	23.91 (19.87-28.49)

[a]All proportions have been weighted to match the Australian population.

[b]N=19,398 (8184), weighted (unweighted) denominators

[c]Individuals with missing data are not shown; this was <5% for all variables except for income, which was incomplete for 24.5% of participants.

[d]N/A: not applicable.

[e]Accessibility/Remoteness Index of Australia.

[f]STI: sexually transmissible infection.

Having met in person was reported by 2.98% (578/19,398) of survey respondents (3.70% men [363/9761], 2.19% women [214/9637]), whereas having had sex with someone first met on a website or a smartphone app was reported by 1.95% (378/19,398) of respondents (2.54% men [248/9761], 1.35% women [130/9637]). Having had sex with someone met on an internet site or a smartphone app in the last year was the highest among respondents who identified as gay, lesbian, or bisexual (15.80%, 107/680) and also more frequent among people aged 16-29 years and 30-39 years, as well as those who had had an STI test or STI diagnosis in the past year, those with a higher number of sexual partners, and those who used a condom at their last sexual event (Table 1). Results from Table 1 stratified by sex are presented in Multimedia Appendices 1 and 2.

When restricted to sexually active respondents,11.41% (1893/16586) had ever searched for potential partners online, 4.92% (815/16586) had done so in the last year, and 2.97% (493/16586) had met with someone in person in the last year (Table 1).

Table 2 displays the proportion of respondents who had sex with someone met on a website or a smartphone app, among those who reported using websites and smartphone apps to search for potential partners in the last year overall and according to selected characteristics. Overall, slightly over one-third (36.07%, 378/1048) of those who searched for partners also reported having had sex with someone they met using these technologies in the last year. There were no statistical differences by sex. Differences were observed for sexual identity, with gay, lesbian, and bisexual respondents being more likely also to report having sex with someone they met on a website or a smartphone app, than heterosexual respondents (62.41% [107/172] vs 30.89% [271/876], P<.001). Differences were also seen for STI testing history (no test: 35.89% [170/473], STI test: 53.77% [171/318], and STI diagnosis: 64.94% [33/51], P<.001), condom use at most recent event (used condoms: 48.43%[151/312] vs not used: 42.85% [122/285], P<.001), and numbers of sexual partners (1 partner: 15.36%, [83/541], 2-3 partners: 48.82% [150/307], and >3 partners 72.44% [145/200], P<.001).

Table 2. Proportion of respondents reporting having had sex with someone met on a website or a smartphone app among people who searched for partners on websites and smartphone apps.

Characteristics	%[a] (95% CI)	*P* value
All participants[b]	36.07 (31.64-40.75)	N/A[c]
Sex		
Men	36.28 (30.90-42.02)	.90
Women	35.68 (28.20-43.93)	N/A
Sexual identity		
Heterosexual	30.89 (26.19-36.03)	<.001
Homosexual or lesbian or bisexual	62.41 (53.24-70.75)	N/A
STI[d] testing in the last year		
No test	35.89 (28.86-43.59)	<.001
STI test	53.77 (43.94-63.31)	N/A
STI diagnosis	64.94 (46.76-79.61)	N/A
Condom use with most recent partner		
Used condoms	48.43 (39.7 57.2)	<.001
Did not use	42.85 (33.5-52.8)	N/A
Missing or don't know or refused	23.26 (18.0-29.5)	N/A
Number of sexual partners in the last year		
1	15.36 (10.50-21.94)	<.001
2-3	48.82 (41.66-56.02)	N/A
>3	72.44 (63.95-79.57)	N/A

[a]All data have been weighted to match the Australian population.

[b]n=1048 (720) weighted (unweighted) denominators.

[c]N/A: not applicable.

[d]STI: sexually transmissible infection.

Among the respondents who had had sex with someone they met on a website or a smartphone app, 62.88% (363/578) were male and 37.12% (214/578) were female. The majority of people who reported having had sex with someone met on a website or a smartphone app identified as heterosexual (77.75%, 129/578). In terms of STI testing, 41.91% (242/578) reported not having had an STI test in the previous year, 38.81% (224/578) reported having had an STI test, and 7.24% (42/578) reported being diagnosed with an STI in the last year (12.03%, 70/578) of respondents either refused to answer, could not recall, or were not asked). Regarding the numbers of sexual partners, 36.45% (211/578) had 1 sexual partner in the last year, 35.28% (204/578) had 2 or 3 sexual partners, and 28.27% (163/578) had >3 sexual partners.

Correlates of Having Sex With Someone Met Online

Men

Among sexually active males, most respondents were heterosexual (97.34%, 8339/8567), spoke English at home (93.80% (8012/8526), and resided in an urban area (68.04%, 5788/8526). Age was distributed as follows: 24.24% (2077/8567) were aged 16-29 years, 21.03% (1802/8567) were aged 30-39 years, 22.12% (1896/8567) were aged 40-49 years, 18.66% (1599/8567) were aged 50-59 years, and 13.92% (1192/8567) were aged 60-69 years. Most respondents had either very high (30.61%, 2623/8567) or high (20.06%, 1719/8567) annual household incomes.

The peak reporting of having had sex with someone met on a website or a smartphone app was among sexually active men aged 16-29 years (4.78%, 100/2077). This declined with increasing age to 0.81% (10/1192) among sexually active men aged 60-69 years. Reporting having had sex with someone met on a website or a smartphone app was substantially higher among homosexual and bisexual men than among heterosexual men (36.23% [83/228] vs 1.86% [155/8339]). In terms of sexual practices, 1.43%(106/7468) of sexually active men with no STI test in the last year reported having had sex with someone met on a website or a smartphone app, compared with 10.92% (111/1017) of men with an STI test in the last year and 31.09% (20/66) of men who reported having had an STI diagnosis in the last year. In addition, 0.61% (45/7352) of sexually active men with 1 sexual partner in the last year reported having had sex with someone met on a website or a smartphone app, compared with 11.85% (91/772) of those with 2-3 sexual partners and 22.85% (101/444) of those with >3 sexual partners in the last year.

Table 3. Sociodemographic and other behavioral characteristics of having sex with someone met on a website or a smartphone app among sexually active men.

Characteristics[a,b]	%[c] (95% CI) in subgroup	%[c] (95% CI) outcome	Odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)[d]	*P* value
Age group (years)						
16-29	24.24 (22.01-26.61)	4.78 (3.54-6.45)	1.60 (0.95-2.68)	.08	0.43 (0.23-0.82)	.01
30-39	21.03 (19.05-23.2)	3.05 (2.06-4.55)	Referent		Referent	
40-49	22.12 (20.26-24.1)	2.39 (1.68-3.41)	0.78 (0.45 - 1.34)	.36	1.01 (0.51-2.01)	.97
50-59	18.66 (17.03-20.41)	1.76 (1.13-2.75)	0.57 (0.31-1.05)	.07	0.62 (0.28-1.40)	.25
60-69	13.96 (12.59-15.46)	0.81(0.46-1.43)	0.26 (0.13-0.52)	<.001	0.28 (0.12-0.65)	.003
Sexual identity						
Heterosexual	97.34 (96.86-97.75)	1.86 (1.48-2.35)	Referent		Referent	
Homosexual, bisexual or other	2.66 (2.25-3.14)	36.23 (28.83-44.34)	30.01 (19.88-45.31)	<.001	15.37 (8.34-28.35)	<.001
Language spoken at home						
English only	93.80 (92.41-94.95)	2.87 (2.44-3.50)	N/A[e]	N/A	N/A	N/A
Other	6.20 (5.05-7.59)	__[f]	N/A	N/A	N/A	N/A
Annual household income						
Very high	30.65 (28.46- 32.92)	30.61 (28.43-32.89)	Referent	N/A	N/A	N/A
High	20.06 (18.20 -22.05)	0.67 (0.34-1.32)	0.49 (0.22-10.9)	.08	N/A	N/A
Middle	16.19 (14.45-18.10)	1.79 (1.11-2.89)	1.28 (0.67-2.45)	.45	N/A	N/A
Very low or low	13.52 (12.02-15.17)	2.1(1.47-3.94)	1.82 (0.93-3.56)	.08	N/A	N/A
Area of residence[g]						
Urban	68.04 (65.86-70.15)	3.05 (2.46-3.76)	Referent		N/A	N/A
Regional or remote	30.34 (28.28-32.49)	2.21 (1.56-3.10)	0.72 (0.48-1.09)	.12	N/A	N/A
High alcohol consumption[h]						
No	N/A	2.65(2.18-3.23)	Referent		N/A	N/A
Yes	N/A	3.22(2.18-4.74)	1.22 (0.78-1.92)	.38	N/A	N/A
Injected drugs in last year						
No	N/A	2.71 (2.26-3.27)	Referent		N/A	N/A
Yes	N/A	5.31(2.59-10.57)	2.00 (0.93-4.33)	.08	N/A	N/A
Smoking status						
Never smoked or former smoker	N/A	2.38 (1.92-2.95)	Referent		N/A	N/A
Current smoker	N/A	4.37 (3.14-6.05)	1.87 (1.25-2.81)	.002	N/A	N/A
STI[i] testing in last year						
No test	N/A	1.43 (1.09-1.88)	Referent		Referent	
STI test	N/A	10.98 (8.42-14.21)	10.92 (8.37-14.12)	<.001	2.02 (1.12-3.38)	.008
STI diagnosis	N/A	31.09(17.57-48.86)	31.30 (14.07- 69.69)	<.001	3.15 (1.25-7.97)	.02
Condom use with most recent partner						
Used condoms	N/A	5.25(3.92-6.99)	3.26 (2.11-5.04)	<.001	N/A	N/A
Did not use	N/A	1.65 (1.23-2.21)	Referent	N/A	N/A	N/A
Number of sexual partners in last year						
1	N/A	0.61 (0.36-1.1)	Referent		Referent	
2-3	N/A	11.85 (9.12-15.24)	24.40 (13.82- 43.07)	<.001	17.86 (9.20-34.68)	<.001

Characteristics[a,b]	%[c] (95% CI) in subgroup	%[c] (95% CI) outcome	Odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)[d]	*P* value
>3	N/A	22.85 (18.3-28.14)	69.64 (36.75-131.97)	<.001	35.77 (18.04-70.94)	<.001

[a]n=8567 (2735), weighted (unweighted) denominators.

[b]Individuals with missing data are not shown; this was <5% for all variables, except for income and condom use with last sexual partner (not available for 42%).

[c]All data have been weighted to match the Australian population.

[d]Adjusted for age group, sexual identity, STI testing in the last year, and numbers of sexual partners in last year.

[e]N/A: not applicable.

[f]Too few responses for analysis (n<15).

[g]Accessibility/Remoteness Index of Australia (ARIA).

[h]≥28 standard drinks per week.

[i]STI: sexually transmissible infection.

In multivariate analyses, among sexually active men, several factors were associated with having had sex with someone met on a website or a smartphone app (Table 3): identifying as homosexual or bisexual compared with heterosexual (adjusted odds ratio, AOR: 15.37, 95% CI 8.34-28.35), having 2-3 sexual partners (AOR: 9.20, 95% CI 9.20-34.68) or >3 sexual partners in the last year (AOR: 35.77, 95% CI 18.04-70.94) compared with 1 partner, having STI in the previous year (AOR: 3.15, 95% CI 1.25-7.97), and having had an STI test in the previous year (AOR: 2.02, 95% CI 1.21-3.38) compared to not having an STI test.

Women

Among sexually active women, most were heterosexual (96.09%, 7838/8158), spoke English at home (96.25%,7852/8158), and lived in urban areas (67.78%, 5530/8158). Furthermore, 28.86% (2355/8158) were aged 16-29 years, 23.45% (1913/8158) were aged 30-39 years, 20.74% (1692/8158) were aged 40-49 years, 17.12% (1397/8158) were aged 50-59 years, and 9.83% (802/8158) were aged 60-69 years. Most sexually active women reported very high (23.16%, 1890/8158]) or high (22.36%, 1824/8158) annual household income compared with middle (17.68%, 1442/8158) or low and very low (16.73%, 1265/8158) annual household income.

Among sexually active women, those aged 16-29 years (2.38%, 55/2355) were most likely to report having had sex with someone met on a website or a smartphone app, and the least likely to report were those aged 60-69 years. Having had sex with someone met on a website or a smartphone app was higher among lesbian and bisexual women than among heterosexual women (5.51% [111/7838] vs 1.42% [18/319]). Women with low and very low annual household income (3.25%, 44/1365) more frequently reported having had sex with someone met using these technologies, compared with women with middle (0.58%, 8/1442), high (0.41% ,7/1824), and very high incomes (0.51%, 10/1890). Those with either an STI test (4.25%, 56/1330) or an STI diagnosis (9.99%, 13/130) in the last year were more likely to report having had sex with someone met online than those who had not had an STI test in the last year (0.89%, 59/6667). Women with 2-3 (13.37%, 59/438) or >3 (25.30%, 37/147) sexual partners in the last year were substantially more likely to report having had sex with someone met online than women with 1 sexual partner (0.44%,33/7574).

Among sexually active women, several factors were associated with having had sex with someone met online (Table 4): having 2-3 sexual partners in the last year (AOR: 32.01, 95% CI 13.17-77.78) or >3 sexual partners in the last year (AOR: 71:03, 95 % CI: 27.48-183.57), reporting a very low and low annual household income compared with very high annual household income (AOR: 3.40, 95% CI 1.12-10.35), and identifying as lesbian or bisexual (AOR: 2.27, 95% CI 1.04-4.49).

Table 4. Sociodemographic and behavioral correlates of having sex with someone met on a website or a smartphone app among sexually active women.

Characteristics[a,b]	%[c] (95% CI) in subgroup	%[c] (95% CI) outcome	Odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)[d]	*P* value
Age group (years)						
16-29	28.86 (26.36-31.51)	2.38 (1.62-3.38)	2.34 (1.62-3.38)	.29	0.43 (0.21 -0.86)	.02
30-39	23.45 (21.39-25.63)	1.69 (1.04-2.72)	Referent	N/A[e]	Referent	N/A
40-49	20.74 (18.91-22.70)	1.28 (0.68-2.38)	0.75 (0.34 - 1.68)	.49	1.16 (0.45 - 3.01)	.75
50-59	17.12 (15.55-18.82)	1.07 (0.38-2.98)	0.63 (0.20-1.99)	.43	1.22 (0.36 -4.18)	.75
60-69	9.83 (8.65-11.14)	0.60 (0.13-2.77)	0.35 (0.07- 1.79)	.21	0.66 (0.11 -3.98)	.65
Sexual identity						
Heterosexual	96.09 (95.35-96.72)	1.42 (1.08 - 1.91)	Referent	N/A	Referent	N/A
Lesbian, bisexual, or other	3.91 (3.28-4.65)	5.51 (3.12- 9.54)	4.05 (2.09 - 7.84)	<.001	2.27 (1.04-4.94)	.04
Language spoken at home						
English only	96.25 (2.78-5.04)	0.84 (0.20 - 3.88)	N/A	N/A	N/A	N/A
Other	3.75 (94.96-97.22)	__[f]	N/A	N/A	N/A	N/A
Annual household income						
Very high	23.16 (21.15-25.31)	0.51 (0.25-1.06)	Referent	N/A	Referent	N/A
High	22.36 (20.33-24.54)	0.41 (0.17-1.00)	0.79 (0.25 - 2.53)	.69	0.86 (0.28 - 2.67)	.80
Middle	17.68 (15.83-19.69)	0.58(0.25-1.31)	1.12 (0.37 - 3.39)	.84	1.00 (0.31 - 3.23)	.99
Very low or low	16.73 (14.95-18.68)	3.25 (2.05)	6.49 (2.71-15.55)	<.001	3.40 (1.12 - 10.35)	.03
Area of residence[g]						
Urban	67.78 (65.44-70.03)	1.74 (1.26 - 2.40)	Referent	N/A	N/A	N/A
Regional or remote	30.54 (28.34-32.82)	1.19 (0.79 - 1.78)	0.68 (0.40 - 1.14)	.12	N/A	N/A
High alcohol consumption[h]						
No	N/A	1.44 (1.02 - 2.05)	Referent	N/A	N/A	N/A
Yes	N/A	1.82 (1.24 - 2.67)	1.27 (0.75 - 2.15)	.38	N/A	N/A
Injected drugs in last year						
No	N/A	1.57 (1.22 - 2.07)	Referent	N/A	N/A	N/A
Yes	N/A	2.09 (0.64 - 6.76)	1.34 (0.39 - 4.62)	.65	N/A	N/A
Smoking status						
Never smoked or former smoker	N/A	1.56 (1.18 - 2.12)	Referent	N/A	N/A	N/A
Current smoker	N/A	1.69(1.03 - 2.81)	1.08 (0.60 - 1.97)	.81	N/A	N/A
STI[i] testing in last year						
No test	N/A	0.89 (0.59 - 1.36)	Referent	N/A	N/A	N/A
STI test	N/A	4.25 (2.97 - 6.16)	4.94 (2.79 - 8.77)	<.001	N/A	N/A
STI diagnosis	N/A	9.99 (4.93 - 19.20)	12.38 (5.17 - 29.62)	<.001	N/A	N/A
Condom use with most recent partner						
Used condoms	N/A	3.41 (2.24 -5.15)	1.06 (0.57-1.96)	.86	N/A	N/A
Did not use	N/A	3.60 (2.34 -5.48)	Referent	N/A	N/A	N/A
Number of sexual partners in last year						
1	N/A	0.44 (0.22-0.88)	Referent	N/A	Referent	N/A
2-3	N/A	13.37 (9.93-17.77)	35.13 (16.12-76.59)	<.001	32.01 (13.17-77.78)	<.001

Characteristics[a,b]	%[c] (95% CI) in subgroup	%[c] (95% CI) outcome	Odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)[d]	*P* value
>3	N/A	25.30 (17.20 -35.58)	77.08 (32.75-181.43)	<.001	71.03 (27.48 - 183.57)	<.001

[a]n=8158 (2580), weighted (unweighted) denominators

[b]Individuals with missing data are not shown; this was <5% for all variables, except for income and condom use with most recent partner (not available for 42%).

[c]All data have been weighted to match the Australian population.

[d]Adjusted for age group, income, and numbers of sexual partners in the last year.

[e]N/A: not applicable.

[f]Too few responses for analysis (n<15).

[g]Accessibility/Remoteness Index of Australia.

[h]≥14 standard drinks per week (for women).

[i]STI: sexually transmissible infection.

Discussion

Principal Findings

Overall, our findings indicate that in 2012-2013, approximately 1 in 10 Australian adults aged from 16 to 69 years had ever looked for potential partners using websites or smartphone apps, of whom approximately half had done so in the in the last year. Among people who searched in the last year, over half had physically met with someone, and approximately two-thirds of these people had had sex with someone they met online in the last year, equating to 1.95% of the population.

These nationally representative estimates of looking for and having had sex with someone met on website and smartphone apps are lower than those of surveys focusing on specific subpopulations and using convenience-sampling frames, which have reported 6%-40% of the population meeting sexual partners using websites [2,5,11,12] and 18%-76% using dating applications and websites [3,13-18]. The difference is almost certainly related to different populations sampled and may also be related to the fact that ASHR2 asked specifically about having sex with someone, rather than only looking for partners or meeting in person. Age was strongly correlated with having had sex with someone met on a website or a smartphone app, especially since younger people tend to have higher levels of mobile phone and Internet access, but this study included a broader range of ages than many others. Just over a third (36.07%) of people who used internet dating and hookup applications in the last year reported having had sex with someone they met on a website or using a smartphone app in the last year. Certain populations were less likely to report having had sex with someone met using these technologies. For example, 30.89% of heterosexual respondents who reported that they had used the Internet or a mobile phone app to search for a potential partner in the last year reported having had sex with someone they met online, compared with 62.41% of gay, lesbian, and bisexual survey respondents (see Table 2). This suggests that other studies reporting on the use of internet dating and hookup apps as a proximal marker of having had sex with someone met online are likely to overestimate the practice's prevalence among lower-risk segments of the population.

In this survey, 15.58% of gay, lesbian, and bisexual respondents reported meeting a sexual partner in the last year. Higher uptake of finding partners using websites among nonheterosexual respondents was also observed in a British population survey [4]. Yet, in our study, after adjusting for age and other characteristics, the strongest correlate of having had sex with someone met using a website or smartphone app was higher numbers of sexual partners in the last year, suggesting that these technologies are favored by the most sexually active respondents. This finding is consistent with other studies reporting that people who look for partners with these tools have increased sexual activity compared with nonusers, including younger age at first sex [2,5] and higher numbers of sexual partners [2,4,5,12,15,16]. Due to the cross-sectional nature of the survey, determining causality is not possible, so findings could mean platforms provide an efficient means for more sexually active individuals to connect with new partners, or alternatively, people who were already more sexually active were attracted to these sites and other ways to meet sexual partners.

In general, people who met partners using websites and apps and had sex with them were more likely to engage in higher-risk practices than those who did not—except for condom use at the last sexual event, which was higher among people who met partners online. There was attenuation of the condom use variable in the multivariate analysis, meaning that the association was not significant in the adjusted analysis after controlling for the numbers of sexual partners and other demographic factors. Higher levels of condom use at the last event could reflect condom use with newer and less established partners, with whom STI prevention is prioritized. This explanation seems highly plausible because online tools are often used to facilitate new sexual partnerships and those who report using websites and apps to find sexual partners also report higher numbers of recent sexual partners. However, the finding contrasts with many other studies that tend to find meeting or seeking partners online is linked to condomless sexual intercourse [4,5,14,15].

In relation to STI history, both STI testing and diagnoses were higher among people who reported meeting partners using websites and smartphone apps and having sex with them. Again, adjusted analysis showed attenuation of STI history for women after adjusting for the number of sexual partners. This implies

that women who reported either an STI test or diagnosis in the last year were also more likely to have multiple sexual partners in the last year. The relationship between STI history and meeting partners using websites and apps remained significant for men, even after adjusting for partner numbers and other demographic factors. Interestingly, this pattern has also been seen in the other two population studies with a significant relationship between STIs and searching for or meeting partners on websites and smartphone apps among men, but not among women [4,5]. Findings from other studies have been mixed as to whether STI diagnoses are related to either finding or searching for sexual partners on internet sites and geosocial apps [1-4,16,19]. A number of studies have reported on HIV testing history among gay men and tend to find that men who use apps are more likely to have been tested for HIV [14,19]. Studies with gay men in Australia suggest that men who use a combination of mobile phone apps, internet websites, and offline places to meet partners appear to be at increased risk of STIs or HIV compared with men who use a narrower range of online and offline methods [20].

In women, low annual household income was associated with meeting partners on websites and apps. Socioeconomic deprivation has been linked with poor health outcomes, including STI acquisition [21-23], and other reports from ASHR2 found lower income related to multiple sexual partners [24]. Aside from age and income level for women, other sociodemographic factors, smoking, high alcohol consumption, and injecting drug use were not associated with having sex with someone met online.

Limitations

To our knowledge, this is the first study to report the prevalence of having had sex with someone met on a website or a mobile phone app in the past year from a representative survey of the general adult population. One of this study's strengths is the capacity to assess the proportion of people who used internet dating and apps, met in person, and had sex with someone met online within the same population. Nonetheless, our study also has several limitations to consider when interpreting findings. First, the study was conducted in 2012-2013, and since that time, the technology landscape and behaviors related to the uptake of technology have changed. Very likely, the uptake of dating and hookup apps has substantially increased since the survey was conducted. Second, the sample of homosexual and bisexual men was not sufficiently large to enable analyses focused on this group. Third, all outcomes including STI outcomes were based on self-report, which is susceptible to recall and other reporting biases. Several similar studies have used biological measurement to ascertain STI prevalence, a more robust measure [2,4,5]. However, when asked, 88.55% of participants reported answering all survey questions truthfully, and a further 9.89% reported that they had answered 90%-99% of the survey honestly. This equates to 98.44% of participants answering at least 90% of questions truthfully [7]. Furthermore, the clear majority (89.97%) of participants reported that they were either not embarrassed or only slightly embarrassed by the questions. Evidence of the relatively low embarrassment and discomfort with questions related to sexual practices is seen in the number of people declining to answer particular questions. The question with the highest rate of refusal involved annual income compared with much lower refusal rates for questions about sexual practices .

Implications

Understanding the number and characteristics of people most likely to use these technologies to meet new sexual partners assists organizations responsible for HIV and STI prevention programs in identifying places and populations wherein they can focus their health promotion and testing initiatives. Our study has also demonstrated that although the prevalence of having had sex with someone met on a website or a smartphone app was 3.03% in people aged 20-29 years, it remained at 2.25% in people aged 30-49 years, suggesting the need for promotional material to cover a broad range of ages, not just younger adults. Although STIs are most prevalent in people aged 16-29 years, recent studies have suggested an increased rate in people over 30 years [25]. Furthermore, despite the finding that people who met partners on dating websites and apps were more likely to have had an STI test in the last year, a substantial proportion of respondents who reported having sex with someone they met online (likely to be a new sexual partner) reported not having an STI test in the past year, suggesting an opportunity to raise awareness about STI testing further and to use targeted advertisements to direct people to easy access points such as new websites where pathology request forms can be downloaded without attending a clinic [26]. Some have also suggested that these platforms have the capacity to enable partner notification and data collection in relation to sexual health [27]. Notably, however, the owners of dating websites and apps have historically been concerned about associating their platforms with STIs and therefore are reluctant to promote public health initiatives [27].

Conclusions

Internet and smartphone technologies are a relatively common way of meeting new sexual partners among highly sexually active survey respondents, homosexual and bisexual men, and younger adults, suggesting that the use of in-app health promotion is a feasible approach to targeting these populations [28,29]. Future research could explore the potential of health promotion in dating websites and geosocial applications. The use of smartphone technologies to search for potential sexual partners may become a normative dating practice among Australian adults in their twenties and thirties, so repeat surveys would be important to document this prevalence over time.

Acknowledgments

The Australian Study of Health and Relationships was funded by the National Health and Medical Research Council (project grant number 1002174).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Prevalence of looking for potential partners on websites or smartphone apps and prevalence of having sex with these partners among males.

[PDF File (Adobe PDF File), 38KB - jmir_v20i12e10683_app1.pdf]

Multimedia Appendix 2

Prevalence of looking for potential partners on websites or smartphone apps and prevalence of having sex with these partners among females.

[PDF File (Adobe PDF File), 38KB - jmir_v20i12e10683_app2.pdf]

References

1. Beymer MR, Weiss RE, Bolan RK, Rudy ET, Bourque LB, Rodriguez JP, et al. Sex on demand: geosocial networking phone apps and risk of sexually transmitted infections among a cross-sectional sample of men who have sex with men in Los Angeles County. Sex Transm Infect 2014 Nov;90(7):567-572 [FREE Full text] [doi: 10.1136/sextrans-2013-051494] [Medline: 24926041]
2. Buhi ER, Klinkenberger N, McFarlane M, Kachur R, Daley EM, Baldwin J, et al. Evaluating the Internet as a sexually transmitted disease risk environment for teens: findings from the communication, health, and teens study. Sex Transm Dis 2013 Jul;40(7):528-533. [doi: 10.1097/OLQ.0b013e31829413f7] [Medline: 23965765]
3. Owen S, Kurka T, Richardson D. UG2 Grindr© use by men who have sex with men (MSM) is associated with high rates of bacterial sexually transmitted infections. Sex Transm Infect 2016 Jun 30;92(Suppl 1):A18.3-A1A19. [doi: 10.1136/sextrans-2016-052718.51]
4. Cabecinha M, Mercer CH, Gravningen K, Aicken C, Jones KG, Tanton C, et al. Finding sexual partners online: prevalence and associations with sexual behaviour, STI diagnoses and other sexual health outcomes in the British population. Sex Transm Infect 2017 Dec;93(8):572-582 [FREE Full text] [doi: 10.1136/sextrans-2016-052994] [Medline: 28396557]
5. Gravningen K, Aicken CR, Schirmer H, Mercer CH. Meeting sexual partners online: associated sexual behaviour and prevalent chlamydia infection among adolescents in Norway: a cross-sectional study. Sex Transm Infect 2016 Mar;92(2):97-103. [doi: 10.1136/sextrans-2015-052152] [Medline: 26268510]
6. IAC. Tinder overview URL: http://iac.com/brand/tinder [accessed 2018-04-05] [WebCite Cache ID 6yRQUrteF]
7. Richters J, Badcock PB, Simpson JM, Shellard D, Rissel C, de Visser RO, et al. Design and methods of the Second Australian Study of Health and Relationships. Sex Health 2014 Nov;11(5):383-396. [doi: 10.1071/SH14115] [Medline: 25376992]
8. Smith AMA, Rissel CE, Richters J, Grulich AE, de Visser RO. Sex in Australia: the rationale and methods of the Australian Study of Health and Relationships. Aust N Z J Public Health 2003;27(2):106-117. [Medline: 14696700]
9. Department of Health and Aged Care. Measuring remoteness: accessibility/remoteness index of Australia (ARIA). Canberra: Commonwealth of Australia; 2001. URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/E2EE19FE831F26BFCA257BF0001F3DFA/$File/ocpanew14.pdf [accessed 2018-10-10] [WebCite Cache ID 733AvVZTD]
10. Grulich AE, de Visser RO, Badcock PB, Smith AMA, Richters J, Rissel C, et al. Knowledge about and experience of sexually transmissible infections in a representative sample of adults: the Second Australian Study of Health and Relationships. Sex Health 2014 Nov;11(5):481-494. [doi: 10.1071/SH14121] [Medline: 25377001]
11. Whiteley LB, Brown LK, Swenson RR, Valois RF, Vanable PA, Carey MP, et al. African American adolescents meeting sex partners online: closing the digital research divide in STI/HIV prevention. J Prim Prev 2012 Feb;33(1):13-18 [FREE Full text] [doi: 10.1007/s10935-012-0262-3] [Medline: 22293979]
12. Ogilvie GS, Taylor DL, Trussler T, Marchand R, Gilbert M, Moniruzzaman A, et al. Seeking sexual partners on the internet: a marker for risky sexual behaviour in men who have sex with men. Can J Public Health 2008;99(3):185-188. [Medline: 18615938]
13. Almeida M, Gibbs J, Estcourt C. Are geosocial networking (GSN) apps associated with increased risk of STIs & HIV: A systematic review. In: Sexually Transmitted Infections.: Sexually Transmitted Infections; 2016 Presented at: BASHH Conference; July 10–12 2016; Oxford.
14. Card KG, Lachowsky NJ, Cui Z, Shurgold S, Gislason M, Forrest JI, et al. Exploring the role of sex-seeking apps and websites in the social and sexual lives of gay, bisexual and other men who have sex with men: a cross-sectional study. Sex Health 2017 Jun;14(3):229-237. [doi: 10.1071/SH16150] [Medline: 27977387]
15. Choi EP, Wong JY, Lo HH, Wong W, Chio JH, Fong DY. The Impacts of Using Smartphone Dating Applications on Sexual Risk Behaviours in College Students in Hong Kong. PLoS One 2016 Nov;11(11):e0165394 [FREE Full text] [doi: 10.1371/journal.pone.0165394] [Medline: 27828997]

16. Douglass C, Tomnay J, Lim M. editors. In: Are geosocial networking (GSN) apps associated with increased risk of STIs & HIV: A systematic review. 2016 Presented at: Australasian Sexual Health Conference; 16-18 November 2016; Adelaide.
17. Goedel WC, Duncan DT. Geosocial-Networking App Usage Patterns of Gay, Bisexual, and Other Men Who Have Sex With Men: Survey Among Users of Grindr, A Mobile Dating App. JMIR Public Health Surveill 2015;1(1):e4 [FREE Full text] [doi: 10.2196/publichealth.4353] [Medline: 27227127]
18. Shilo G, Mor Z. Seeking sex online: social and sexual risk factors among adolescent and young gay and bisexual men. Sex Health 2015 Jun;12(3):217-223. [doi: 10.1071/SH14229] [Medline: 25751619]
19. Zou H, Fan S. Characteristics of Men Who Have Sex With Men Who Use Smartphone Geosocial Networking Applications and Implications for HIV Interventions: A Systematic Review and Meta-Analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]
20. Hull P, Mao L, Prestage G, Zablotska I, de Wit J, Holt M. The use of mobile phone apps by Australian gay and bisexual men to meet sex partners: an analysis of sex-seeking repertoires and risks for HIV and STIs using behavioural surveillance data. Sex Transm Infect 2016 Nov 19;92(7):502-507. [doi: 10.1136/sextrans-2015-052325] [Medline: 27095378]
21. Sales JM, Smearman EL, Swartzendruber A, Brown JL, Brody G, DiClemente RJ. Socioeconomic-related risk and sexually transmitted infection among African-American adolescent females. J Adolesc Health 2014 Nov;55(5):698-704 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.05.005] [Medline: 24974317]
22. Andreatos N, Grigoras C, Shehadeh F, Pliakos EE, Stoukides G, Port J, et al. The impact of HIV infection and socioeconomic factors on the incidence of gonorrhea: A county-level, US-wide analysis. PLoS One 2017;12(9):e0183938 [FREE Full text] [doi: 10.1371/journal.pone.0183938] [Medline: 28863154]
23. Furegato M, Chen Y, Mohammed H, Mercer CH, Savage EJ, Hughes G. Examining the role of socioeconomic deprivation in ethnic differences in sexually transmitted infection diagnosis rates in England: evidence from surveillance data. Epidemiol Infect 2016 Nov;144(15):3253-3262. [doi: 10.1017/S0950268816001679] [Medline: 27511704]
24. Rissel C, Badcock PB, Smith AMA, Richters J, de Visser RO, Grulich AE, et al. Heterosexual experience and recent heterosexual encounters among Australian adults: the Second Australian Study of Health and Relationships. Sex Health 2014 Nov;11(5):416-426. [doi: 10.1071/SH14105] [Medline: 25376995]
25. The Kirby Institute. HIV, Viral Hepatitis and Sexually Transmissible Infections in Australia Annual Surveillance Report. Sydney: The Kirby Institute, UNSW Sydney; 2016.
26. Gaydos CA, Jett-Goheen M, Barnes M, Dize L, Hsieh Y. Self-testing for Trichomonas vaginalis at home using a point-of-care test by women who request kits via the Internet. Sex Health 2016 Aug;13(5):491-493 [FREE Full text] [doi: 10.1071/SH16049] [Medline: 27491592]
27. Belluz J. Tinder and Grindr don't want to talk about their role in rising STDs. 2017 URL: https://www.vox.com/science-and-health/2017/11/13/16620286/online-dating-stds-tinder-grindr [accessed 2018-04-05] [WebCite Cache ID 6yRQzadA9]
28. Ubrihien A, Davies S, Hadlow B. Use of location based mobile application Grindr to increase MSM testing rates in Northern Sydney, Australia. 2014 Presented at: Australasian Sexual Health Conference Sydney; 9-11 October 2014; Sydney.
29. Rosengren AL, Huang E, Daniels J, Young SD, Marlin RW, Klausner JD. Feasibility of using Grindr(TM) to distribute HIV self-test kits to men who have sex with men in Los Angeles, California. Sex Health 2016 May;13(4):389-392. [doi: 10.1071/SH15236] [Medline: 27209076]

Abbreviations

ASHR2: Second Australian Study of Health and Relationships
AOR: adjusted odds ratio
STI: sexually transmissible infection

Edited by G Eysenbach; submitted 04.04.18; peer-reviewed by H Fontenot, Z Ma; comments to author 23.05.18; revised version received 05.06.18; accepted 16.06.18; published 18.12.18.

Please cite as:
Watchirs Smith L, Guy R, Degenhardt L, Yeung A, Rissel C, Richters J, Liu B
Meeting Sexual Partners Through Internet Sites and Smartphone Apps in Australia: National Representative Study
J Med Internet Res 2018;20(12):e10683
URL: https://www.jmir.org/2018/12/e10683/
doi:10.2196/10683
PMID:30563809

Original Paper

The Effect of Using Geosocial Networking Apps on the HIV Incidence Rate Among Men Who Have Sex With Men: Eighteen-Month Prospective Cohort Study in Shenyang, China

Junjie Xu[1], PhD; Huan Yu[1], MPH; Weiming Tang[2], MD, PhD; Sequoia I Leuba[3], MD, MPH; Jing Zhang[1], PhD; Xiang Mao[1], MPH; Hongyi Wang[1], MPH; Wenqing Geng[1], PhD; Yongjun Jiang[1], PhD; Hong Shang[1], PhD

[1]Key Laboratory of AIDS Immunology of National Health and Family Planning Commission, Department of Laboratory Medicine, The First Affiliated Hospital, China Medical University, Shenyang, China

[2]Dermatology Hospital, Southern Medical University, Guangzhou, China

[3]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Hong Shang, PhD
Key Laboratory of AIDS Immunology of National Health and Family Planning Commission
Department of Laboratory Medicine
The First Affiliated Hospital, China Medical University
155 Nanjing North Street
Heping District
Shenyang,
China
Phone: 86 8328 2634
Fax: 86 8328 2634
Email: hongshang100@hotmail.com

Abstract

Background: Men who have sex with men (MSM) frequently seek partners through mobile apps (geosocial networking [GSN] apps). However, it is unclear whether GSN apps' use is associated with the increase in HIV incidence among MSM.

Objective: The aim of this study was to clarify the characteristics of GSN apps' users and to determine the association and putative mechanisms between GSN apps' use behavior and HIV incidence.

Methods: We conducted an 18-month prospective cohort study of MSM in Shenyang, China, and the participants were surveyed every 3 months from March 2015 to December 2016. An in-person interview collected information on sociodemographics, GSN apps' use, recreational drug use, and sexual behaviors. In addition, blood was drawn to test for HIV and syphilis. We used a multivariable Cox regression model to determine possible predictors for increased HIV incidence.

Results: Of the enrolled 686 HIV-negative MSM, 431 (431/686, 62.8%) were GSN apps' users. Compared with GSN apps' nonusers, GSN apps' users were younger; had an earlier age of sexual debut; and in the past 3 months, were more likely to have used recreational drugs, more likely to have had 5 or more casual partners (CPs), more likely to have had group sex with males, and more likely to have had condomless anal intercourse (CAI) with male steady partners (SPs). In addition, 59.4% (256/431) of the GSN apps' users were willing to accept HIV and AIDS prevention information push services through GSN apps. In total, 19 MSM seroconverted to HIV during the follow-up period; the HIV incidence density rate was 8.5 (95% CI 5.0-13.5) per 100 person-years (PY) among GSN apps' users and 2.0 (95% CI 0.4-5.8) per 100 PY among nonusers. New HIV infections were independently associated with ever using GSN apps to seek male sexual partners (P=.04) and in the past 3 months, using recreational drugs (P=.048), having group sex with males (P=.01), and having CAI with male CPs (P=.02).

Conclusions: GSN apps' use is associated with higher HIV incidence and may be mediated through recreational drug use and having multiple CPs. Researchers must develop an intervention propagated through GSN apps to reach this high-risk population to mitigate the HIV epidemic in the MSM community.

(J Med Internet Res 2018;20(12):e11303) doi:10.2196/11303

KEYWORDS

geosocial networking apps; dating apps; HIV; incidence; homosexuality; male; cohort study

Introduction

Background

Men who have sex with men (MSM) are disproportionately affected by HIV [1-3]. The percentage of MSM with HIV and AIDS in China increased from 13.7% in 2011 to 28.3% in 2015 [4,5]. Understanding the driving forces of this epidemic is essential to provide tailored interventions to MSM.

In the past decade, how MSM seek sexual partners has changed dramatically. In the 1990s, MSM mainly socialized and sought sexual partners in public facilities such as gay bars and bathrooms. Today, with the current popularity of smartphones, tablets, or computers with geosocial networking (GSN) abilities, a large number of GSN apps have been developed and are widely used by MSM. These GSN apps have revolutionized social communication and how MSM seek casual partners (CPs) or multiple sexual partners [6-10]. On average, 36.0% to 63.6% of MSM in the United States [7,11,12] and 40.6% of MSM in mainland China [6] have sought male sexual partners by using GSN apps. Previous research has shown that people who use GSN apps have more sexual partners and more frequent casual (ie, a quick, unplanned encounter without inquiring about the partner's HIV serostatus) sexual intercourse compared with people who do not, leading to concerns about GSN apps' use affecting increases in HIV transmission around the world [12-16].

However, the prevalence of HIV and other sexually transmitted infections (STIs) among GSN apps' users and nonusers is inconsistent with expectations based on the above behavioral differences. A recent meta-analysis concluded that GSN apps' users compared with nonusers had significantly higher prevalence of STIs, including gonorrhea and chlamydia, but had lower HIV prevalence [16]. Possible reasons for these differences in the prevalence of specific STIs among GSN apps' users and among nonusers may be related to the designs of these studies. First, in most of these studies, the HIV/STI infection history was developed through the respondents' self-report [14,17,18], which is subject to recall bias and social desirability bias. Second, most studies were cross-sectional surveys and, thus, only examined the relationship between GSN apps' use and prevalent HIV infection among MSM. Furthermore, the cross-sectional study design can neither define the time duration of current GSN apps' use among GSN apps' users nor can it clarify whether these GSN apps are linked to HIV-related behavioral changes or HIV incidence over time. Finally, compared with using recent HIV infection, using overall HIV infection (includes both recent and established cases) as the outcome of interest does not accurately reflect the effects of GSN apps' use among MSM. Previous studies have found that among MSM, GSN apps' users were younger compared with GSN apps' nonusers, and as this younger population has a shorter duration of potential exposure to HIV, they are expected to have lower HIV prevalence [6,11,14,19,20]. Although 1 recent study reported finding casual sex partners on the internet was an independent risk factor for incidence of HIV infection among MSM in Bangkok, Thailand [21], finding partners through the GSN apps is a more innovative way for MSM to seek sexual partners [6]. No publications have examined the association between using GSN apps and HIV incidence rates among MSM so far. Longitudinal studies of MSM are needed to clarify the HIV incidence among GSN apps' users over time and to compare the HIV incidence rate among GSN apps' users with the incidence rate among GSN apps' nonusers to determine if using GSN apps is contributing to the HIV epidemic among MSM [16].

Objectives

We conducted an 18-month prospective cohort study among MSM in Shenyang, China, to clarify the characteristics of GSN apps' users, the association and putative mechanisms between GSN apps' use behavior and HIV incidence, and their willingness to accept an HIV prevention information dissemination service through a GSN app platform.

Methods

Recruitment

Between March 2015 and December 2016, MSM participants in Shenyang, Liaoning province, were recruited through a mixed recruitment method of internet sampling, venue-based sampling, or chain-referral sampling [22]. The inclusion criteria of this cohort were as follows: (1) being 18 years or older; (2) born male; (3) had anal and/or oral intercourse with male partners in the past 6 months; (4) tested as serologically negative for both HIV antibodies and HIV nucleic acid amplification testing (NAAT); and (5) willing and able to provide a written informed consent.

The survey was conducted at The First Affiliated Hospital of China Medical University in Shenyang, China.

Follow-Up of the Prospective Men Who Have Sex With Men Cohort

All eligible participants were prospectively followed-up at a 3-month frequency. After the initial eligibility interview screening, trained staff interviewed in-person eligible participants in a private counseling room and assigned each participant a unique 6-number identification code to be linked to their laboratory testing results. Venous blood specimens were then drawn and tested for HIV and syphilis. All participants who tested positive for HIV or syphilis received posttest counseling for the infection and referrals to relevant clinics. Each participant received 50 RMB (US $7.4), free condoms, and 1 free lubricant after each completed study visit. Each participant was asked to provide at least two different current methods of contact, and reminder phone calls were made before each follow-up visit.

Data Collection and Related Measures

Baseline and each follow-up questionnaires repeatedly asked for the following information: (1) sociodemographics, including

age, marital status, ethnicity, education, and monthly income; (2) sexual practices in the past 3 months, including number of male sexual partners, how many were steady partners (SPs, sexual activity that takes place between partners in a romantic relationship and usually implies commitment, emotional attachment, or familiarity between sexual partners), how many were casual partners (CPs, affairs like one-night stands or casual sex between males who have little or no knowledge of each other), condom use, and group sex with males (sexual behavior involving more than 2 male participants); (3) recreational drug use in the past 3 months, including using the following types of drugs: poppers (alkyl nitrites), ecstasy, ice (methamphetamine), amphetamine, tramadol, and ketamine; and (4) the names, numbers, and current length of use of GSN apps to seek male sexual partners.

We defined the main outcome of incident HIV infection as seroconversion determined by the presence of HIV antibody during a visit after a previous visit with a laboratory-confirmed HIV-negative serostatus.

We calculated the sample size of MSM participants based on a Cox regression of the log hazard ratio analysis model [23]. When GSN-app-use behavior (X1) is estimated to be .50, the estimated HIV incidence rate (outcome event) is 0.07. We used the parameters of 85% power at a two-sided .05 significance level, an assumed hazard ratio of 3, an SD of X1=0.5, and an R^2 (*R*-squared of X1 with other Xs)= 0.18, and calculated that the smallest sample size was 432 observations. We used PASS (Power Analysis & Sample Size) software version 11 (NCSS, Kaysville, UT, USA) to calculate the sample size.

Laboratory Testing

After obtaining informed consent at the baseline survey and each follow-up time point, we drew 10 mL of venous blood from each participant to test for HIV and syphilis. HIV-1 antibody screening was performed by enzyme-linked immunosorbent assay, and positive cases were further confirmed through a Western blot test. Specimens that had negative or indeterminate HIV antibody results were further tested using RNA with pooled NAAT (COBAS AmpliPrep, COBAS TaqMan HIV-1 Test, Roche, Germany). Syphilis serology was performed with the rapid plasma reagin (RPR) test (Shanghai Kehua, China), and positive cases were further confirmed by the *Treponema pallidum particle agglutination* assay (TPPA, Serodia, Japan). Participants with plasma positive for both RPR and TPPA were concluded to be currently infected with syphilis.

All related biological tests were conducted in the Key Laboratory of AIDS Immunology of National Health and Family Planning Commission of The First Affiliated Hospital of China Medical University.

Statistical Analysis

Data were entered twice and checked for accuracy using EpiData Entry software. All data analyses were performed using IBM SPSS (International Business Machines Corporation Statistical Product and Service Solutions) 20.0. MSM who self-reported ever using GSN apps to seek male sexual partners were defined as GSN apps' users, and the men who self-reported never using GSN apps to seek male sexual partners were considered GSN apps' nonusers. Comparisons between GSN apps' users and nonusers and between MSM retained for at least one 3-month follow-up and MSM who withdrew from the cohort were analyzed by chi-square tests. The time of HIV seroconversion was defined as the middle time point between the last laboratory-confirmed HIV seronegative date and the first laboratory-confirmed HIV seropositive date. We measured the follow-up in person-years (PY), and the follow-up spanned from the date of enrollment to either the date of HIV seroconversion or the date of the last follow-up session. We used *a mixed* Cox proportional hazards model to assess cumulative hazard ratios (HRs), both crude (cHR) and adjusted (aHR), for high-risk factors for HIV infection to determine their effects on HIV incidence rates. Time-dependent covariates for the Cox proportional hazards model included over the past 3 months, condom use with male SPs, condom use with male CPs, group sex with males, number of CPs, recreational drug use, and use of GSN apps to seek male sexual partners. The models were adjusted for age, level of education, registered residence, ethnicity, marital status, and monthly income. A two-sided *P* value of less than .05 was considered statistically significant.

Ethics Statement

This study protocol was reviewed and approved by the Institutional Review Board of the First Affiliated Hospital of China Medical University, with ethical review number of 2011-36. The study protocol, contents, and procedure were explained to each participant before the survey. Written informed consent was obtained from all participants before the interview and blood collection. The procedures in the study were performed in accordance with the study protocol and relevant regulations.

Results

Sociodemographic Characteristics of the Participants

A total of 761 MSM who had no prior positive HIV test were screened for HIV, of which 9.1% (69/761) individuals were detected as HIV-positive and excluded from this study. The HIV-negative MSM were invited to participate in an 18-month prospective cohort study, of which 0.9% (6/692) HIV-negative MSM declined to participate. Thus, a total of 686 eligible HIV-negative MSM were included in this prospective cohort study, of which 431 (431/686, 62.8%) self-identified as GSN apps' users and 255 (255/686, 37.2%) as nonusers (Figure 1). Table 1 summarizes the sociodemographics, sexual behaviors, recreational drug use, and HIV testing behaviors of the baseline eligible HIV-negative GSN apps' users and nonusers. Most GSN apps' users were older than 24 years (294/431, 68.2%), were originally not from Shenyang city (253/431, 58.7%), had college-level education or above (236/431, 54.8%), were ≤20 years at their sexual debut with males (234/431, 54.3%), had male SPs in the past 3 months (250/431, 58.0%), had male CPs in the past 3 months (224/431, 52.0%), and were willing to receive HIV and AIDS prevention information through a push service conducted through a GSN app platform (256/431, 59.4%), but not many GSN apps' users had been tested for HIV before (113/431, 26.2%). In addition, in the past 3 months, 85.8% (370/431) of GSN apps' users sought male sexual

partners through at least one GSN app, 32.7% (141/431) GSN apps' users used recreational drugs, 10.9% (47/431) had at least five male CPs, and 8.1% (35/431) had group sex with males. At baseline, 10.4% (45/431) of GSN apps' users were infected with syphilis.

GSN apps' users were younger (P<.001), had higher proportion of college or above education level (P<.001), had higher HIV testing rates (P=.02), had higher proportion of recreational drug use (P<.001), had higher proportion of group sex with males (P=.01), and had higher proportions of having 5 or more CPs in the past 3 months (P=.01) compared with its counterpart group.

In total, GSN apps' users reported using mainly 7 types of GSN apps, in which Blued was the most popular GSN app used to seek male partners (403/431, 93.5%), followed by Zank (91/431, 21.1%), WeChat (55/431, 12.8%), Jack'd (51/431, 11.8%), Tencent QQ (10/431, 2.3%), Momo (8/431, 1.9%), and Gpark (4/431, 0.9%). In addition, 3.7% (16/431) of GSN apps' users used other kinds of GSN apps (Figure 2). In addition, 35.0% (151/431) of GSN apps' users had once used at least two GSN apps to seek partners, and the median duration for which GSN apps were used was 12 (interquartile range: 4-30) months.

Figure 1. Flowchart of the prospective cohort study examining the relationship between using geosocial networking apps and HIV incidence among men who have sex with men population. MSM: men who have sex with men, GSN: geosocial networking, PY: person-years.

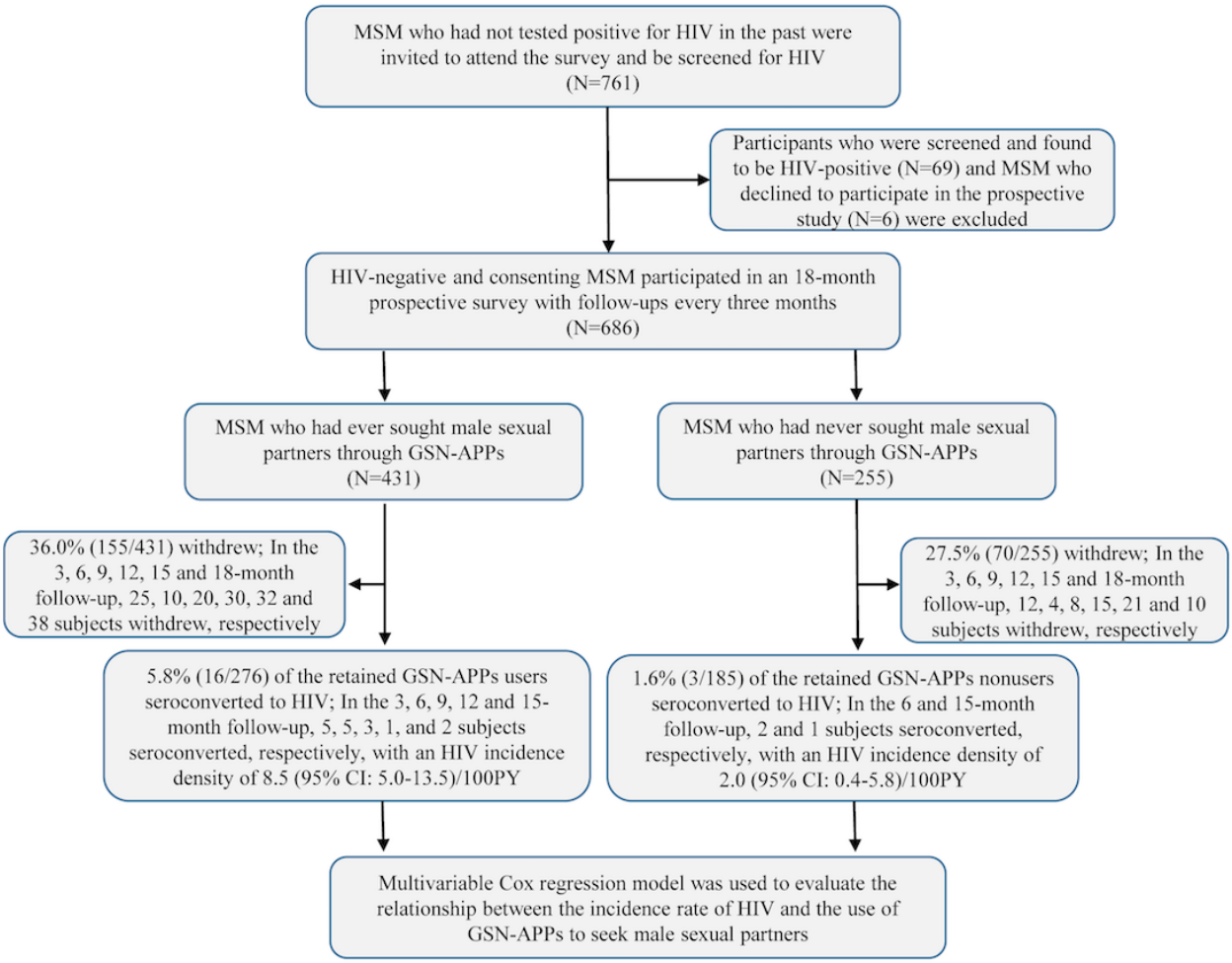


Table 1. Sociodemographics and behavioral characteristics of geosocial networking apps' users and nonusers (N=686).

Characteristics	Total, N (%)	GSN[a] apps users, n (%)	GSN apps nonusers, n (%)
Total	686 (100)	431 (62.8)[b]	255 (37.2)[b,c]
Age (years)			
18-24	176 (25.7)	137 (32)	39 (15.3)[c]
>24	510 (74.3)	294 (68.2)	216 (84.7)
Residency			
Shenyang city	269 (39.2)	178 (41.3)	91 (35.7)
Other	417 (60.8)	253 (58.7)	164 (64.3)
Ethnicity			
Han	570 (83.1)	362 (84.0)	208 (81.6)
Other	116 (16.9)	69 (16.0)	47 (18.4)
Education level being college or above	319 (46.5)	236 (54.8)	83 (32.5)[c]
Marital status			
Single	466 (67.9)	311 (72.2)	155 (60.8)[d]
Married or divorced or widowed or cohabiting	220 (32.1)	120 (27.8)	100 (39.2)
Monthly income			
≤3000 RMB/Yuan	392 (57.1)	254 (58.9)	138 (54.1)
>3000 RMB/Yuan	294 (42.9)	177 (41.1)	117 (45.9)
Currently a university student	79 (11.5)	67 (15.5)	12 (4.7)[c]
Age of sexual debut with males ≤20 years	320 (46.6)	234 (54.3)	86 (33.7)[c]
Anal sex position			
Versatile	287 (41.8)	185 (42.9)	102 (40.0)
Bottom	264 (38.5)	151 (35.0)	113 (44.3)[e]
Top	135 (19.7)	95 (22.0)	40 (15.7)
Ever been tested for HIV	160(23.3)	113(26.2)	47 (18.4)[e]
Used recreational drugs in the past 3 months	169 (24.6)	141 (32.7)	28 (11.0)[c]
Had male SPs[f] in the past 3 months	407 (59.3)	250 (58.0)	157 (61.6)
Had male CPs[g] in the past 3 months	334 (48.7)	224 (52.0)	110 (43.1)[e]
Two or more male SPs in the past 3 months	68 (9.9)	40 (9.3)	28 (11.0)
Five or more male CPs in the past 3 months	59 (8.6)	47 (10.9)	12 (4.7)[e]
CAI[h] with male CPs in the past 3 months	114 (16.6)	67 (15.5)	47 (18.4)
CAI with male SPs in the past 3 months	174 (25.4)	94 (21.8)	80 (31.4)[d]
Had group sex with males in the past 3 months	43 (6.3)	35 (8.1)	8 (3.1)[d]
Positive for syphilis at baseline	77 (11.2)	45 (10.4)	32 (12.5)
Willing to accept HIV prevention information push service through GSN apps	259 (37.8)	256 (59.4)	3 (1.2)

[a]GSN: geosocial networking.

[b]These percentages are out of the overall total number of men who have sex with men (N=686). The rest of the percentages are out of the specified group (Total, GSN apps' users, or GSN apps' nonusers).

[c]*P*<.001. Statistical significance was set at alpha=.05.

[d]*P*<.01.

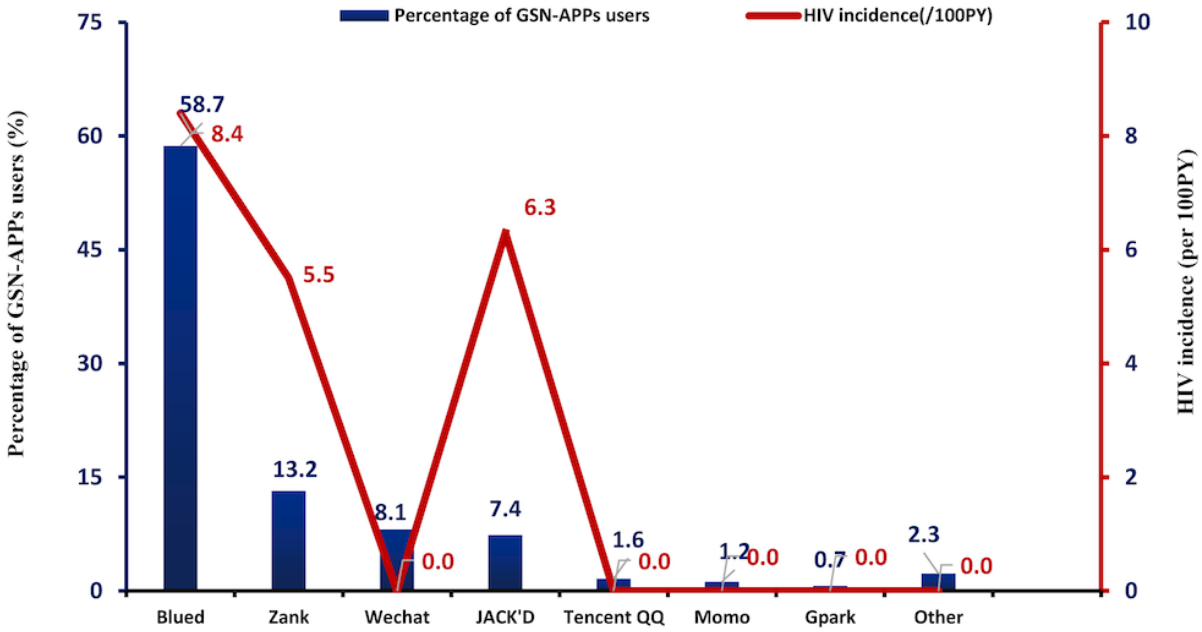
[e]The statistical significance of the difference between GSN apps' users and GSN apps nonusers $P<.05$.

[f]SPs: steady partners (sexual activity that takes places between male partners in a romantic relationship and usually implies commitment, emotional attachment, or familiarity between sexual partners).

[g]CPs: casual partners (a one-night stand or casual sex between males who have little or no history with each other).

[h]CAI: condomless anal intercourse.

Figure 2. The percentage of men who have sex with men who used geosocial networking apps over the past 3 months to seek male sexual partners and the HIV incidence density for each specific geosocial networking app. GSN: geosocial networking, PY: person-years.



HIV and Syphilis Incidence and Factors Correlated With HIV Seroconversion

During the follow-up period, 36.0% (155/431) of GSN apps' users and 27.5% (70/255) of nonusers withdrew from the study. In total, 19 MSM who remained in the study seroconverted to HIV during the study period, of which 16 were GSN apps' users and 3 were GSN apps' nonusers. The pooled HIV incidence densities were 8.5 (95% CI 5.0-13.5) per 100 PY among GSN apps' users and 2.0 (95% CI 0.4-5.8) per 100 PY among nonusers and were significantly different from each other (aHR 3.7, 95% CI 1.1-13.1, P=.04). Figure 2 showed the percentage of MSM GSN apps users to seek male sexual partners and the corresponding HIV incidence density in each group.

In total, 56 MSM who remained in the study became seropositive to syphilis during the follow-up, of which 34 were GSN apps' users and 22 were GSN apps' nonusers. The syphilis incidence densities were similar among GSN apps' users (17.4, 95% CI 12.0-22.7 per 100 PY) and nonusers (16.1, 95% CI 9.8-22.3 per 100 PY).

The Kaplan-Meier curves show the differences of the cumulative hazard ratios of HIV seroconversion between GSN apps' users and nonusers (Figure 3, Panel A), between those who used recreational drugs in the past 3 months and those who did not (Figure 3, Panel B), between MSM who had CAI with CPs in the past 3 months and those who did not (Figure 3, Panel C), and between MSM who had group sex with males in the past 3 months and those who did not (Figure 3, Panel D).

Table 2 contains the HIV incidence among MSM retained in our cohort. Table 3 shows the results of the multivariable Cox regression analysis for predictors correlating with HIV incidence after adjusting for age, level of education, registered residence, ethnicity, marital status, and monthly income. The following characteristics were independently associated with HIV incidence: have ever used GSN apps to seek male sexual partners (i.e. GSN-apps' users vs nonusers) (aHR 3.7, 95% CI 1.1-13.1, P=.04), have used recreational drugs in the past 3 months (aHR 2.6, 95% CI 1.0-6.9, P=.048), have had group sex with males in the past 3 months (aHR 4.8,95% CI 1.6-15.0, P=.01), and have had CAI with male CPs in the past 3 months (aHR 3.2, 95% CI 1.2-8.4, P=.02). The covariate of using GSN apps to seek male sexual partners in the past 3 months had a marginally significant statistical association with HIV incidence (aHR 2.6, 95% CI 0.9-7.9, P=.08). In contrast, age of sexual debut (P=.90), ever had sexual intercourse with females (P=.28), having anal sex position of versatile (P=.24) or bottom (P=.27) compared with top, having CAI with male SPs in the past 3 months (P=.12), the number of CPs (P=.30), ever being tested for HIV (P=.75), and syphilis infection at baseline (P=.67) were not statistically associated with HIV incidence (Table 3).

Figure 3. Kaplan-Meier estimates for high-risk factors of HIV seroconversion. GSN: geosocial networking, CAI: condomless anal intercourse, CP: casual partner.

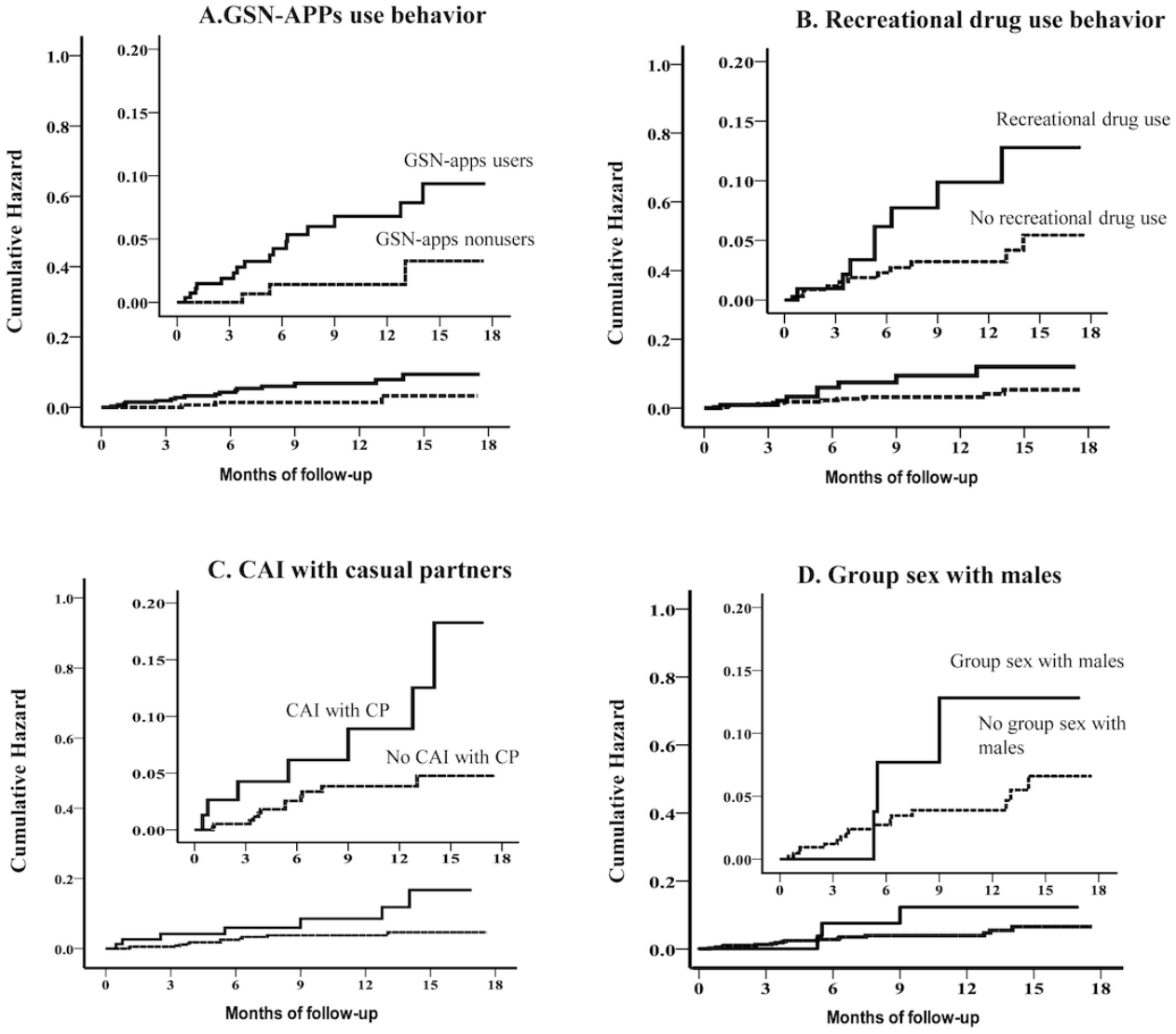


Table 2. HIV incidence among men who have sex with men retained in our cohort in Shenyang (N=461).

Characteristics	Total (N=461[a])	Incidence, n (%)	Observed person-years (PY)	Incidence rate (per 100 PY)
Age of sexual debut (years)				
≤20	205	9 (4.4)	153.1	5.9
>20	256	10 (3.9)	181.5	5.5
Ever had sexual intercourse with females				
Yes	246	7 (2.8)	177.2	4.0
No	215	12 (5.6)	157.4	7.6
Anal sex position				
Versatile	208	7 (3.4)	162.3	4.3
Bottom	168	6 (3.6)	118.1	5.1
Top	85	6 (7.1)	54.2	11.1
Ever been tested for HIV				
Yes	95	5 (5.3)	63.9	7.8
No	366	14 (3.8)	270.7	5.2
Used recreational drugs in the past 3 months				
Yes	141	10 (7.1)	75.3	13.3
No	320	9 (2.8)	259.3	3.5
Number of male steady partners in the past 3 months				
≥2	51	0 (0.0)	39.4	0.0
<2	410	19 (4.6)	295.2	6.4
Number of male casual partners in the past 3 months				
≥5	43	3 (7.0)	31.9	9.4
<5	418	16 (3.8)	302.7	5.3
Condomless anal intercourse with male casual partners in the past 3 months				
Yes	77	7 (9.1)	55.4	12.6
No	384	12 (3.1)	279.2	4.3
Condomless anal intercourse with male steady partners in the past 3 months				
Yes	120	2 (1.7)	89.2	2.2
No	341	17 (5.0)	245.5	6.9
Had group sex with males in the past 3 months				
Yes	31	4 (12.9)	28.6	14.0
No	430	15 (3.5)	306.1	4.9
Positive for syphilis at baseline				
Yes	59	3 (5.1)	47.0	6.4
No	402	16 (4.0)	287.6	5.6
Ever used geosocial networking apps to seek male sexual partners (ie, GSN-apps' users vs nonusers)				
Yes	264	16 (6.1)	187.3	8.5
No	197	3 (1.5)	147.3	2.0
Used geosocial networking apps to seek male sexual partners in the past 3 months				
Yes	239	14 (5.9)	175.5	8.0
No	222	5 (2.3)	159.2	3.1

[a]The number of GSN apps' users and nonusers retained to at least one 3-month follow-up visit.

Table 3. Multivariable Cox regression analysis of HIV incidence among men who have sex with men retained in our cohort in Shenyang (N=461).

Characteristics of users and nonusers[a]	Crude analysis	Multivariable analysis[b], cHR[c] (95% CI) aHR[d] (95% CI)	*P* value
Age of sexual debut (years)			
≤20	1.3 (0.5-3.3)	1.1 (0.4-2.8)	.90
>20	Reference	Reference	—[e]
Ever had sexual intercourse with females			
Yes	0.5 (0.2-1.3)	0.6 (0.2-1.6)	.28
No	Reference	Reference	—
Anal sex position			
Versatile	0.4 (0.1-1.2)	0.5 (0.1-1.6)	.24
Bottom	0.5 (0.1-1.4)	0.5 (0.1-1.7)	.27
Top	Reference	Reference	—
Ever been tested for HIV			
Yes	1.5 (0.5-4.1)	1.2 (0.4-3.4)	.75
No	Reference	Reference	—
Used recreational drugs in the past 3 months			
Yes	2.5 (1.0-6.2)	2.6 (1.0-6.9)	.048
No	Reference	Reference	—
Number of male SPs[f] in the past 3 months			
≥2	N/A[g]	N/A	N/A
<2	N/A	N/A	—
Number of male CPs[h] in the past 3 months			
≥5	1.7 (0.5-6.0)	1.9 (0.6-6.7)	.30
<5	Reference	Reference	—
CAI[i] with male CPs in the past 3 months			
Yes	2.9 (1.1-7.4)	3.2 (1.2-8.4)	.02
No	Reference	Reference	—
CAI with male SPs in the past 3 months			
Yes	0.3 (0.1-1.4)	0.3 (0.1-1.4)	.12
No	Reference	Reference	—
Had group sex with males in the past 3 months			
Yes	4.3 (1.4-13.1)	4.8 (1.6-15.0)	.01
No	Reference	Reference	—
Positive for syphilis at baseline			
Yes	1.2 (0.4-4.1)	1.3 (0.4- 4.8)	.67
No	Reference	Reference	—
Ever used GSN[j] apps to seek male sexual partners			
Yes	4.1 (1.2-14.2)	3.7 (1.1-13.1)	.04
No	Reference	Reference	—
Used GSN apps to seek male sexual partners in the past 3 months			

Characteristics of users and nonusers[a]	Crude analysis	Multivariable analysis[b], cHR[c] (95% CI)	
		aHR[d] (95% CI)	*P* value
Yes	2.9 (1.0-7.9)	2.6 (0.9-7.9)	.08
No	Reference	Reference	—

[a]The number of GSN apps' users and nonusers retained to at least one 3-month follow-up visit.

[b]The multivariable Cox regression model adjusted for age, residence status, ethnicity, education, income, and marital status.

[c]cHR: crude hazard ratio.

[d]aHR: adjusted hazard ratio.

[e]Not applicable.

[f]SPs: steady partners.

[g]N/A: not available.

[h]CPs: casual partners.

[i]CAI: condomless anal intercourse.

[j]GSN: geosocial networking.

Factors Correlated With Cohort Retention

Compared with MSM who withdrew from the cohort, MSM who were retained to at least one follow-up visit had lower HIV proportion of using GSN apps to seek male sexual partners (57.1% [264/461] vs. 74.2% [167/225], *P*<.01), displayed marginally higher rates of syphilis at baseline (12.8% [59/461] vs 8.0% [18/225], *P*=.06), had marginally higher rates of being older than 20 years at the age of sexual debut (55.5% [256/461] vs 48.4% [109/225], *P*=.08), and had marginally higher proportion of ≥2 SPs in the past 3 months (11.1% [51/461] vs 6.7% [15/225], *P*=.07). There were no statistically significant differences between these 2 groups over the 3 months before the baseline interview in terms of recreational drug use (*P*=.65), having CAI with male CPs (*P*=.93), having CAI with male SPs (*P*=.57), the number of CPs (*P*=.33), having group sex with males (*P*=.48), and having used GSN apps to seek male sexual partners (*P*=.12; Table 4).

Table 4. Comparisons of high-risk factors for HIV infection between men who have sex with men who were retained and who withdrew from the follow-up (N=686).

High-risk factors for HIV infection	Retained (n=461), n (%)	Withdrew (n=225), n (%)	Chi-square	*df*	*P* value
Used recreational drugs in the past 3 months					
Yes	141 (30.6)	65 (28.9)	0.2	1	.65
No	320 (69.4)	160 (71.1)	N/A[a]	N/A	N/A
Age of sexual debut (years)					
≤20	205 (44.5)	116 (51.6)	3.1	1	.08
>20	256 (55.5)	109 (48.4)	N/A	N/A	N/A
CAI[b] with male CPs[c] in the past 3 months					
Yes	77 (16.7)	37 (16.4)	0.0	1	.93
No	384 (83.3)	188 (83.6)	N/A	N/A	N/A
CAI with male SPs[d] in the past 3 months					
Yes	120 (26.0)	54 (24.0)	0.3	1	.57
No	341 (74.0)	171 (76.0)	N/A	N/A	N/A
Had group sex with males in the past 3 months					
Yes	31 (6.7)	12 (5.3)	0.5	1	.48
No	430 (93.3)	213 (94.7)	N/A	N/A	N/A
Number of SPs in the past 3 months					
≥2	51 (11.1)	15 (6.7)	3.4	1	.07
<2	410 (88.9)	210 (93.3)	N/A	N/A	N/A
Number of CPs in the past 3 months					
≥5	43 (9.3)	16 (7.1)	0.9	1	.33
<5	418 (90.7)	209 (92.9)	N/A	N/A	N/A
Positive for syphilis at baseline					
Yes	59 (12.8)	18 (8.0)	3.5	1	.06
No	402 (87.2)	207 (92.0)	N/A	N/A	N/A
Ever used GSN[f] apps to seek male sexual partners					
Yes	264 (57.1)	167 (74.2)	18.6	1	<.01
No	197 (42.7)	70 (25.8)	N/A	N/A	N/A
Used GSN apps to seek male sexual partners in the past 3 months					
Yes	239 (51.8)	131 (58.2)	2.5	1	.12
No	222 (48.2)	94 (41.8)	N/A	N/A	N/A

[a]N/A: not applicable.

[b]CAI: condomless anal intercourse.

[c]CPs: casual partners.

[d]SPs: steady partners.

[e]GSN: geosocial networking.

Discussion

Principal Findings

This 18-month prospective cohort study found that GSN apps' users had significantly higher HIV incidence compared with GSN apps' nonusers. We also determined possible mechanisms for how GSN apps' use causes higher HIV incidence. GSN apps' users were more likely to participate in group sex with males, have CAI with male CPs, use recreational drugs, and could have used GSN apps to facilitate these risky behaviors that are associated with HIV infection. In addition, 59.4% (256/431) of GSN apps' users were willing to accept HIV prevention information disseminated through these GSN apps.

Comparison With Prior Work

Almost all previous peer-reviewed GSN apps–related surveys in MSM have been conducted in the United States and China. In our study, 62.8% (431/686) of MSM participants had used GSN apps at least once to seek male sexual partners, which is similar to the percentage previously published in the United States (36.0%-63.6%) [7,11,12], but slightly higher than a prior study of Chinese MSM (40.6%) [6]. In addition, 85.8% (370/431) GSN apps' users had sought male sexual partners through GSN apps in the past 3 months; this percentage is higher than the percentage of MSM GSN apps' users in the United States (56.0%), and the median duration for which MSM use GSN apps to seek male sexual partners in our cohort is similar to that used by MSM in the United States (about 12 months) [17,18]. These results suggest that GSN apps' use in China is similar to that in the United States; thus, China's HIV prevention strategies targeting MSM using GSN apps can build on previous experiences of using GSN-APP platforms in the United States to conduct improved novel HIV prevention approaches focused on MSM [24], although the types of GSN apps used by MSM to seek sexual partners may be different.

Significance of the Study Results

Since the emergence of GSN apps, it is unclear whether their use increases the risk of HIV among their users. It has been speculated that as GSN apps allow for easier access to casual sexual relationships, their use increases the number of sexual partners and, thus, increases the risk of HIV infection [12-16]. In contrast, others have argued that specifically GSN apps' use does not increase HIV-related high-risk behaviors including CAI [6,19] and, thus, does not increase the risk of HIV infection [11,14,16]. As all previous peer surveys were cross-sectional studies, they were unable to establish temporality and, thus, were unable to draw conclusions on the relationship between putative causes and the outcome of HIV infection. This study showed that among MSM in China, GSN apps' users have nearly 4 times the HIV incidence rate of nonusers. As this study is a prospective cohort survey, it can not only evaluate whether the HIV incidence rate is linked with GSN apps' use but can also control the influence of related confounding factors. The multivariable Cox regression model indicated that certain high-risk behaviors are significantly correlated with higher HIV incidence rates after adjusting for potential confounding factors.

In addition, the prospective study allowed for a temporal sequence between putative cause and outcome and, thus, addressed a critical gap in the available literature about GSN apps' use and new HIV infections among MSM. We were able to determine potential mechanisms underpinning how GSN apps' use may lead to new HIV infections for its users. We found that GSN apps' users were more likely to use recreational drugs, have larger numbers of male CPs, and have group sex with males compared with GSN apps' nonusers. These high-risk behaviors for HIV infection were later confirmed in the multivariable Cox regression analysis to be independent correlates of HIV incidence, and these results were consistent with previous publications [8,10-12,16,25-27]. These results suggest that GSN apps' use increases the HIV incidence rate among their users through facilitating recreational drug use and higher numbers of sexual partners. Interestingly, we found ever using GSN apps to seek male sexual partners at baseline was an independent significant predictor of HIV seroconversion (P=.04), but the covariate of using GSN apps to seek male sexual partners in the past 3 months only had a marginal statistical association with study outcome (.05<P<.10). One of the possible reasons for the above difference may be attributed to insufficient efficiency of statistical power for the latter covariate. Statistical power is positively associated with sample size, and the number of participants who ever used GSN apps to seek male sexual partners in this study was just relatively higher than that of participants who used GSN apps to seek male sexual partners in the past 3 months (264 vs 226), which may partly explain the above inconsistence of P values. In this study, we used time-dependent Cox regression model to analyze the influence of GSN apps' use on HIV incidence. The baseline life-time GSN-app using behavior and the GSN app using behavior in past 3 months was set as a fixed covariate and time-dependent covariate, respectively. This data analysis strategy may help public health workers to fully understand the influence of GSN app use behavior within different window periods on HIV seroconversion risk.

Encouragingly, we also found that 59.4% (256/431) of GSN apps' users in this survey were willing to accept HIV prevention information disseminated through GSN apps. These results have important implications considering the severe social discrimination toward MSM, low sexual orientation disclosure rate, and low HIV testing rate in China [28]. Recently, some social media platforms, including Facebook and Grindr, have collaborated with researchers to disseminate HIV prevention information, promote HIV testing, and link MSM to medical care [24,29]. Thus, future steps include developing interventions circulated through these GSN-APP platforms to reach the target high-risk MSM population to mitigate the HIV epidemic in this community. Further studies need to evaluate the relative impact of HIV prevention interventions disseminated through GSN apps used by MSM compared with traditional facility-based interventions at voluntary counseling and testing clinics or hospitals.

Our study indicated that MSM who use GSN apps compared with nonusers were more likely to be younger than 24 years (31.8% [137/431] vs 15.3% [39/255]) and to be university students (15.5% [67/431] vs 4.7% [12/255]). Currently in China, rates of new HIV infections among young MSM, especially university students, have greatly increased [30]. The Chinese government reported that the number of 15- to 24-year-olds in China who live with HIV more than doubled from 8354 people living with HIV (PLWH) in 2008 to 16,986 PLWH in 2015. Furthermore, the proportion of university students among PLWHs aged between 15 and 24 years increased from 5.8% in 2008 to 19.1% in 2015. Our study results indicate that many young MSM using GSN apps suggest that using these platforms to promote HIV prevention strategies could be effective at targeting young MSM in China.

Future Studies

The results suggest that GSN apps' users were significantly associated with higher education levels and higher HIV testing

rates compared with nonusers. Studies have shown that people with higher levels of education tend to have higher incomes [31] and, thus, are more likely to have the income needed to purchase expensive smartphones that recognize GSN app software. Less than 50% of Chinese MSM in a prior study got tested for HIV in the previous 12 months [32]; this low HIV testing rate is a serious obstacle in controlling the HIV epidemic [33]. Studies in the United States and in the United Kingdom have shown that promotion of HIV testing can be effectively conducted through GSN app platforms [34,35]. However, currently, there is no published research discussing using GSN app platforms to promote HIV testing among Chinese MSM. The study results support integrating GSN app platforms, in particular Blued, into public health HIV testing promotion strategies to reach MSM.

As there were no significant differences in high-risk sexual behaviors between those who were retained in the study and those who withdrew, it is possible that missing data from those who withdrew from the cohort did not lead to serious bias. Thus, the HIV incidence rate derived from the MSM who were retained to at least one follow-up visit may accurately represent the HIV incidence of the overall recruited MSM population.

Study Strengths

The study design was a prospective cohort study; this study was conducted among a relatively large sample of MSM and controlled for the influence of many relevant confounders. In addition, it included information on sociodemographics, high-risk behavior for HIV infection, and laboratory testing for HIV and syphilis. Moreover, this study explored possible mechanisms through which GSN apps' use leads to an increase in the HIV incidence rate; the study results suggested the association between GSN apps' use and higher HIV incidence rate is possibly mediated through GSN apps facilitating recreational drug use and multiple male CPs.

Limitations

A potential limitation of this study is reporting bias because of social expectation about the self-reported HIV-related high-risk behaviors, thus leading to underestimation of these behaviors. Second, this study was conducted at a single site, thus limiting extrapolation of its results. Third, participants were not recruited randomly, so the characteristics of participants in this study may not represent well the entire MSM population in Shenyang. Although this study found that ever using GSN apps was correlated with higher HIV incidence rate, GSN apps' use in the past 3 months was only marginally correlated with HIV incidence. Thus, a larger prospective cohort is needed to further examine the causal relationship between GSN apps' use and HIV incidence. Finally, approximately 30% of participants withdrew from the prospective cohort during the follow-up period, and the prevalence of syphilis among MSM who withdrew at baseline was marginally lower than that among those who were retained. As syphilis infection can be used as a proxy for unprotected sex, the HIV incidence of Shenyang MSM may be slightly overestimated.

Conclusions

The GSN apps' users had higher incidence rates of HIV seroconversion than nonusers, which may be influenced by their higher rates of HIV-related high-risk behavior, including recreational drug use and multiple CPs. Thus, public health workers must collaborate with GSN-app operators to develop an Web-based and offline comprehensive HIV intervention strategy targeting users of these platforms to mitigate the HIV epidemic among MSM.

Acknowledgments

This study was supported by China's Ministry of Science and Technology, the Mega-projects of National Science Research for the thirteenth 5-Year Plan (2017ZX10201101-002-007), National Natural Science Foundation of China (81872674), and National Science and Technology Major Project (2018ZX10101-001-001-003). The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication. The findings and conclusions in this study are those of the authors and do not necessarily represent the views of the National Health and Family Planning Commission.

Authors' Contributions

HS, HY, and YJJ conceived and designed the experiments; JJX, HY, WQG, and YJJ performed the study and experiments; HY, JZ, WMT, XM, and HYW analyzed the data; and HS, JJX, HY, JZ, XM, WMT, HYW, and SIL wrote and revised the manuscript. All authors reviewed the manuscript.

Conflicts of Interest

None declared.

References

1. Sullivan PS, Hamouda O, Delpech V, Geduld JE, Prejean J, Semaille C, Annecy MSM Epidemiology Study Group. Reemergence of the HIV epidemic among men who have sex with men in North America, Western Europe, and Australia, 1996-2005. Ann Epidemiol 2009 Jun;19(6):423-431. [doi: 10.1016/j.annepidem.2009.03.004] [Medline: 19460672]
2. Millett GA, Peterson JL, Flores SA, Hart TA, Jeffries WL, Wilson PA, et al. Comparisons of disparities and risks of HIV infection in black and other men who have sex with men in Canada, UK, and USA: a meta-analysis. Lancet 2012 Jul 28;380(9839):341-348. [doi: 10.1016/S0140-6736(12)60899-X] [Medline: 22819656]

3. Baral SD, Friedman MR, Geibel S, Rebe K, Bozhinov B, Diouf D, et al. Male sex workers: practices, contexts, and vulnerabilities for HIV acquisition and transmission. Lancet 2015 Jan 17;385(9964):260-273 [FREE Full text] [doi: 10.1016/S0140-6736(14)60801-1] [Medline: 25059939]
4. Shang H, Xu J, Han X, Spero Li J, Arledge KC, Zhang L. HIV prevention: bring safe sex to China. Nature 2012 May 30;485(7400):576-577. [doi: 10.1038/485576a] [Medline: 22660303]
5. Pan S, Xu JJ, Han XX, Zhang J, Hu QH, Chu ZX, et al. Internet-based sex-seeking behavior promotes HIV infection risk: a 6-year serial cross-sectional survey to MSM in Shenyang, China. Biomed Res Int 2016 Jan;2016(1):2860346-2860345 [FREE Full text] [doi: 10.1155/2016/2860346] [Medline: 28105415]
6. Bien CH, Best JM, Muessig KE, Wei C, Han L, Tucker JD. Gay apps for seeking sex partners in China: implications for MSM sexual health. AIDS Behav 2015 Jun;19(6):941-946 [FREE Full text] [doi: 10.1007/s10461-014-0994-6] [Medline: 25572834]
7. Phillips G, Magnus M, Kuo I, Rawls A, Peterson J, Jia Y, et al. Use of geosocial networking (GSN) mobile phone applications to find men for sex by men who have sex with men (MSM) in Washington, DC. AIDS Behav 2014 Sep;18(9):1630-1637. [doi: 10.1007/s10461-014-0760-9] [Medline: 24682866]
8. Francisco Luz Nunes Queiroz AA, Lopes de Sousa Á, Evangelista de Araújo TM, Milanez de Oliveira FB, Batista Moura ME, Reis RK. A review of risk behaviors for HIV infection by men who have sex with men through geosocial networking phone apps. J Assoc Nurses AIDS Care 2017;28(5):807-818. [doi: 10.1016/j.jana.2017.03.009] [Medline: 28456472]
9. Beyrer C, Sullivan P, Sanchez J, Baral SD, Collins C, Wirtz AL, et al. The increase in global HIV epidemics in MSM. AIDS 2013 Nov 13;27(17):2665-2678. [doi: 10.1097/01.aids.0000432449.30239.fe] [Medline: 23842129]
10. Tang W, Best J, Zhang Y, Liu FY, Tso LS, Huang S, et al. Gay mobile apps and the evolving virtual risk environment: a cross-sectional online survey among men who have sex with men in China. Sex Transm Infect 2016 Nov;92(7):508-514 [FREE Full text] [doi: 10.1136/sextrans-2015-052469] [Medline: 27288414]
11. Beymer MR, Weiss RE, Bolan RK, Rudy ET, Bourque LB, Rodriguez JP, et al. Sex on demand: geosocial networking phone apps and risk of sexually transmitted infections among a cross-sectional sample of men who have sex with men in Los Angeles County. Sex Transm Infect 2014 Nov;90(7):567-572 [FREE Full text] [doi: 10.1136/sextrans-2013-051494] [Medline: 24926041]
12. Holloway IW. Substance use homophily among geosocial networking application using gay, bisexual, and other men who have sex with men. Arch Sex Behav 2015 Oct;44(7):1799-1811 [FREE Full text] [doi: 10.1007/s10508-015-0581-6] [Medline: 26216146]
13. Lehmiller JJ, Ioerger M. Social networking smartphone applications and sexual health outcomes among men who have sex with men. PLoS One 2014;9(1):e86603 [FREE Full text] [doi: 10.1371/journal.pone.0086603] [Medline: 24466166]
14. Landovitz RJ, Tseng CH, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. J Urban Health 2013 Aug;90(4):729-739 [FREE Full text] [doi: 10.1007/s11524-012-9766-7] [Medline: 22983721]
15. Holloway IW, Pulsipher CA, Gibbs J, Barman-Adhikari A, Rice E. Network influences on the sexual risk behaviors of gay, bisexual and other men who have sex with men using geosocial networking applications. AIDS Behav 2015 Jun;19 Suppl 2:112-122 [FREE Full text] [doi: 10.1007/s10461-014-0989-3] [Medline: 25572832]
16. Zou H, Fan S. Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: a systematic review and meta-analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]
17. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296 [FREE Full text] [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
18. Winetrobe H, Rice E, Bauermeister J, Petering R, Holloway IW. Associations of unprotected anal intercourse with Grindr-met partners among Grindr-using young men who have sex with men in Los Angeles. AIDS Care 2014;26(10):1303-1308. [doi: 10.1080/09540121.2014.911811] [Medline: 24754563]
19. Grosskopf NA, LeVasseur MT, Glaser DB. Use of the Internet and mobile-based "apps" for sex-seeking among men who have sex with men in New York City. Am J Mens Health 2014 Nov;8(6):510-520. [doi: 10.1177/1557988314527311] [Medline: 24658284]
20. Burrell ER, Pines HA, Robbie E, Coleman L, Murphy RD, Hess KL, et al. Use of the location-based social networking application GRINDR as a recruitment tool in rectal microbicide development research. AIDS Behav 2012 Oct;16(7):1816-1820 [FREE Full text] [doi: 10.1007/s10461-012-0277-z] [Medline: 22851153]
21. Piyaraj P, van Griensven F, Holtz TH, Mock PA, Varangrat A, Wimonsate W, et al. The finding of casual sex partners on the internet, methamphetamine use for sexual pleasure, and incidence of HIV infection among men who have sex with men in Bangkok, Thailand: an observational cohort study. Lancet HIV 2018 Jul;5(7):e379-e389. [doi: 10.1016/S2352-3018(18)30065-1] [Medline: 29861202]
22. Xu JJ, Yu YQ, Hu QH, Yan HJ, Wang Z, Lu L, et al. Treatment-seeking behaviour and barriers to service access for sexually transmitted diseases among men who have sex with men in China: a multicentre cross-sectional survey. Infect Dis Poverty 2017 Jan 18;6(1):15 [FREE Full text] [doi: 10.1186/s40249-016-0219-5] [Medline: 28095879]

23. Hsieh F, Lavori PW. Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. Control Clin Trials 2000 Dec;21(6):552-560. [Medline: 11146149]
24. Cao B, Gupta S, Wang J, Hightow-Weidman LB, Muessig KE, Tang W, et al. Social media interventions to promote HIV testing, linkage, adherence, and retention: systematic review and meta-analysis. J Med Internet Res 2017 Nov 24;19(11):e394 [FREE Full text] [doi: 10.2196/jmir.7997] [Medline: 29175811]
25. Xu JJ, Zhang C, Hu QH, Chu ZX, Zhang J, Li YZ, et al. Recreational drug use and risks of HIV and sexually transmitted infections among Chinese men who have sex with men: mediation through multiple sexual partnerships. BMC Infect Dis 2014 Dec 02;14(11-12):642-800 [FREE Full text] [doi: 10.1186/s12879-014-0642-9] [Medline: 25443542]
26. Tang W, Tang S, Qin Y, Zhang Y, Zhang W, Liu C, et al. Will gay sex-seeking mobile phone applications facilitate group sex? A cross-sectional online survey among men who Have sex with men in China. PLoS One 2016;11(11):e0167238 [FREE Full text] [doi: 10.1371/journal.pone.0167238] [Medline: 27880823]
27. Yeo TE, Ng YL. Sexual risk behaviors among apps-using young men who have sex with men in Hong Kong. AIDS Care 2016;28(3):314-318. [doi: 10.1080/09540121.2015.1093597] [Medline: 26444390]
28. Zhao Y, Ma Y, Chen R, Li F, Qin X, Hu Z. Non-disclosure of sexual orientation to parents associated with sexual risk behaviors among gay and bisexual MSM in China. AIDS Behav 2016 Jan;20(1):193-203. [doi: 10.1007/s10461-015-1135-6] [Medline: 26174317]
29. Young LE, Fujimoto K, Schneider JA. HIV prevention and sex behaviors as organizing mechanisms in a Facebook group affiliation network among young black men who have sex with men. AIDS Behav 2018 Mar 13;22(10):3324-3334. [doi: 10.1007/s10461-018-2087-4] [Medline: 29536285]
30. McLaughlin K. HIV infections are spiking among young gay Chinese. Science 2017 Mar 31;355(6332):1359. [doi: 10.1126/science.355.6332.1359] [Medline: 28360273]
31. Kim D, Han K, Choi JY, Ersek J, Liu J, Jo SJ, et al. Age- and sex-specific relationships between household income, education, and diabetes mellitus in Korean adults: the Korea National Health and Nutrition Examination Survey, 2008-2010. PLoS One 2015;10(1):e0117034 [FREE Full text] [doi: 10.1371/journal.pone.0117034] [Medline: 25622031]
32. Zou H, Hu N, Xin Q, Beck J. HIV testing among men who have sex with men in China: a systematic review and meta-analysis. AIDS Behav 2012 Oct;16(7):1717-1728. [doi: 10.1007/s10461-012-0225-y] [Medline: 22677975]
33. Choi KH, Lui H, Guo Y, Han L, Mandel JS. Lack of HIV testing and awareness of HIV infection among men who have sex with men, Beijing, China. AIDS Educ Prev 2006 Feb;18(1):33-43. [doi: 10.1521/aeap.2006.18.1.33] [Medline: 16539574]
34. Lampkin D, Crawley A, Lopez TP, Mejia CM, Yuen W, Levy V. Reaching suburban men who have sex with men for STD and HIV services through online social networking outreach: a public health approach. J Acquir Immune Defic Syndr 2016 May 01;72(1):73-78. [doi: 10.1097/QAI.0000000000000930] [Medline: 27097365]
35. Elliot E, Rossi M, McCormack S, McOwan A. Identifying undiagnosed HIV in men who have sex with men (MSM) by offering HIV home sampling via online gay social media: a service evaluation. Sex Transm Infect 2016 Dec;92(6):470-473. [doi: 10.1136/sextrans-2015-052090] [Medline: 27221475]

Abbreviations

aHR: adjusted hazard ratio
CAI: condomless anal intercourse
cHR: crude hazard ratio
CP: casual partner
GSN: geosocial networking
HR: hazard ratio
MSM: men who have sex with men
NAAT: nucleic acid amplification test
PLWH: people living with HIV
PY: person-years
RPR: rapid plasma regain
SP: steady partner
STI: sexually transmitted infection
TPPA: *Treponema pallidum particle agglutination assay*

Edited by G Eysenbach; submitted 19.06.18; peer-reviewed by Z Wang, HF Raymond; comments to author 30.08.18; revised version received 10.09.18; accepted 27.09.18; published 21.12.18.

Please cite as:
Xu J, Yu H, Tang W, Leuba SI, Zhang J, Mao X, Wang H, Geng W, Jiang Y, Shang H
The Effect of Using Geosocial Networking Apps on the HIV Incidence Rate Among Men Who Have Sex With Men: Eighteen-Month Prospective Cohort Study in Shenyang, China
J Med Internet Res 2018;20(12):e11303
URL: http://www.jmir.org/2018/12/e11303/
doi:10.2196/11303
PMID:30578225

Original Paper

Clinical Feasibility of a Just-in-Time Adaptive Intervention App (iREST) as a Behavioral Sleep Treatment in a Military Population: Feasibility Comparative Effectiveness Study

I Wayan Pulantara[1], PhD; Bambang Parmanto[1], PhD; Anne Germain[2], PhD

[1]Health and Rehabilitation Informatics Laboratory, Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA, United States

[2]Sleep and Behavioral Neuroscience Center, Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Anne Germain, PhD
Sleep and Behavioral Neuroscience Center
Department of Psychiatry
University of Pittsburgh
Sterling Plaza 240
Pittsburgh, PA,
United States
Phone: 1 412 383 2150
Email: germax@upmc.edu

Abstract

Background: Although evidence-based cognitive behavioral sleep treatments have been shown to be safe and effective, these treatments have limited scalability. Mobile health tools can address this scalability challenge. iREST, or interactive Resilience Enhancing Sleep Tactics, is a mobile health platform designed to provide a just-in-time adaptive intervention (JITAI) in the assessment, monitoring, and delivery of evidence-based sleep recommendations in a scalable and personalized manner. The platform includes a mobile phone–based patient app linked to a clinician portal.

Objective: The first aim of the pilot study was to evaluate the effectiveness of JITAI using the iREST platform for delivering evidence-based sleep interventions in a sample of military service members and veterans. The second aim was to explore the potential effectiveness of this treatment delivery form relative to habitual in-person delivery.

Methods: In this pilot study, military service members and veterans between the ages of 18 and 60 years who reported clinically significant service-related sleep disturbances were enrolled as participants. Participants were asked to use iREST for a period of 4 to 6 weeks during which time they completed a daily sleep/wake diary. Through the clinician portal, trained clinicians offered recommendations consistent with evidence-based behavioral sleep treatments on weeks 2 through 4. To explore potential effectiveness, self-report measures were used, including the Insomnia Severity Index (ISI), the Pittsburgh Sleep Quality Index (PSQI), and the PSQI Addendum for Posttraumatic Stress Disorder.

Results: A total of 27 participants completed the posttreatment assessments. Between pre- and postintervention, clinically and statistically significant improvements in primary and secondary outcomes were detected (eg, a mean reduction on the ISI of 9.96, t_{26}=9.99, P<.001). At posttreatment, 70% (19/27) of participants met the criteria for treatment response and 59% (16/27) achieved remission. Comparing these response and remission rates with previously published results for in-person trials showed no significant differences.

Conclusion: Participants who received evidence-based recommendations from their assigned clinicians through the iREST platform showed clinically significant improvements in insomnia severity, overall sleep quality, and disruptive nocturnal disturbances. These findings are promising, and a larger noninferiority clinical trial is warranted.

(J Med Internet Res 2018;20(12):e10124) doi:10.2196/10124

KEYWORDS

just-in-time adaptive intervention; insomnia; sleep; mHealth; mobile health; interactive Resilience Enhancing Sleep Tactics (iREST); behavioral therapy; brief behavioral therapy for insomnia; cognitive behavioral therapy for insomnia

Introduction

Sleep disturbances such as insomnia and nightmares are among the most prevalent complaints reported by post–9/11 military service members (SMs) and veterans [1-3]. Insomnia affects between 40% and 70% of SMs and veterans [4] and can compromise readiness by impairing critical cognitive and moral reasoning abilities while increasing the risk of injuries and costly mishaps due to the resulting fatigue [5].

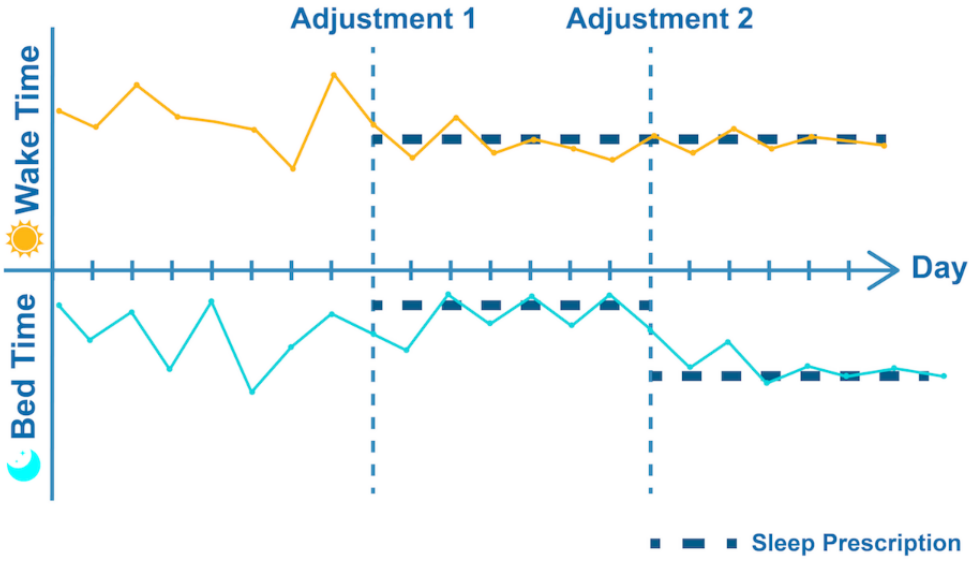
Insomnia also constitutes a robust risk factor for poor psychological health outcomes, including posttraumatic stress disorder (PTSD), major depressive disorder, suicidal tendencies, hazardous alcohol use, and addictive disorders [6-8]. Furthermore, insomnia impedes the response to treatment of those aforementioned conditions and increases the risk of onset or recurrence [9].

Insomnia is a treatable sleep disorder and a modifiable risk factor of compromised readiness and health. The National Institutes of Health and the American College of Physicians recommend nonpharmacological treatments for insomnia [10,11]**.** Nonpharmacological treatments are commonly the core of cognitive behavioral therapy for insomnia (CBTI). These treatment protocols are typically delivered in person over 1 to 4 sessions (brief version) or to 5 to 8 sessions (standard version) and are usually delivered by a licensed psychologist trained in behavioral sleep medicine [12-14] or a master's level clinician [15-17]. CBTI has been shown to be safe, effective, and associated with durable improvements [18] in the general population [19-21] and in military samples [22-25].

Nevertheless, the scalability of CBT for sleep disturbances remains limited. One of the main barriers in making CBTI widely available is the shortage of trained clinicians and availability of expertise outside urban centers. For instance, specialty sleep care clinics are not readily available in rural areas in the United States, where approximately 25% of veterans are located [26], and the more than 150 countries where US Armed Forces are stationed. Furthermore, the traditional in-person treatment format often creates barriers to receiving or adhering to treatment visit schedules due to travel distance (to and from the clinics), conflict with work and family schedules, childcare availability, etc. To reduce these potential burdens, brief behavioral treatment protocols (1 to 4 sessions) and online programs have been developed and tested. Brief in-person programs yield comparable or greater benefits as standard longer 6- to 8-week CBTI protocols [27-31]. Telehealth programs [32] and online commercial treatment programs such as SHUTi (BeHealth Solutions LLC) [33,34], Sleepio (Sleepio Ltd) [35], and RESTore (CCBT Ltd) [36] have also been shown to be efficacious and typically require anywhere between 5 weeks to several months of patient engagement [25,33,37,38]. From the patient perspective, traditional CBTI typically requires them to keep a paper sleep diary, which is cumbersome. Most importantly, the rigid schedule of the current CBTI delivery formats (ie, weekly in-person visits) limits the clinician's ability to personalize the intervention (ie, deliver the right intervention, at the right time, to the right patient).

Current advancements in mobile technology and increases in its adoption have the potential to increase access to evidence-based behavioral sleep treatments as well as to enhance the efficacy of these interventions by tailoring them to each individual's dynamic moment-to-moment needs. For example, a patient with insomnia typically shows high night-to-night variability in wake times and bedtimes, which leads to irregular sleep duration and unpredictable sleep quality (Figure 1).

Figure 1. An illustration of just-in-time adjustment of sleep recommendations consistent with sleep restriction and stimulus control based on changes in a patient's sleep pattern. Before the start of the treatment, a high night-to-night variability in wake times and bedtimes was observed. A sleep restriction recommendation was sent to the patient (Adjustment 1). After several nights, the patient adjusted to this restriction and achieved a reduction in the night-to-night wake/bedtimes variability. At this point, the interactive Resilience Enhancing Sleep Tactics (iREST) portal would suggest a reduction in the amount of sleep restriction (lengthening the recommended time allowed in bed). With clinician approval, this recommendation was sent to the patient's iREST app (Adjustment 2).



Baseline bedtime and wake times are determined by data collected by the patients using their mobile phone. Based on these parameters, a clinician is likely to use principles of sleep restriction [39] and/or stimulus control [40]. Sleep restriction is one of the most common behavioral treatments for insomnia and aims to optimize the predictability and quality of sleep by implementing regular wake times and bedtimes. Stimulus control aims to reinforce learned associations between sleep and specific environmental cues (eg, bed, bedroom). Based on data collected with daily sleep/wake diaries, the interactive Resilience Enhancing Sleep Tactics (iREST) portal calculates and suggests personalized recommendations for the implementation of sleep restriction and stimulus control. The clinician can then review and approve or modify these recommendations (or prescriptions) before they are forwarded to the patient through the iREST app (Figure 1, Adjustment 1). Recommendations sent to a patient include specific information on what new sleep behaviors should be adopted, how to implement the recommendations, and the rationale supporting each recommendation. As the patient adopts these recommendations, changes in behaviors and improvements in sleep quality and predictability are detected by the iREST system (Figure 1, Adjustment 2), which iteratively reassesses what behavioral changes may be required and, again with the clinician's approval, sends an adjustment in the personalized recommendations in a just-in-time fashion [41] until the desired sleep outcomes are achieved (ie, regular sleep behaviors and satisfactory sleep quality). This sort of adaptability includes personalization of the intervention not only at the beginning of the episode of care but throughout the intervention period, in the form of frequent iterative adjustments based on patient-reported data. This type of adaptability and personalization in delivering evidence-based interventions is known as a just-in-time adaptive intervention (JITAI) [42].

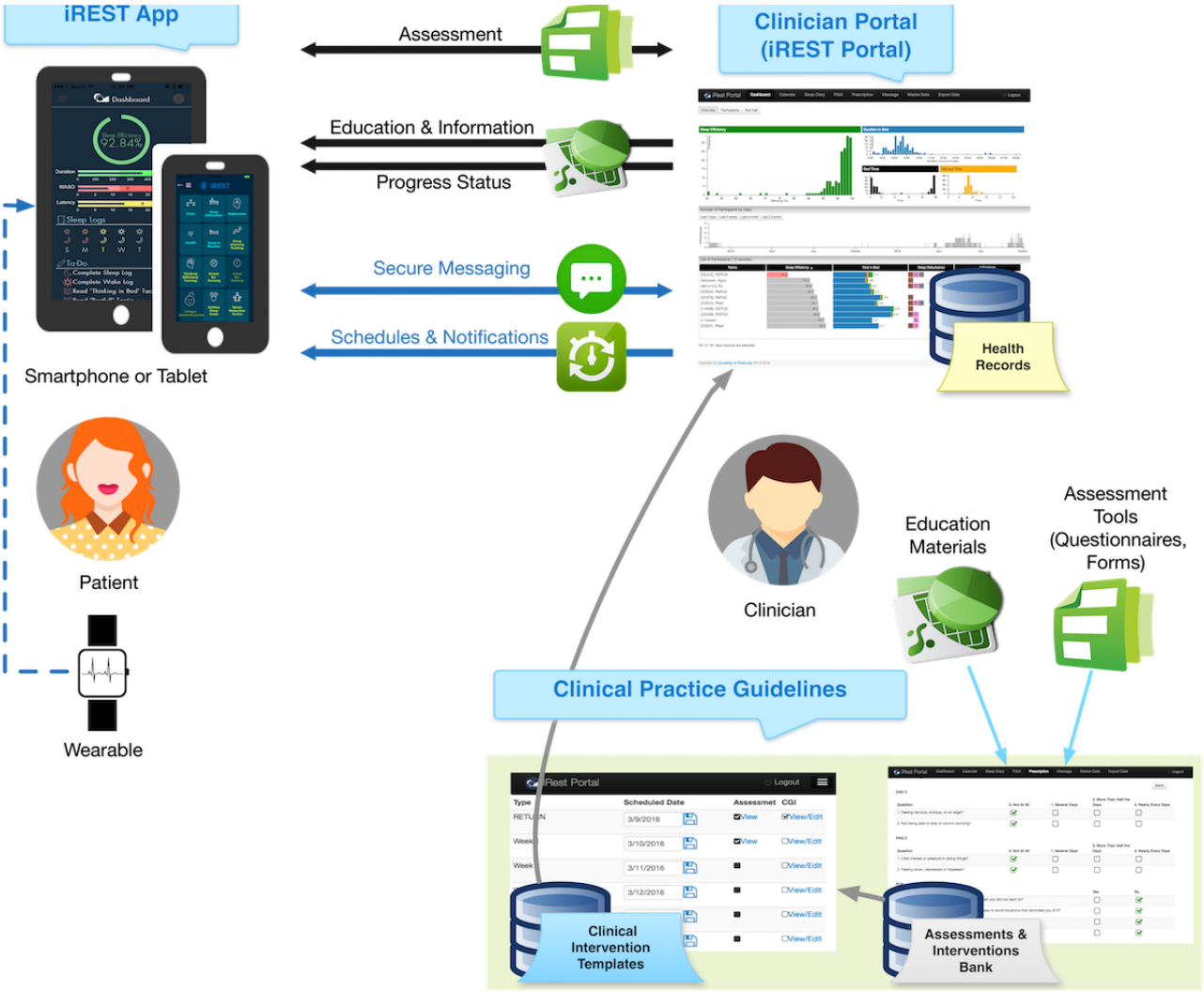
With this in mind, we developed iREST [43]. iREST is a JITAI implementation of existing behavioral sleep intervention techniques, particularly the military-version brief behavioral therapy for insomnia [17], an intervention that has been found to be effective in SMs and veterans. The iREST system consists of the following (Figure 2):

- Cross-platform mobile phone app [43] that records sleep data, shows feedback and related educational materials, and provides cues and notifications
- Web-based portal that allows therapists to monitor sleep information, prescribe treatment, and engage participants via secure messaging
- Wearable integration that allows objective measurement of the patient's sleep-wake pattern. The preliminary feasibility report on this integration has been published elsewhere [43]
- Communication protocol that allows real-time bidirectional exchange of data among the app, portal, and wearable sensors

iREST, as a mobile phone–based intervention, has the potential to improve the delivery of traditional CBTI with such novel features as personalization and context awareness. Assessments and interventions are best delivered when they are personalized to fit each individual's needs and conditions [44,45]. iREST can further tailor the treatment by dynamically adapting both the assessment and intervention. This ability to adapt the intervention can expand to accommodating the environment and social situation, especially important for the military population where training or deployment may not be compatible with prescribed sleep treatments. Such continual adaptation requires personalization of the intervention not only at the beginning of the episode of care but also frequent iterative adjustments during the course of care—something for which a JITAI such as iREST may reveal promising potential.

This pilot study first sought to evaluate the potential effectiveness of digital monitoring and delivery of evidence-based CBT for sleep disturbances in this sample using an open-trial design. In addition, to provide a comparative effectiveness framework, results were compared with previously published effect sizes and rates of treatment responses and remission following traditional [46] in-person CBT for sleep disturbances in this population [17].

Figure 2. A model representing the interactive Resilience Enhancing Sleep Tactics (iREST) app and clinician portal's two-way interactions including assessment, education/information delivery, progress reporting, scheduling, notification delivery, and secure messaging. The model also shows objective data gathering using wearable devices.



Methods

Participants

The University of Pittsburgh Institutional Review Board approved this study. SMs and veterans between the ages of 18 and 60 years were recruited from other studies (Military Operational Medicine Research Program proposal number PT130572, PI: Reifman/Germain; Military Operational Medicine Research Program log number 11293006, PI: Germain; log number 13154004, PI: Okonkwo) that used postcards, flyers, study websites, social media/Facebook (San Francisco, CA), and public television advertisements for recruiting purposes. Since our study required participants to use their own device, eligible SMs and veterans had to both own a mobile phone with internet access and be fluent in the use of that mobile phone. Other eligibility criteria included the presence of a clinically significant sleep complaint as determined by a baseline score of 10 or higher on the Insomnia Severity Index (ISI) [47] and having consistently experienced sleep disturbances for at least 1 month. Participants who were diagnosed with obstructive sleep apnea or who scored greater than or equal to 4 on the STOP-BANG (snoring, tiredness, observed apnea, blood pressure, body mass index, age, neck size, and gender) questionnaire [48-50] were excluded from the study. Other exclusion criteria included a history of psychotic disorder or bipolar disorder, the presence of symptoms of narcolepsy or any other sleep disorder requiring further evaluation and treatment, the presence of any severe or untreated psychiatric disorders associated with marked impairments in functioning, and any scheduled/imminent military deployment during the course of the study. Finally, pregnant or breastfeeding women were not included in the study.

Screening Procedures

After obtaining each participant's verbal consent, a telephone screening was conducted to assess eligibility prior to the initial in-person visit. Screening questions were related to the current use of a mobile phone, past and current psychiatric and physical health, and the presence of any suspected or diagnosed physiological sleep disorders or sleep apnea. Eligible participants were invited for an in-person consent and assessment visit.

After obtaining written informed consent, participants underwent a 2-part diagnostic evaluation that included a diagnostic interview and a series of screening questionnaires. The diagnostic interview focused on assessing the insomnia, presence and severity of trauma history, alcohol/substance use disorders, other psychiatric disorders, and current physical health. A weekly consensus meeting was held to review diagnostic information and establish the participant eligibility for the open-trial phase of the study. Participants also completed a series of self-report screening questionnaires:

- Structured Clinical Interview for the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV), nonpatient version [51]: used to assess the participant's past and current psychiatric history
- DSM Sleep Disorder: developed locally and similar to the Structured Clinical Interview, this instrument assesses the presence of core symptoms of sleep disorders as defined by the International Classification of Sleep Disorders [52], including insomnia, sleep-disordered breathing, restless legs syndrome and other sleep-related movement disorders, and parasomnias
- PTSD Checklist–Civilian version (PCL-C) [53]: used to measure PTSD symptoms; only participants with PCL-C less than 51 were included in the study
- STOP-BANG [54]: a set of 8 yes/no questions performed to assess the participant's risk for developing sleep apnea
- Insomnia Severity Index (ISI) [47]: used to assess the subjective severity of participant's insomnia symptoms
- Pittsburgh Sleep Quality Index (PSQI) [55]: administered to assess different components of a participant's sleep quality (cutoff of 5 differentiating between good and bad sleepers)
- PSQI Addendum for PTSD (PSQI-A) [56,57]: performed to assess the frequency of disruptive nocturnal behaviors commonly experienced by trauma-exposed individuals
- Epworth Sleepiness Scale (ESS) [58]: used to assess a participant's daytime sleepiness, with higher scores indicating greater sleepiness

Outcome Measures

Because insomnia is the most prevalent sleep disorder among post–9/11 SMs and veterans [59,60], the ISI [47] was used as the primary sleep outcome metric. The ISI is a 7-item self-administered questionnaire that subjectively assesses the severity of a participant's insomnia symptoms, including level of satisfaction with sleep, noticeability and extent of daytime impairment, and additional concerns caused by sleep problems. Each item has a scale with a range from 0 to 4, with a total score of 10 or higher reflecting the presence of clinically significant insomnia [47]. Daytime sleepiness was assessed using the ESS [58], an 8-item self-report questionnaire where respondents are asked to rate, on a scale from 0 to 4, their usual chances of dozing off or falling asleep while engaged in 8 different activities. The overall ESS score can range from 0 to 24, where a score of 10 or higher indicates clinically significant somnolence.

Overall sleep quality was assessed using the PSQI [55] and PSQI-A [56,57]. The PSQI is an 18-item self-administered questionnaire that assesses different components of sleep quality with scores ranging from 0 to 21. A score of 5 or higher has been shown to reflect clinically significant sleep complaints. Disruptive nocturnal behaviors were assessed using the PSQI-A, which is a 7-item self-report measure that assesses the severity of 7 disruptive nocturnal behaviors commonly experienced by trauma-exposed individuals [56,57]. PSQI-A scores range from 0 to 21, with a score of 4 or higher indicating clinically significant disruptive nocturnal behaviors.

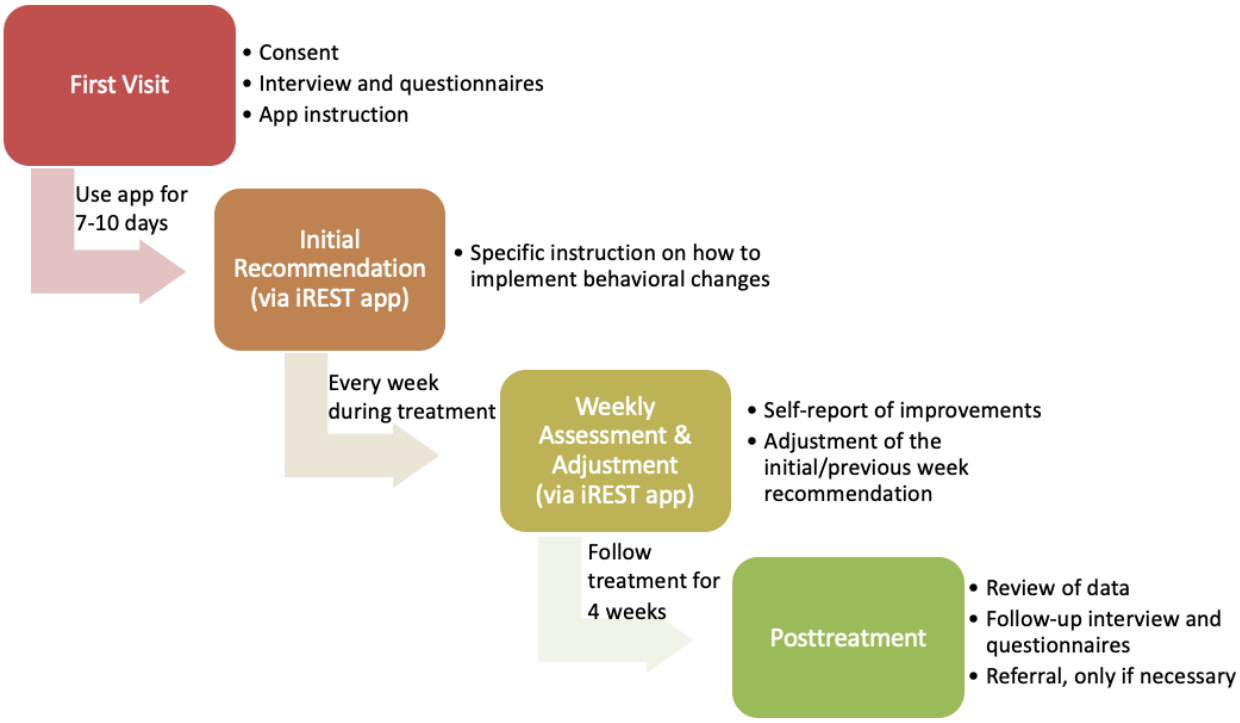
Given the common comorbidity between sleep disturbances and psychiatric symptoms, participants also completed the PCL-C [53], the Patient Health Questionnaire 9-item (PHQ-9) [61] to measure symptoms of depression, and the Generalized Anxiety Disorder 7-item (GAD-7) [62] to measure symptoms of anxiety. The PCL-C is a 17-item self-report rating scale of PTSD symptom severity, with higher scores reflecting more severe symptomatology. The PHQ-9 item assesses the frequency of 9 symptoms of depression over the preceding 2 weeks. Scores of 5, 10, 15, and 20 represent no to mild, moderate, and moderately severe depression, respectively. Finally, the GAD-7 is a brief self-report measure of symptoms of generalized anxiety.

Consistent with the previous trials [17,46], treatment response was defined as a reduction of 8 or more points on the ISI [63]. Furthermore, remission was defined as meeting the treatment response criteria and achieving a posttreatment ISI score below the clinical threshold of 7. Treatment response was also assessed with more global measures of improvements, using the PSQI (defined as a decrease of at least 3 points from pre- to posttreatment) [21,46] and the Patient- and Clinician-Rated Clinical Global Improvement Scales [64-66].

Exploratory Evaluation of Noninferiority

The second aim of the study was to compare clinical improvements in sleep and psychiatric symptoms using iREST relative to habitual, in-person delivery formats of evidence-based behavioral sleep treatments. To do so, we extracted data from 2 previously published trials [17,46]. The first trial included an 8-week treatment arm (CBTI+IRT) that combined in-person CBTI and imagery rehearsal therapy (IRT) for nightmares [46]. The second trial tested an abbreviated, 4-week CBTI protocol specifically designed for SMs and veterans [17]. In this trial, the intervention was delivered during a 45-minute session in week 1 followed by a booster telephone session in week 3 [67]. Both of these trials enrolled SMs and veterans presenting chronic, service-related sleep complaints and employed the same sleep measures (ie, ISI, PSQI, PSQI-A, and ESS). Symptoms of PTSD were assessed with the PCL-C. Symptoms of depression and anxiety were assessed with the Beck Depression Inventory [68] and the Beck Anxiety Inventory (BAI) [69].

Figure 3. The interactive Resilience Enhancing Sleep Tactics (iREST) study workflow.



Treatment Conditions

Figure 3 depicts this study's overall workflow. After providing their written informed consent, participants were issued a 6-digit ID and then asked to complete the series of self-report questionnaires aimed at assessing their military history and any demographic variables, baseline sleep quality, current sleep habits and behaviors, current psychological well-being, and overall perceived physical health. Participants also completed clinician-administered interviews to assess the presence of psychiatric disorders. After the interviews, participants downloaded the iREST app on their personal mobile phone and received instructions on how to use the app. They were instructed to complete the morning and evening sleep diary for the next 7 to 10 days, at which point they would receive their personalized sleep recommendations via the app. They were also instructed to contact their clinicians via the text messaging function or by telephone as needed. This first visit took approximately 90 minutes.

After this initial period of 7 to 10 days with the app, participants received their individualized recommendations via the app, with specific instructions on how to implement recommended behavioral changes. Each week, they completed a short battery of self-report measures to assess overall perceived improvements in sleep, side effects, and symptoms of depression, anxiety, and PTSD to monitor progress. Clinicians reviewed participant sleep information and reported symptoms and improvements via the portal. Adjustments to the initial recommendations were provided on average on a weekly basis. During the intervention phase, participants who reported an exacerbation of symptoms were scheduled for telephone or in-person visit. After the intervention phase, participants who continued to experience significant sleep complaints were offered in-person sleep consults with the clinician and/or referral to a sleep clinic or mental health services.

Statistical Analysis

Descriptive statistics were performed to describe the demographic characteristics of the study participants using frequencies for categorical variables and means and standard deviations for continuously measured demographic variables.

SPSS Statistics software version 24.0 (IBM Corp) was used to assess pre- to postintervention changes in sleep and psychiatric symptom severity. For the first aim, paired *t* tests were used to test pre- to posttreatment differences on self-reported sleep and psychiatric symptom measures. To better contextualize the magnitude of improvement, Cohen *d* effect sizes were also computed. In addition, mixed model analyses of variance (ANOVAs) were performed to explore whether the improvement in outcomes differed based on the presence or absence of comorbid disorders.

For the second aim, descriptive statistics were performed to describe the demographic characteristics of the study participants using frequencies for categorical variables and means and standard deviations for continuously measured demographic variables. Using a chi-square test for categorical variables and ANOVAs for continuous variables, each demographic variable was compared with the same variable from the previously published in-person trials [17,46] to determine whether there existed any statistically significant differences between the distribution in this study sample and the previous one. Rates of treatment response and remission across 3 delivery formats were compared using the chi-square test. Finally, a mixed model ANOVA was conducted on the primary sleep outcome (ISI) to explore on whether there were different effects between the groups.

Results

Participant Flow

A total of 111 SMs and veterans expressed interest in participating in this pilot feasibility study (Figure 4). Of these, 40 provided written consent and 33 were found eligible to complete the study. Out of the participants who started the intervention, 84% (27/33) completed the posttreatment/follow-up assessment and were included in the follow-up analysis.

Demographics

Of the 5 participants who did not finish the treatment, 3 were excluded due to poor compliance and 2 withdrew because they were no longer interested in the study. There were no significant differences between those who completed the study and those who did not. Table 1 shows the demographics and baseline scores for the 27 completers.

Pre- to Posttreatment Changes in Sleep and Psychiatric Symptoms

The pre- and postintervention tests show statistically significant improvement in primary and secondary sleep outcomes. As shown in Table 2, the mean reduction on the ISI was 9.96, t_{26}=9.99, P<.001, which reflects a decrease by at least 1 severity category on this measure. Additionally, there was marked improvement in sleep quality, with a mean reduction on the PSQI of 6.67, t_{26}=8.22, P<.001, and mean reduction on the overall severity of disruptive nocturnal disturbances on the PSQI-A of 2.37, t_{26}=3.55, P=.001. Finally, there was a decrease in daytime sleepiness, with a mean reduction on the ESS of 2.04, t_{26}=2.98, P=.006. Clinically and statistically significant improvements in symptoms of depression, anxiety, and PTSD were also detected (see Table 2 for details on these improvements).

Figure 4. Participant flow diagram.

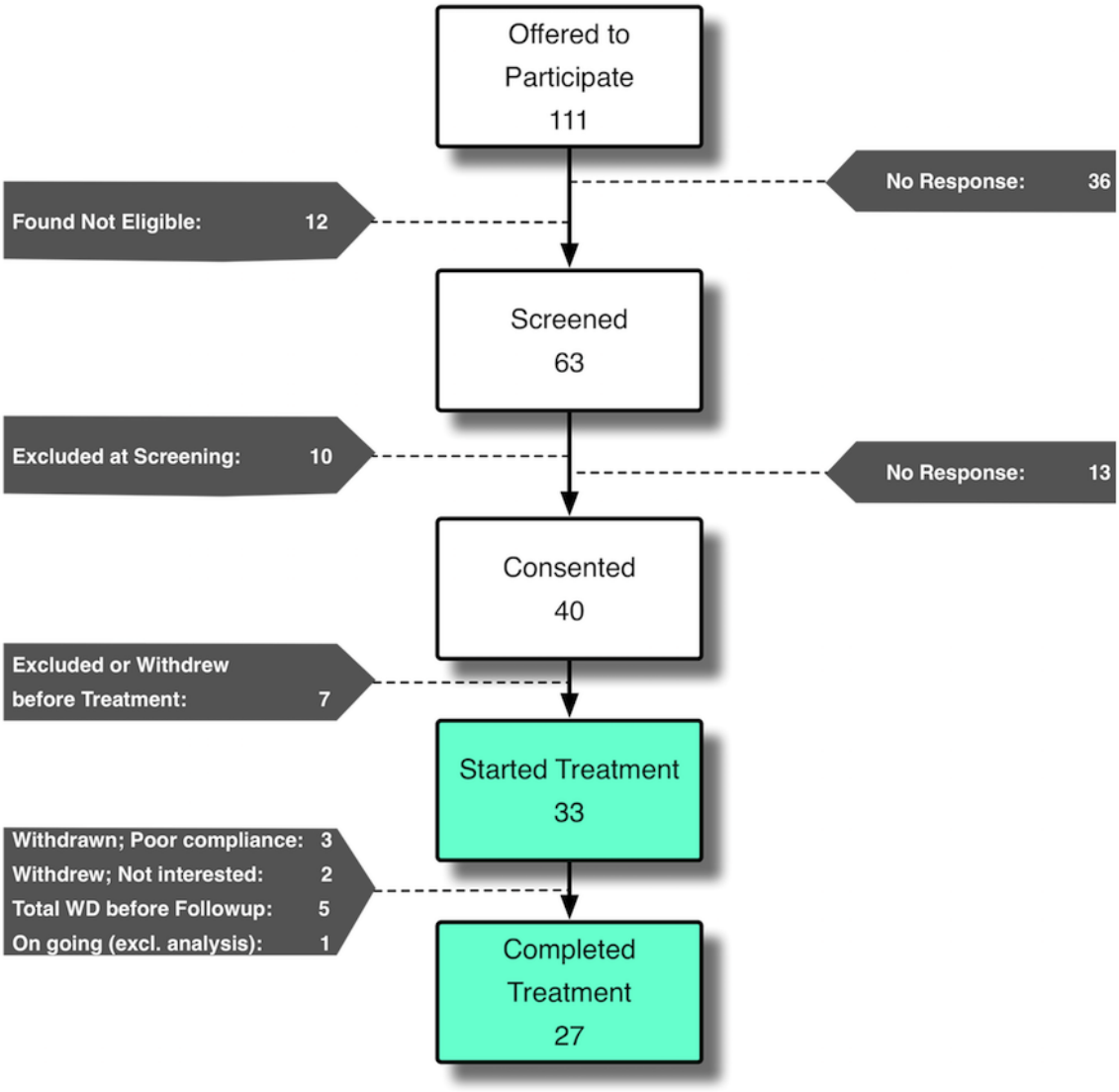


Table 1. Participant demographics and baseline scores (N=27).

Variable	Value
Male, n (%)	24 (89)
White, n (%)	20 (74)
Age (years), mean (SD)	36.48 (9.64)
Army, n (%)	15 (56)
Current posttraumatic stress disorder, n (%)	8 (30)
Using psychotropic medications, n (%)	7 (26)
Current mood or anxiety disorder, n (%)	10 (37)

Table 2. Mean score changes pre- and postintervention.

Variable	Baseline score, mean (SD)	Posttreatment score, mean (SD)	Mean change (SE)	*t* statistics (*df*)	*P* value	Cohen *d* effect sizes
ISI[a]	15.59 (4.13)	5.63 (4.76)	9.96 (1.00)	9.99 (26)	<.001	1.93
ESS[b]	7.07 (4.51)	5.04 (3.82)	2.04 (0.69)	2.98 (26)	.006	0.57
PSQI[c]	11.81 (3.19)	5.15 (3.43)	6.67 (0.81)	8.22 (26)	<.001	1.58
PSQI-A[d]	4.59 (3.83)	2.22 (2.81)	2.37 (0.67)	3.55 (26)	.001	0.71
PCL-C[e,f,g]	38.41 (14.10)	27.22 (11.87)	11.19 (1.86)	6.58 (26)	<.001	1.19
PHQ-9[h]	8.41 (5.22)	3.63 (5.34)	4.78 (0.81)	5.89 (26)	<.001	1.13
GAD-7[i]	6.17 (5.32)	2.91 (2.94)	3.26 (0.90)	3.64 (22)	.001	0.74

[a]ISI: Insomnia Severity Index.

[b]ESS: Epworth Sleepiness Scale.

[c]PSQI: Pittsburgh Sleep Quality Index.

[d]PSQI-A: Pittsburgh Sleep Quality Index–Addendum for Posttraumatic Stress Disorder.

[e]PCL-C: Posttraumatic Stress Disorder Checklist–Civilian.

[f]PCL-C scores were not normally distributed and a natural log transformation was used in the analyses.

[g]Raw scores are presented.

[h]PHQ-9: Patient Health Questionnaire 9-itm.

[i]GAD-7: Generalized Anxiety Disorder 7-item.

Table 3. Insomnia improvement grouped by comorbidity diagnoses.

Grouping variable and effect	*F*	*P* value
Posttraumatic stress disorder diagnosis		
Time	84.50	<.001
Group	0.22	.64
Time × group	0.25	.62
Mood and anxiety diagnosis		
Time	87.86	<.001
Group	3.01	.09
Time × group	0.25	.62

Rates of Treatment Response and Remission With the Interactive Resilience Enhancing Sleep Tactics App

Using the global measures of clinical improvement, 74% (20/27) of participants reported that they were much or very much improved posttreatment, whereas clinicians rated 82% (22/27) of participants much or very much improved posttreatment. Using the criterion of a decrease of at least 3 points on the PSQI, 85% (23/27) of patients showed improvements in global sleep quality and, of those, 19 achieved full remission, defined as a posttreatment PSQI score of less than 5. Using the stringent definition of treatment response of a reduction by 8 points or more on the ISI, 70% (19/27) of participants met the criterion for treatment response and 16 presented full remission of

insomnia (ie, ISI score less than 7 posttreatment—59% (16/27) of the full sample and 84% (16/19) of responders).

Exploratory Assessment of Noninferiority of the Interactive Resilience Enhancing Sleep Tactics App Relative to Standard and Abbreviated Cognitive Behavioral Therapy for Insomnia

As shown in Table 4, there were no statistically significant demographic differences between participants in this study (iREST) and those in the traditional trials used as a control. Furthermore, participants for the iREST study and prior traditional studies were drawn from the same geographical area. For all 3 studies, participants had to be able to attend an in-person assessment at the University of Pittsburgh Military Sleep Tactics and Resilience Research Team office.

Mixed model ANOVAs were also performed within-group to explore whether the improvement in outcomes differed based on the presence or absence of comorbid PTSD and mood and anxiety disorders. After controlling for baseline ISI scores, the only significant effect observed was expressed in terms of time (pre- to post-). No significant effect from comorbidity conditions or any significant interactions between time and these conditions was observed. See Table 3 for additional results.

On the ISI, the CBTI+IRT, brief CBTI, and iREST yielded large and clinically significant improvements, with Cohen *d* effect sizes of *d*=1.50, *d*=1.96, and *d*=1.93, respectively (shown in Figure 5). Additionally, significant improvements in sleep quality, or reductions in PSQI score, were also observed in all 3 groups, with *d*=1.45 in the CBTI+IRT group, *d*=1.56 in the brief CBTI group, and *d*=1.58 in the iREST group. There were also significant improvements in PSQI-A scores for the CBTI+IRT and the iREST groups, with *d*=0.87 and *d*=0.71, respectively.

PTSD symptoms were similarly reduced in the CBTI+IRT and iREST groups pre- to posttreatment, with Cohen *d* effect sizes of *d*=1.08 and *d*=1.19, respectively; a lower Cohen *d* effect size of *d*=0.20 was reported on the brief CBTI group. Depression symptom severity was also significantly reduced in all groups, with *d*=0.65 in the CBTI+IRT group and *d*=0.69 in the brief CBTI group compared with *d*=1.13 in the iREST group. For symptoms of anxiety measured with the BAI, pre- to posttreatment changes in both the CBTI+IRT and brief CBTI groups were nonsignificant (*d*=0.08 and *d*=0.14, respectively), whereas the pre- to posttreatment changes in symptoms of generalized anxiety as measured by the GAD-7 were significantly more pronounced in the iREST group (*d*=0.89).

Furthermore, a mixed model ANOVA conducted before and after treatment, with time functioning as a within-subject repeated measure on the primary clinical outcome (ISI), showed no significant group × time interaction ($F_{2,53}$=0.36, *P*=.70) and no main effect of group (iREST vs brief CBTI vs CBTI+IRT; $F_{2,53}$=1.02, *P*=.37). Instead, only a main effect of time was detected ($F_{1,53}$= 140.5, *P*<.001). This further suggests that iREST may be noninferior to the in-person brief CBTI and standard CBTI.

Table 4. Demographic and clinical information at baseline compared with in-person standard (8 weeks) [46] and brief (4 weeks) cognitive behavioral therapy for insomnia trials in military samples [17].

Characteristics	iREST[a] (n=27)	CBTI[b] + IRT[c] (n=17)	Brief CBTI (n=20)	Statistics	
				χ^2	$F_{2,61}$
Variable					
Male, n (%)	24 (88.9)	14 (88.9)	19 (95)	1.51	—
White, n (%)	20 (74.1)	12 (70.6)	14 (70)	0.11	—
Age (years), mean (SD)	36.48 (9.6)	40.0 (14.1)	40.9 (12.0)	—	0.94
Army, n (%)	15 (55.6)	NR[d]	16 (80)	3.06	
Current posttraumatic stress disorder, n (%)	8 (29.6)	7 (41.2)	4 (20)	1.18	—
Using psychotropic medications, n (%)	7 (25.9)	6 (35.3)	5 (25)	0.59	—
Current mood or anxiety disorder, n (%)	10 (37.0)	2 (11.8)	2 (10)	4.14	—
Baseline sleep assessment, mean (SD)					
Epworth Sleepiness Scale[e]	7.4 (4.6)	NR	7.3 (4.4)	—	—
Insomnia Severity Index	17.4 (4.0)	16.5 (4.0)	16.3 (3.9)	—	0.52
Pittsburgh Sleep Quality Index	11.9 (3.9)	10.3 (2.9)	11.3 (3.5)	—	2.14

[a]iREST: interactive Resilience Enhancing Sleep Tactics.

[b]CBTI: cognitive behavioral therapy for insomnia.

[c]IRT: imagery rehearsal therapy.

[d]NR: value for this category was not reported on the CBTI+IRT study.

[e]t_{46}=0.08

Figure 5. Reductions in the Insomnia Severity Index from baseline to posttreatment with cognitive behavioral therapy for insomnia (CBTI; 8 in-person visits over 8 weeks), brief CBTI (2 in-person visits over 4 weeks), and the interactive Resilience Enhancing Sleep Tactics (iREST) app (visits=interventions through the app over 4 weeks).

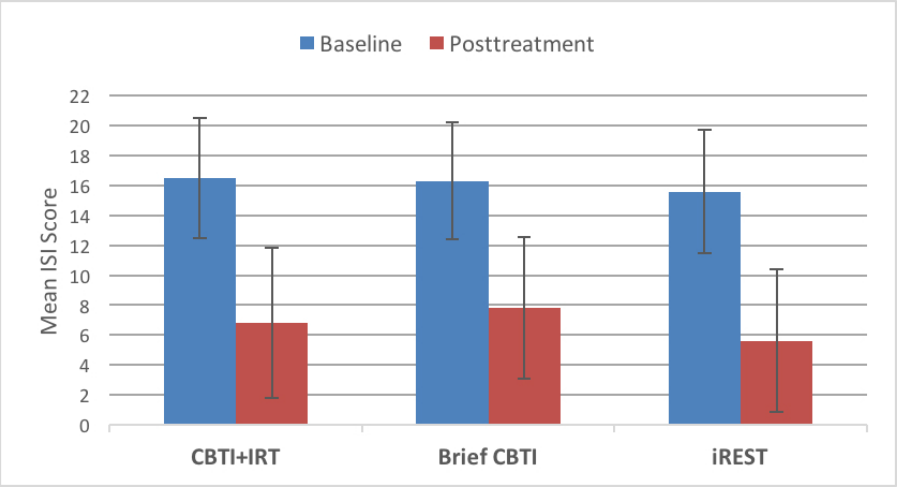
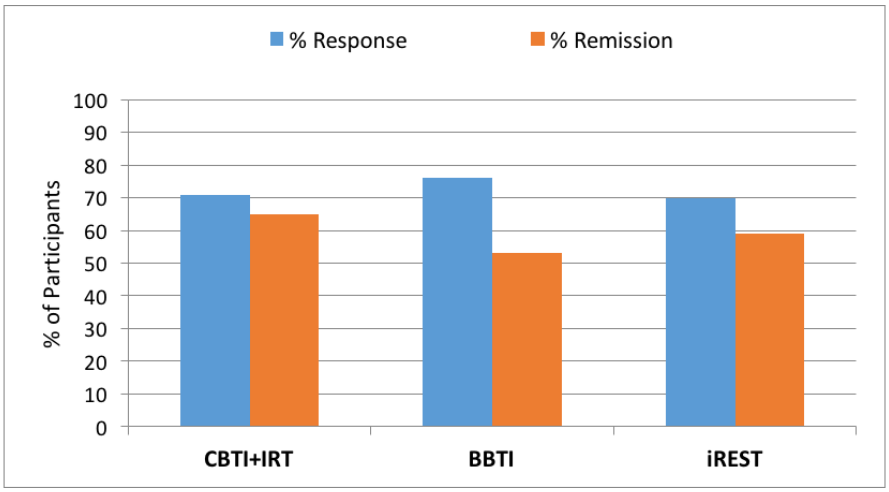


Figure 6. Comparison of the interactive Resilience Enhancing Sleep Tactics (iREST) app and traditional intervention remission and treatment response rates. BBTI: brief behavioral therapy for insomnia; CBTI: cognitive behavioral therapy for insomnia; IRT: imagery rehearsal therapy.



The response rate for the previous in-person brief CBTI study was 76.47% with a remission rate of 52.94%, while the rates for the CBTI+IRT trial were 70.59% and 64.70%, respectively. No significant difference exists between the rates in our trial and the previous in-person brief and standard CBTI studies (illustrated in Figure 6). Here the chi-square values are χ^2=0.22, *P*=.90, and χ^2=0.49, *P*=.78, for response and remission, respectively.

Discussion

Principal Findings

The purpose of this pilot study was to evaluate clinical feasibility and potential benefits of a novel JITAI app (iREST) for the delivery of evidence-based behavioral sleep treatments. In this study, clinically significant improvements in insomnia, general sleep quality, and disruptive nocturnal behaviors were detected pre- to posttreatment with iREST. Clinically meaningful improvements in symptoms of PTSD, depression, and anxiety were also detected. These findings suggest that iREST, a novel mobile health (mHealth) platform, has a high potential for augmenting the current inventory of evidence-based recommended behavioral sleep treatments.

To explore the potential noninferiority of iREST, we compared clinical outcomes observed with iREST in this study to improvements we observed in previous randomized clinical trials with standard or abbreviated in-person delivery of CBTI [17,46]. Consistent with previous JITAI studies [45,70-73], we found that the implementation of personalized and time-varying JITAI approaches with iREST yielded noninferior outcomes with regard to insomnia or measures of sleep quality and disruptive nocturnal behaviors. The rates of treatment response and remission were also comparable to previously reported rates in CBTI trials [17,46,74-76]. Finally, this exploratory comparison suggested that the magnitude of improvements detected for psychiatric symptoms of PTSD and depression were noninferior with IREST as previously detected in clinical trials. Although improvements in symptoms of anxiety seem to be superior with iREST relative to the standard in-person treatments, the different measures used across trials warrant caution.

The use of a mobile phone app and clinician Web portal, features that allow for real-time monitoring and delivery of personalized treatment prescriptions (eg, bedtime reminders, wake-up alarms, appropriate bibliotherapies, and additional assessments) matching the needs of the individual, contributed to the promising results of this study. Furthermore, delivering the behavioral sleep intervention digitally (through mHealth/mobile app) reduced or eliminated the costs associated with an in-person visit to a sleep clinic (eg, loss of wages and cost for travel, accommodations, and child care). It also potentially addressed access and scalability barriers; through the personalization and prioritization embedded in a JITAI-based system, clinicians could optimize their service so it would reach a greater number of patients while maintaining the same level of care. Last, digital interventions such as JITAI allow seamless integration of patients' clinical progress and outcomes into the medical center or clinic's electronic health record system and each patient's own personal health record. This integration is important in maintaining the continuity of care, especially since insomnia is highly comorbid with other health and psychological conditions [6,77,78].

Limitations

Inherent to the pilot nature of the study, a first limitation is the relatively small sample size. Therefore, the effect sizes detected in this sample, albeit moderate to large, may be attenuated in a larger, confirmatory noninferiority clinical trial. A second limitation relates to the exclusion of individuals with severe psychiatric disorders or sleep apnea. The high rate of exclusion in this study and previous clinical trials highlights the fact that these disorders are highly prevalent among SMs and veterans [59,60] and hence limit the generalizability of the findings to the more severely affected populations. In a related manner, the inclusion of military SMs and veterans may limit the generalizability of the findings to the general civilian population. Future studies should include a wider set of participants and narrower exclusion criteria to assess the effectiveness and generalizability of the treatment for patients with comorbidities.

Conclusions

In this preliminary study, iREST, a novel JITAI app, was associated with statistically significant and clinically meaningful improvements in sleep and psychiatric symptoms in a sample of SMs and veterans with chronic, service-related insomnia. Exploratory comparisons strongly suggest that iREST is noninferior to the traditional in-person delivery formats for CBTI for sleep and related psychiatric symptoms. Together, these findings support the notion that iREST and the JITAI approach can be an acceptable and effective approach to enhance the scalability of evidence-based behavioral sleep treatments. Larger confirmatory noninferiority trials are needed in order to fully understand the effectiveness of JITAI-based iREST among military and civilian populations.

Acknowledgments

We wish to thank the SMs and veterans who participated in the study. We also wish to acknowledge the research team members that helped coordinate and run the study including Hassen Khan, Noelle Rode, Mark Jones, Robin Richardson, and Megan Wolfson. We would also like to thank Justin Stec for his contribution in the editing of this manuscript. This project was supported in part by grants from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant numbers 90RE5018 and 90DP0064). NIDILRR is a center within the Administration for Community Living, Department of Health and Human Services.

Conflicts of Interest

BP and AG own equity in Rehat, LLC. IWP, BP, and AG receive royalties for iREST. AG has served as a consultant for Jazz Pharmaceuticals, Inc.

References

1. Epstein DR, Babcock-Parziale JL, Haynes PL, Herb CA. Insomnia treatment acceptability and preferences of male Iraq and Afghanistan combat veterans and their healthcare providers. J Rehabil Res Dev 2012;49(6):867-878 [FREE Full text] [Medline: 23299258]
2. Hoge C, Terhakopian A, Castro CA, Messer SC, Engel CC. Association of posttraumatic stress disorder with somatic symptoms, health care visits, and absenteeism among Iraq war veterans. Am J Psychiatry 2007 Jan;164(1):150-153. [doi: 10.1176/ajp.2007.164.1.150] [Medline: 17202557]
3. Ocasio-Tascón ME, Alicea-Colón E, Torres-Palacios A, Rodríguez-Cintrón W. The veteran population: one at high risk for sleep-disordered breathing. Sleep Breath 2006 Jun;10(2):70-75. [doi: 10.1007/s11325-005-0043-9] [Medline: 16496117]
4. Livingston WS, Rusch HL, Nersesian PV, Baxter T, Mysliwiec V, Gill JM. Improved sleep in military personnel is associated with changes in the expression of inflammatory genes and improvement in depression symptoms. Front Psychiatry 2015 Apr;6:59 [FREE Full text] [doi: 10.3389/fpsyt.2015.00059] [Medline: 25983695]
5. Haslam DR. Sleep loss, recovery sleep, and military performance. Ergonomics 1982 Feb;25(2):163-178. [doi: 10.1080/00140138208924935] [Medline: 7084209]
6. Chakravorty S, Grandner MA, Mavandadi S, Perlis ML, Sturgis EB, Oslin DW. Suicidal ideation in veterans misusing alcohol: relationships with insomnia symptoms and sleep duration. Addict Behav 2014 Feb;39(2):399-405 [FREE Full text] [doi: 10.1016/j.addbeh.2013.09.022] [Medline: 24169371]

7. Raskind MA, Peskind ER, Hoff DJ, Hart KL, Holmes HA, Warren D, et al. A parallel group placebo controlled study of prazosin for trauma nightmares and sleep disturbance in combat veterans with post-traumatic stress disorder. Biol Psychiatry 2007 Apr 15;61(8):928-934. [doi: 10.1016/j.biopsych.2006.06.032] [Medline: 17069768]
8. Maher MJ, Rego SA, Asnis GM. Sleep disturbances in patients with post-traumatic stress disorder: epidemiology, impact and approaches to management. CNS Drugs 2006;20(7):567-590. [Medline: 16800716]
9. Gironda RJ, Clark ME, Ruff RL, Chait S, Craine M, Walker R, et al. Traumatic brain injury, polytrauma, and pain: challenges and treatment strategies for the polytrauma rehabilitation. Rehabil Psychol 2009 Aug;54(3):247-258. [doi: 10.1037/a0016906] [Medline: 19702423]
10. National Institutes of Health. NIH State-of-the-Science Conference statement on manifestations and management of chronic insomnia in adults. Sleep 2005 Sep;28(9):1049-1057. [Medline: 16268373]
11. Qaseem A, Kansagara D, Forciea MA, Cooke M, Denberg TD, Clinical Guidelines Committee of the American College of Physicians. Management of chronic insomnia disorder in adults: a clinical practice guideline From the American College of Physicians. Ann Intern Med 2016 Jul 19;165(2):125-133. [doi: 10.7326/M15-2175] [Medline: 27136449]
12. Edinger J, Carney C. Overcoming Insomnia: A Cognitive-Behavioral Therapy Approach, Therapist Guide. Oxford: Oxford University Press; 2014.
13. Morin C. Insomnia: Psychological Assessment and Management. New York: Guilford Press; 1993.
14. Perlis M, Aloia M, Millikan A, Boehmler J, Smith M, Greenblatt D, et al. Behavioral treatment of insomnia: a clinical case series study. J Behav Med 2000 Apr;23(2):149-161. [Medline: 10833677]
15. Germain A, Buysse D. Brief behavioral treatment of insomnia. In: Pertis ML, Aloia M, Kuhn B, editors. Behavioral Treatments for Sleep Disorders: A Comprehensive Primer of Behavioral Sleep Medicine Interventions (Practical Resources for the Mental Health Professional). Burlington: Academic Press; 2011:143-150.
16. Trockel M, Karlin BE, Taylor CB, Manber R. Cognitive behavioral therapy for insomnia with veterans: evaluation of effectiveness and correlates of treatment outcomes. Behav Res Ther 2014 Feb;53:41-46. [doi: 10.1016/j.brat.2013.11.006] [Medline: 24412462]
17. Germain A, Richardson R, Stocker R, Mammen O, Hall M, Bramoweth AD, et al. Treatment for insomnia in combat-exposed OEF/OIF/OND military veterans: preliminary randomized controlled trial. Behav Res Ther 2014 Oct;61:78-88 [FREE Full text] [doi: 10.1016/j.brat.2014.07.016] [Medline: 25194223]
18. 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] [Medline: 22654196]
19. Edinger JD, Wohlgemuth WK, Radtke RA, Marsh GR, Quillian RE. Cognitive behavioral therapy for treatment of chronic primary insomnia: a randomized controlled trial. JAMA 2001 Apr 11;285(14):1856-1864. [Medline: 11308399]
20. 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]
21. Buysse DJ, Germain A, Moul DE, Franzen PL, Brar LK, Fletcher ME, et al. Efficacy of brief behavioral treatment for chronic insomnia in older adults. Arch Intern Med 2011 May 23;171(10):887-895 [FREE Full text] [doi: 10.1001/archinternmed.2010.535] [Medline: 21263078]
22. Taylor DJ, Peterson AL, Pruiksma KE, Hale WJ, Young-McCaughan S, Wilkerson A, STRONG STAR Consortium. Impact of cognitive behavioral therapy for insomnia disorder on sleep and comorbid symptoms in military personnel: a randomized clinical trial. Sleep 2018 Jun 01;41(6):1. [doi: 10.1093/sleep/zsy069] [Medline: 29618098]
23. Phelps AJ, Varker T, Metcalf O, Dell L. What are effective psychological interventions for veterans with sleep disturbances? A rapid evidence assessment. Mil Med 2017 Dec;182(1):e1541-e1550. [doi: 10.7205/MILMED-D-16-00010] [Medline: 28051971]
24. Capaldi VF, Kim JR, Grillakis AA, Taylor MR, York CM. Insomnia in the military: application and effectiveness of cognitive and pharmacologic therapies. Curr Psychiatry Rep 2015 Oct;17(10):85. [doi: 10.1007/s11920-015-0622-9] [Medline: 26364060]
25. Taylor DJ, Peterson AL, Pruiksma KE, Young-McCaughan S, Nicholson K, Mintz J, STRONG STAR Consortium. Internet and in-person cognitive behavioral therapy for insomnia in military personnel: a randomized clinical trial. Sleep 2017 Dec 01;40(6):1. [doi: 10.1093/sleep/zsx075] [Medline: 28472528]
26. US Department of Veterans Affairs. Office of Rural Health: rural veterans. URL: https://www.ruralhealth.va.gov/aboutus/ruralvets.asp [accessed 2018-10-22] [WebCite Cache ID 73MZlwg8P]
27. Buysse DJ, Germain A, Moul DE, Franzen PL, Brar LK, Fletcher ME, et al. Efficacy of brief behavioral treatment for chronic insomnia in older adults. Arch Intern Med 2011 May 23;171(10):887-895 [FREE Full text] [doi: 10.1001/archinternmed.2010.535] [Medline: 21263078]
28. Edinger JD, Sampson WS. A primary care “friendly” cognitive behavioral insomnia therapy. Sleep 2003 Mar 15;26(2):177-182. [Medline: 12683477]
29. Edinger JD, Wohlgemuth WK, Radtke RA, Coffman CJ, Carney CE. Dose-response effects of cognitive-behavioral insomnia therapy: a randomized clinical trial. Sleep 2007 Feb;30(2):203-212. [Medline: 17326546]

30. Germain A, Moul DE, Franzen PL, Miewald JM, Reynolds CF, Monk TH, et al. Effects of a brief behavioral treatment for late-life insomnia: preliminary findings. J Clin Sleep Med 2006 Oct 15;2(4):403-406. [Medline: 17557467]
31. Germain A, Shear MK, Hall M, Buysse DJ. Effects of a brief behavioral treatment for PTSD-related sleep disturbances: a pilot study. Behav Res Ther 2007 Mar;45(3):627-632. [doi: 10.1016/j.brat.2006.04.009] [Medline: 16777060]
32. Gehrman P, Shah MT, Miles A, Kuna S, Godleski L. Feasibility of group cognitive-behavioral treatment of insomnia delivered by clinical video telehealth. Telemed J E Health 2016 Dec;22(12):1041-1046. [doi: 10.1089/tmj.2016.0032] [Medline: 27286369]
33. 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] [Medline: 19581560]
34. Ritterband LM, Thorndike FP, Ingersoll KS, Lord HR, Gonder-Frederick L, Frederick C, et al. Effect of a web-based cognitive behavior therapy for insomnia intervention with 1-year follow-up: a randomized clinical trial. JAMA Psychiatry 2017 Jan 01;74(1):68-75. [doi: 10.1001/jamapsychiatry.2016.3249] [Medline: 27902836]
35. Espie CA, Hames P, McKinstry B. Use of the internet and mobile media for delivery of cognitive behavioral insomnia therapy. Sleep Medicine Clinics 2013 Sep;8(3):407-419. [doi: 10.1016/j.jsmc.2013.06.001]
36. Holmqvist M, Vincent N, Walsh K. Web- vs. telehealth-based delivery of cognitive behavioral therapy for insomnia: a randomized controlled trial. Sleep Med 2014 Feb;15(2):187-195. [doi: 10.1016/j.sleep.2013.10.013] [Medline: 24461370]
37. Hagatun S, Vedaa O, Nordgreen T, Smith ORF, Pallesen S, Havik OE, et al. The short-term efficacy of an unguided internet-based cognitive-behavioral therapy for insomnia: a randomized controlled trial with a six-month nonrandomized follow-up. Behav Sleep Med 2017 Mar 27:1-23. [doi: 10.1080/15402002.2017.1301941] [Medline: 28345961]
38. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
39. Spielman AJ, Saskin P, Thorpy MJ. Treatment of chronic insomnia by restriction of time in bed. Sleep 1987 Feb;10(1):45-56. [Medline: 3563247]
40. Bootzin R. Stimulus control treatment for insomnia. Proc Am Psychol Assoc 1972 Sep 02;7:395-396. [doi: 10.1037/e465522008-198]
41. Intille SS, Kukla C, Farzanfar R, Bakr W. Just-in-time technology to encourage incremental, dietary behavior change. AMIA Annu Symp Proc 2003:874 [FREE Full text] [Medline: 14728379]
42. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2016 Sep 23:1. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
43. Pulantara IW, Parmanto B, Germain A. Development of a Just-in-Time adaptive mHealth intervention for insomnia: usability study. JMIR Hum Factors 2018 May 17;5(2):e21 [FREE Full text] [doi: 10.2196/humanfactors.8905] [Medline: 29773529]
44. Rabbi M, Aung M, Zhang M. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. 2015 Presented at: UbiComp '15 Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2015; Osaka. [doi: 10.1145/2750858.2805840]
45. 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] [Medline: 23638127]
46. Germain A, Richardson R, Moul DE, Mammen O, Haas G, Forman SD, et al. Placebo-controlled comparison of prazosin and cognitive-behavioral treatments for sleep disturbances in US Military Veterans. J Psychosom Res 2012 Feb;72(2):89-96 [FREE Full text] [doi: 10.1016/j.jpsychores.2011.11.010] [Medline: 22281448]
47. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [Medline: 11438246]
48. Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. Br J Anaesth 2012 May;108(5):768-775 [FREE Full text] [doi: 10.1093/bja/aes022] [Medline: 22401881]
49. Silva GE, Vana KD, Goodwin JL, Sherrill DL, Quan SF. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. J Clin Sleep Med 2011 Oct 15;7(5):467-472 [FREE Full text] [doi: 10.5664/JCSM.1308] [Medline: 22003341]
50. Farney RJ, Walker BS, Farney RM, Snow GL, Walker JM. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: relation to polysomnographic measurements of the apnea/hypopnea index. J Clin Sleep Med 2011 Oct 15;7(5):459-465 [FREE Full text] [doi: 10.5664/JCSM.1306] [Medline: 22003340]
51. First ME, Spitzer R, Gibbon M. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Nonpatient Edition. Washington: American Psychiatric Press; 2002.
52. American Academy of Sleep Medicine. The International Classification of Sleep Disorders: Diagnostic and Coding Manual, 3rd Edition. Darien: American Academy of Sleep Medicine; 2005.

53. Weathers F, Litz B, Huska J, Keane T. PTSD Checklist—Civilian version. 1994. URL: https://www.mirecc.va.gov/docs/visn6/3_ptsd_checklist_and_scoring.pdf [accessed 2018-10-31] [WebCite Cache ID 73ZScpeTk]
54. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology 2008 May;108(5):812-821. [doi: 10.1097/ALN.0b013e31816d83e4] [Medline: 18431116]
55. 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]
56. Insana SP, Hall M, Buysse DJ, Germain A. Validation of the Pittsburgh Sleep Quality Index Addendum for posttraumatic stress disorder (PSQI-A) in U.S. male military veterans. J Trauma Stress 2013 Apr;26(2):192-200 [FREE Full text] [doi: 10.1002/jts.21793] [Medline: 23512653]
57. Germain A, Hall M, Krakow B, Katherine Shear M, Buysse DJ. A brief sleep scale for Posttraumatic Stress Disorder: Pittsburgh Sleep Quality Index Addendum for PTSD. J Anxiety Disord 2005;19(2):233-244. [doi: 10.1016/j.janxdis.2004.02.001] [Medline: 15533706]
58. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [Medline: 1798888]
59. Mysliwiec V, Gill J, Lee H, Baxter T, Pierce R, Barr TL, et al. Sleep disorders in US military personnel: a high rate of comorbid insomnia and obstructive sleep apnea. Chest 2013 Aug;144(2):549-557 [FREE Full text] [doi: 10.1378/chest.13-0088] [Medline: 23681455]
60. Capaldi VF, Kim JR, Grillakis AA, Taylor MR, York CM. Insomnia in the military: application and effectiveness of cognitive and pharmacologic therapies. Curr Psychiatry Rep 2015 Oct;17(10):85. [doi: 10.1007/s11920-015-0622-9] [Medline: 26364060]
61. 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]
62. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
63. 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]
64. Guy W. ECDEU Assessment Manual for Psychopharmacology: Revised (DHEW publication no ADM 76-338). US Department of Health, Education and Welfare: Rockville; 1976.
65. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. Psychiatry (Edgmont) 2007 Jul;4(7):28-37 [FREE Full text] [Medline: 20526405]
66. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. Am J Obstet Gynecol 2003 Jul;189(1):98-101. [Medline: 12861145]
67. Troxel WM, Germain A, Buysse DJ. Clinical management of insomnia with brief behavioral treatment (BBTI). Behav Sleep Med 2012 Oct;10(4):266-279 [FREE Full text] [doi: 10.1080/15402002.2011.607200] [Medline: 22946736]
68. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961;4(6):561-571. [doi: 10.1001/archpsyc.1961.01710120031004]
69. 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]
70. Witkiewitz K, Desai SA, Bowen S, Leigh BC, Kirouac M, Larimer ME. Development and evaluation of a mobile intervention for heavy drinking and smoking among college students. Psychol Addict Behav 2014 Sep;28(3):639-650. [doi: 10.1037/a0034747] [Medline: 25000269]
71. Consolvo S, McDonald D. Activity sensing in the wild: a field trial of ubifit garden. 2008 Presented at: Proceedings of the SIGCHI conference on human factors in computing systems; 2008; Florence p. 1797-1806. [doi: 10.1145/1357054.1357335]
72. Saeb S, Lattie EG, Kording KP, Mohr DC. Mobile phone detection of semantic location and its relationship to depression and anxiety. JMIR Mhealth Uhealth 2017 Aug 10;5(8):e112 [FREE Full text] [doi: 10.2196/mhealth.7297] [Medline: 28798010]
73. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A context-sensing mobile phone app (Q Sense) for smoking cessation: a mixed-methods study. JMIR Mhealth Uhealth 2016 Sep 16;4(3):e106 [FREE Full text] [doi: 10.2196/mhealth.5787] [Medline: 27637405]
74. Ulmer CS, Edinger JD, Calhoun PS. A multi-component cognitive-behavioral intervention for sleep disturbance in veterans with PTSD: a pilot study. J Clin Sleep Med 2011 Feb 15;7(1):57-68 [FREE Full text] [Medline: 21344046]
75. Ong JC, Kuo TF, Manber R. Who is at risk for dropout from group cognitive-behavior therapy for insomnia? J Psychosom Res 2008 Apr;64(4):419-425 [FREE Full text] [doi: 10.1016/j.jpsychores.2007.10.009] [Medline: 18374742]
76. Edinger JD, Olsen MK, Stechuchak KM, Means MK, Lineberger MD, Kirby A, et al. Cognitive behavioral therapy for patients with primary insomnia or insomnia associated predominantly with mixed psychiatric disorders: a randomized clinical trial. Sleep 2009 Apr;32(4):499-510 [FREE Full text] [Medline: 19413144]

77. Cappuccio FP, Stranges S, Kandala N, Miller MA, Taggart FM, Kumari M, et al. Gender-specific associations of short sleep duration with prevalent and incident hypertension: the Whitehall II Study. Hypertension 2007 Oct;50(4):693-700 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.107.095471] [Medline: 17785629]
78. Buysse DJ, Angst J, Gamma A, Ajdacic V, Eich D, Rössler W. Prevalence, course, and comorbidity of insomnia and depression in young adults. Sleep 2008 Apr;31(4):473-480 [FREE Full text] [Medline: 18457234]

Abbreviations

ANOVA: analysis of variance
BAI: Beck Anxiety Inventory
CBTI: cognitive behavioral therapy for insomnia
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
ESS: Epworth Sleepiness Scale
GAD-7: Generalized Anxiety Disorder 7-item
iREST: interactive Resilience Enhancing Sleep Tactics
IRT: imagery rehearsal therapy
ISI: Insomnia Severity Index
JITAI: just-in-time adaptive intervention
mHealth: mobile health
PCL-C: Posttraumatic Stress Disorder Checklist–Civilian
PHQ-9: Patient Health Questionnaire 9-item
PSQI: Pittsburgh Sleep Quality Index
PSQI-A: Pittsburgh Sleep Quality Index–Addendum for Posttraumatic Stress Disorder
PTSD: posttraumatic stress disorder
SM: service member
STOP-BANG: snoring, tired, observed apneas, blood pressure, body mass index, age, neck size, and gender

Edited by G Wadley, R Calvo, M Czerwinski, J Torous; submitted 15.02.18; peer-reviewed by C Stepnowsky, K Stawarz, A Heraz; comments to author 26.03.18; revised version received 22.08.18; accepted 10.09.18; published 07.12.18.

Please cite as:
Pulantara IW, Parmanto B, Germain A
Clinical Feasibility of a Just-in-Time Adaptive Intervention App (iREST) as a Behavioral Sleep Treatment in a Military Population: Feasibility Comparative Effectiveness Study
J Med Internet Res 2018;20(12):e10124
URL: https://www.jmir.org/2018/12/e10124/
doi:10.2196/10124
PMID:30530452

Corrigenda and Addenda

Multimedia Appendix Removal and Editorial Warning Regarding the MMAS Scale (Two-Way Social Media Messaging in Post-Operative Cataract Surgical Patients: A Prospective Interventional Study)

Thuss Sanguansak[1], MD; Katharine E Morley[2], MD, MPH; Michael G Morley[3,4], ScM, MD; Kavin Thinkhamrop[4], DrPH; Jaruwan Thuanman[4], MPH; Isha Agarwal[5,6], ScD, MD

[1]Department of Ophthalmology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[2]Department of Medicine, Massachusetts General Hospital, Boston, MA, United States

[3]Harvard Medical School, Boston, MA, United States

[4]Ophthalmic Consultants of Boston, Department of Ophthalmology, Harvard Medical School, Boston, MA, United States

[5]Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University, Khon Kaen, Thailand

[6]Harvard School of Public Health, Boston, MA, United States

Corresponding Author:
Katharine E Morley, MD, MPH
Massachusetts General Hospital
Department of Medicine
Harvard University
15 Parkman St
Boston, MA, 02114
United States
Phone: 1 617 726 2707
Fax: 1 617 724 0189
Email: kemorley@mgh.harvard.edu

Related Article:

Correction of: http://www.jmir.org/2017/12/e413/

(J Med Internet Res 2018;20(12):e12120) doi:10.2196/12120

Authors' Correction

The authors of "Two-Way Social Media Messaging in Postoperative Cataract Surgical Patients: Prospective Interventional Study" (J Med Internet Res 2017;19(12):e413) advise that in order to address concerns raised by the purported holders of intellectual property interests in the Morisky Medication Adherence Scale ("MMAS"), Multimedia Appendix 7 (Morisky Medication Adherence Scale-8 Item questionnaire for postoperative cataract surgery) should be removed from the article. As a result, Appendix 7 was removed and all subsequent appendices 8-10 were renumbered.

The correction will appear in the online version of the paper on the JMIR website on December 10, 2018, together with the publication of this correction notice. Because this was made after submission to PubMed, Pubmed Central, and other full-text repositories, the corrected article also has been re-submitted to those repositories.

Editorial Notice

This is the second correction we have to publish due to the actions by Steven Trubow and Donald Morisky from the company MMAS Research LLC, the copyright holder of the instrument. The developers of this scale are known to comb the literature and ask those who used the scale for research to pay for a retroactive license which may cost thousands or tens of thousands of dollars, to add references to their work, or to remove details such as the actual instrument used from publications [1].

While it is certainly the prerogative of copyright holders of research instruments to enforce their rights, the Committee on Publication Ethics (COPE) has recently discussed the ethics of the behavior of certain copyright holders who "hold authors to ransom", and recommends that affected journals emphasize "the fact that this is not good for the advancement of scientific knowledge or in the public interest" [2]. As open access and open science publisher, JMIR Publications could not agree more

and we remind our authors of our policies and preference for public and free availability of research tools, including questionnaires [3], in the interest of reproducability and transparency of research. Given the flurry of legal disputes and correction notices related to MMAS (a third one affecting a JMIR journal is forthcoming), we now actively discourage use of MMAS and other instruments which are not available under a Creative Commons Attribution license, and encourage our authors to use or develop/validate new instruments which can be freely reproduced.

We are also hereby issuing a special call for papers for short paper instruments or electronic tools licensed under Creative Commons (or available under an Open Source license) that can be used instead of MMAS to measure medication adherence, and will waive the article submission fee for such development and validation papers describing new instruments that can be used as a free alternative to MMAS.

References

1. Marcus A. Pay up or retract? Survey creator's demands for money rile some health researchers. Science 2017 Sep 13;15(3):169-182. [doi: 10.1126/science.aap9445]
2. Committee on Publication Ethics (COPE). License for a Published Scale. 2018. URL: https://publicationethics.org/case/licence-published-scale [accessed 2018-11-29]
3. JMIR Publications. What is your policy regarding access to critical research tools and instruments (eg, questionnaires)?. 2018. URL: https://jmir.zendesk.com/hc/en-us/articles/360000547811 [accessed 2018-11-29]

Edited by G Eysenbach; submitted 24.09.18; this is a non–peer-reviewed article;accepted 04.10.18; published 20.12.18.

Please cite as:
Sanguansak T, Morley KE, Morley MG, Thinkhamrop K, Thuanman J, Agarwal I
Multimedia Appendix Removal and Editorial Warning Regarding the MMAS Scale (Two-Way Social Media Messaging in Post-Operative Cataract Surgical Patients: A Prospective Interventional Study)
J Med Internet Res 2018;20(12):e12120
URL: http://www.jmir.org/2018/12/e12120/
doi:10.2196/12120
PMID:30578176

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org